

Essays on Rural-to-Urban Migration and Urban Industrial Performance in Sub-Saharan Africa

Yuya Kudo

St Cross College, University of Oxford

Submitted in partial fulfillment of
the degree of D.Phil. in Economics

Hilary Term, 2010-2011

Essays on Rural-to-Urban Migration and Urban Industrial Performance in Sub-Saharan Africa

Yuya Kudo

St Cross College, University of Oxford

Submitted in partial fulfillment of the degree of D.Phil. in Economics

Hilary Term, 2010-2011

Abstract

This thesis consists of three independent but thematically related papers exploring the income determination process in African labour markets from spatial and sectoral perspectives.

Using long-run household panel data from rural Tanzania, chapter 2 investigates the extent to which education can explain migrants' income and consumption gains. We expect that the higher return to schooling at the destination primarily drives migrants' gains, suggesting that those who cannot afford the cost of schooling cannot reap the benefits of migration. We find that education indeed plays the role, but that it does not appear to be a major factor in limiting the internal migration as a source of raising income and consumption.

Exploiting data drawn from urban household panel surveys in Ghana and Tanzania, chapter 3 investigates how rural-to-urban migrants' earnings compare with those of natives in urban labour markets. The chapter attempts to identify the growth of migrants' earnings at the destination (assimilation), making a distinction between wage and self-employed migrants. We find that wage-dependent migrants would achieve higher lifetime earnings if they entered a self-employed sector from their arrival, conditional on individuals' attributes and the varying returns to those attributes across urban residents. The evidence points towards the importance of capital constraints in a decision to start a business.

Using firm-level data of manufacturing and retailing from the Enterprise Surveys conducted in seven Sub-Saharan African countries, chapter 4 attempts to improve our understanding of enterprise performance in urban Africa by investigating three aspects of firms' productive structure: technology, total factor productivity (TFP), and firm size. We find that the technology is similar between sectors, that retailing firms are smaller and less capital intensive but not, on average, ones with lower TFP, and that TFP differences are primarily within sectors. All these findings might point towards the importance of factor prices in characterising the industrial structure in urban Africa (308 words).

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Francis Teal, for his exceptional guidance on my research projects. He has always been supportive and broadening my horizons. To me, working with him meant taking a series of advanced lectures on how to implement applied micro-econometric research. I also thank to participants at seminars in Oxford (chapters 3 and 4) and international conferences - CSAE Conference: Economic Development in Africa 2010 in Oxford (chapter 3) and 3rd International Conference: Migration & Development 2010 in Paris (chapter 2) for helpful comments. I do appreciate Christopher Heady and Ian Crawford providing an assessment of the thesis as examiners. I am also grateful to the donors of Oxford Kobe Scholarship for giving me an opportunity to continue my research. Without the financial assistance, I could not have completed my projects. Thanks finally go to my family and friends. Especially, I would like to express my sincere gratitude to my parents, Mutsuo and Junko, and my wife, Tamaki, for their unceasing support and genuine affection.

Contents

1	Introduction	1
2	Migration, Educational Attainment and Welfare Growth: Evidence from Rural Tanzania	5
2.1	Introduction	6
2.2	Conceptual Framework	11
2.2.1	Migration for Schooling	11
2.2.2	Migration of Educational Attainers	17
2.2.3	Summary of the Model	18
2.3	Empirical Strategy	19
2.3.1	IV Approach: Structural Modelling Approach	21
2.3.2	Control Function (CF) Approach	25
2.4	Data	27
2.4.1	Migration	29
2.4.2	Occupational Mobility	31
2.4.3	Income Measurement Issues	31
2.4.4	Welfare Changes	34
2.5	Validating the Theoretical Model	35
2.6	Estimation Results	37
2.6.1	IV Approach: Structural Modelling Approach	37
2.6.2	CF Approach	42
2.6.3	Summary of the Estimation Results	46
2.7	Conclusion	47
2.8	Appendix: Model in Section 2.2	50
2.8.1	Proof	50
2.8.2	Comparative Statics of the Model	50
2.8.3	Some Thoughts on Modelling Migration	52
2.9	Appendix: Impact Evaluation Approach	53
2.9.1	Literature Review	53
2.9.2	Structural Interpretation of Impact Evaluation Approach	55
2.9.3	Estimation Results	57
2.10	Appendix: Income Measurement Issues	58
2.10.1	Suggestive Evidence on Latent Income	59
2.10.2	Empirical Test for Latent Income	61
2.10.3	Per Worker Earnings Functions of Each Activity	67
2.11	Appendix: More on Descriptive Statistics	68
2.11.1	Occupational Mobility	68
2.11.2	Hours Worked, Hourly Income, and Occupational Disaggregation	69
2.12	Appendix: Descriptions of Welfare Variables	70

3	Heterogeneous Population and Labour Market Outcome in Urban Africa: Mi- grants' Earnings in Ghana and Tanzania	103
3.1	Introduction	104
3.2	Review on Assimilation Literature	107
3.2.1	Theoretical Explanations	107
3.2.2	Empirical Model	112
3.2.3	Econometric Concerns	114
3.3	Empirical Strategy	119
3.3.1	Specification	119
3.3.2	Identification	123
3.4	Data	125
3.5	Estimation Results	128
3.5.1	Main Results	128
3.5.2	Robustness Checks	130
3.5.3	Interpreting the Results	133
3.6	Conclusion	133
3.7	Appendix: Model in Section 3.2.1(ii)	135
3.7.1	Derivation of equation (3.3)	135
3.7.2	Extending the Model: Influx of Skilled Immigrants and Occupational Upgrading	135
3.8	Appendix: Recent Arrivals Artificially Created by Repeated Migration	137
3.8.1	Model	137
3.8.2	Proof	139
3.8.3	Derivation of equation (3.30)	140
3.8.4	Ability Group and Destination Experience in Each Period	143
4	Technology, Productivity and Size in African Manufacturing and Retailing Firms: Similar or Different?	160
4.1	Introduction	160
4.2	Data	162
4.2.1	Data Source and Structure	162
4.2.2	Measurement Issues	164
4.2.3	Descriptive Summary	166
4.3	Modelling Technology	168
4.3.1	Model and Baseline Results	168
4.3.2	More on Panel Manufacturers	170
4.3.3	More on Retailers	175
4.3.4	Summary of Technology	176
4.4	Productivity and Size	176
4.4.1	Are Retailing Firms Less Productive?	176
4.4.2	Are Retailing Firms Smaller?	179
4.4.3	How are Productivities Distributed in Each Sector?	180
4.5	Conclusion	185
4.6	Appendix: Literature Review on Productivity and Size in African Manufacturing	188
4.7	Appendix: Levinsohn and Petrin Estimations for Panel Manufacturers	191
4.8	Appendix: Labour Productivity Specification	192
4.9	Appendix: Sensitivity of Productivity Analysis	193
5	Conclusion	223

List of Figures

2.1	Optimal Choice of Migrants	16
2.2	Transtion of Activity Shares in Total Hours Worked	74
3.1	The Estimate of Assimilation Effect Based on Hypothetical Cross-Sectional Data in 2010	115
3.2	Artificially Created Recent Arrivals	119
3.3	Histogram of Reasoning Ability and Cognitive Skills in Ghana	145
3.4	Histogram of Reasoning Ability and Cognitive Skills in Tanzania	145
3.5	Appendix: Simulated Earnings of Wage-employed Migrants with an Arrival Age of 20 Relative to Otherwise Comparable Wage-employed Natives	154
3.6	Appendix: Transition of Rural Population	159
4.1	Empirical Cumulative Distribution Functions of (Demeaned) TFP	206
4.2	Empirical Cumulative Distribution Functions of (Demeaned) Labour Productivity	206
4.3	Histogram of TFP Efficiency Score Based upon Gross Output	208
4.4	Histogram of TFP Efficiency Score Based upon Value Added	208
4.5	Appendix: Empirical Cumulative Distribution Functions of (Demeaned) TFP (Exclude Malawi and Zambia)	220
4.6	Appendix: Empirical Cumulative Distribution Functions of (Demeaned) Labour Productivity (Exclude Malawi and Zambia)	220
4.7	Appendix: Histogram of TFP Efficiency Score Based upon Gross Output (Exclude Malawi and Zambia)	222
4.8	Appendix: Histogram of TFP Efficiency Score Based upon Value Added (Exclude Malawi and Zambia)	222

List of Tables

2.1	Migrants by Origin	72
2.2	Migrants by Age at Migration	72
2.3	Migrants by Year of Migration	72
2.4	Reason for Migration	73
2.5	Summary of Key Variables - Individual Level	75
2.6	Summary of Key Variables - Household Level	76
2.7	Migration and Educational Attainment by Age in Wave 1	77
2.8	Reason for Migration by Educational Attainment	78
2.9	Informative Evidence on Exclusion Restrictions	79
2.10	Reduced Form Equations: Migration and Schooling	80
2.11	IV Approach: Structural Estimation Results	81
2.12	Control Function (CF) Approach: Earnings	82
2.13	Control Function (CF) Approach: Consumption	83
2.14	Appendix: Checking the Correlation between the Pre-migration Level of Schooling and Wealth	84
2.15	Appendix: 2SLS Estimations on Earnings	85
2.16	Appendix: Impact Evaluation Results	86
2.17	Appendix: Detailed Descriptive Statistics - Individual Level	87
2.18	Appendix: Detailed Descriptive Statistics - Household Level	89
2.19	Appendix: Per Worker Weekly Income, Weekly Hours Worked, and Hourly Income by Occupation	92
2.20	Appendix: Diversification of Activities - Individual Level (Aged 23 to 40)	95
2.21	Appendix: Diversification of Activities - Household Level	96
2.22	Appendix: Correlation between Income and Consumption (OLS)	97
2.23	Appendix: Empirical Test for Latent Income in Wave 5	98
2.24	Appendix: Per Worker Wage Earnings Functions (OLS)	99
2.25	Appendix: Per Worker Farming Self-employment Income Functions (OLS)	100
2.26	Appendix: Per Worker Non-farm Self-employment Income Functions (OLS)	100
2.27	Appendix: Transition of Activity Shares by Age in Wave 1	101
2.28	Appendix: Correlation of the Transition of Activity Shares with Educational Attainment by Age in Wave 1	102
3.1	Urban Population Data	144
3.2	Descriptive Statistics - Ghana	146
3.3	Descriptive Statistics - Tanzania	147
3.4	Main Estimation Results (OLS)	148
3.5	The Estimated π_1 in the Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS)	149
3.6	Appendix: Main Estimation Results (OLS) - Continuous Measure: Ghana	150

3.7	Appendix: Main Estimation Results (OLS) - Continuous Measure: Tanzania . . .	151
3.8	Appendix: Main Estimation Results (OLS) - Discrete Measure: Ghana	152
3.9	Appendix: Main Estimation Results (OLS) - Discrete Measure: Tanzania	153
3.10	Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Ghana (1)	155
3.11	Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Ghana (2)	156
3.12	Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Tanzania (1)	157
3.13	Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Tanzania (2)	158
3.14	Appendix: Migrants' Urban Experience (years)	159
4.1	Industrial Structure and Labour Productivity	194
4.2	Sample Composition	195
4.3	Descriptive Statistics	196
4.4	Baseline Production Functions and Size Regressions (OLS)	198
4.5	Production Functions of Panel Manufacturers	199
4.6	Production Functions of Retailers	200
4.7	TFP Comparison - Pooling Regressions (OLS)	201
4.8	Top Manager's Characteristics	202
4.9	The Largest Owner's Characteristics	202
4.10	Match Quality between Firms and Workers	203
4.11	The Biggest Business Obstacle	203
4.12	Size Comparison - Pooling Regressions (OLS)	204
4.13	Sensitivity Check of (Demeaned) TFP Distributions across Estimations - Manufac- turing	205
4.14	Sensitivity Check of (Demeaned) TFP Distributions across Estimations - Retail . .	205
4.15	Test for the Equality of (Demeaned) Productivity Distributions between Sectors .	207
4.16	TFP Efficiency Scores	207
4.17	Appendix: Transition of GDP Share by Sector (%)	209
4.18	Appendix: Full Regression Results of Table 4.4(OLS)	210
4.19	Appendix: Full Regression Results of Table 4.5	211
4.20	Appendix: Marginal Effects in Table 4.7	212
4.21	Appendix: Levinsohn and Petrin Estimation Results	213
4.22	Appendix: Baseline Production Functions and Size Regressions (OLS) - Labour Productivity Specification	214
4.23	Appendix: Production Functions of Panel Manufacturers - Labour Productivity Specification	215
4.24	Appendix: TFP Comparison - Pooling Regressions (OLS) - Labour Productivity Specification	216
4.25	Appendix: TFP Comparison - Pooling Regressions (OLS) (Exclude Malawi and Zambia)	217
4.26	Appendix: Marginal Effects in Table 4.25	218
4.27	Appendix: Sensitivity Check of (Demeaned) TFP Distributions across Estimations - Manufacturing (Exclude Malawi and Zambia)	219
4.28	Appendix: Sensitivity Check of TFP Distributions across Estimations - Retail (Ex- clude Malawi and Zambia)	219

4.29	Appendix: Test for the Equality of (Demeaned) Productivity Distributions (Exclude Malawi and Zambia)	221
4.30	Appendix: Sensitivity Check of TFP Efficiency Scores (Exclude Malawi and Zambia)	221

Chapter 1

Introduction

World history shows that economic growth is associated with two types of important economic transformations: sectoral transformation and spatial transformation. Starting from a traditional agricultural economy, an industrial sector firstly takes over the primacy as people's income rises (industrialisation). This process normally involves technological progress in the agricultural sector which makes that sector more labour productive than before, people's demand shift to more industrial goods, and workers' mobility from the agricultural to the industrial sector. Once people's demand for industrial goods becomes saturated, they begin to want less material goods (i.e., services), which in turn generate much more employment opportunities in a service sector. Since service workers are not simply substitutable by machines and further technological progress makes both agricultural and industrial sectors more labour productive, the service sector continues to absorb the country's labour force and eventually becomes a leading sector of the economy (post-industrialisation). In contrast to the shift in the production structure, less clear but equally important is the spatial transformation. As fertile plain attracts farming, coastal areas are suited for export-industries and urban agglomeration contributes to creating knowledge spillovers in high-tech service industries, each industry has a preferable place to do. Thus, the structural shift in the economy is usually followed by the spatial transformation, which shapes the pattern of population distribution. Some researchers now agree that spatially efficient economic geography is as important for economic growth as technological advances and the sectoral transformation.

In contrast to the world history, many African countries appear to reveal a somewhat different picture: urbanisation without industrialisation. Then, whether or not the economic mobility observed in Africa is a part of the growth process of the economy is a fundamental question. The thesis implicitly attempts to improve our understanding of this question by exploring the income

determination process in labour markets from these spatial and sectoral aspects.

The thesis consists of three independent but thematically related core chapters. The first two chapters (chapters 2 and 3), with an emphasis on the spatial perspective, explores labour market outcomes associated with internal migration. Specifically, chapter 2 investigates migrants' labour market outcomes from an angle of rural economy by comparing migrants with non-migrants, whereas chapter 3 does so from an angle of urban labour market by comparing migrants with urban natives. The third chapter (chapter 4) focuses on the sectoral perspective, and attempts to understand the income determination process from an angle of enterprise performance in urban labour markets. Specifically, the chapter explores the urban industrial structure by comparing manufacturing firms with retailing. All chapters are empirical research, and use micro-level data drawn from either household or enterprise surveys.

It has been around thirty years since Stark (1982) reviewed the past studies on rural-to-urban migration and advocated the importance of “effectively manipulating the phenomenon so as to turn it into a vehicle of national development and personal betterment.” Whilst early research on labour migration, which was preoccupied with urban unemployment, social tensions and squalor in the fast-growing cities, has led governments to institute migration abatement policies, some economists now acknowledge the benefits that labour migration (or agglomeration economies, more generally) create, and are in agreement to that the past restrictions on labour mobility (whether direct or indirect) imposed the costs of foregone opportunities for economic growth and convergence in living standards. Accordingly, international policy debates have been shifting from how to discourage labour mobility into how to effectively create and reap the benefits of agglomeration and how to integrate leading and lagging areas in a country, while managing undesirable agglomeration outcomes (e.g., congestion, pollution, urban slums), reducing the widening or persistent differences in living standards across regions, and avoiding unnecessary intervention failures (World Bank, 2009).

Since policy imperatives change as urbanisation (or agglomeration) advances, having a different set of policy instruments for a different level of urbanisation is important to achieve those goals. However, one lesson from historical experience is that spatially targeted policy efforts (e.g., slum-upgrading programme, public housing programme, etc.) may not bear fruit unless economies maintain sound and credible institutions, such as the provision of basic social services (education,

health care, etc), the security of property rights, and the efficient operation of fluid land markets (spatially blind institutions - World Bank, 2009). This suggests that one fundamental research to improve our understanding of the pattern of spatial transformation and link it with an effective policy implementation is to evaluate the interaction between human mobility and these institutions set by governments.

Chapter 2 focuses on the role of education, one of these institutions. A traditional human capital theory of migration studies regard migration as an investment in human capital because it places workers from one place to another where the return to their human capital is higher at the cost of movement. Based on this view, we expect that the higher return to education at the destination primarily drives migration, eventually benefiting (educated) migrants. If people move just to take advantage of the higher return, we would find no gains from migration in an outcome equation (income or consumption), after controlling for the current level of education and allowing for the difference in the return to schooling between the origin and the destination. Then, this finding suggests that those who cannot afford the cost of schooling (as well as the cost of migration) cannot reap the benefits of migration. Using long-run household panel data from rural Tanzania, Kagera, from 1991 to 2004, chapter 2 investigates the extent to which education can explain migrants' income and consumption gains. By exploring the role of education, we equivalently test the extent to which an inability to have an access to education limits the internal mobility as a source of raising income and consumption.

A policy debate in favour of restricting rural-to-urban migration often results from the image of overcrowded cities, the concentration of urban poverty associated with emerging shantytowns, and appalling environmental degradation. Nevertheless, how internal mobility is affecting the structure of urban labour markets in less developed economies is not well understood due to lack of good migration and labour force data. Chapter 3 sheds light on the point by asking the question of how rural-to-urban migrants' earnings compare with those of natives in urban labour markets in Africa. The human capital theory of migration studies implies that migrants earn less than otherwise comparable natives on their arrival, but they eventually achieve earnings parity after their adjustment period at the destination (assimilation). Using pooled data drawn from three-round household panel surveys conducted in urban Ghana and Tanzania in 2003/2004, 2005, and 2006, chapter 3 attempts to identify the transition process by estimating an earnings function.

The novelty in the chapter is to make a distinction between wage-dependent and self-employed migrants. We think that this distinction is important because it seems that migrating to start a business is more costly and needs stronger determination than migrating for wage jobs, thus the two types of mobility are fundamentally different processes. Interestingly, the analysis reveals the difference in migrants' earnings profile at the destination between sectors, pointing to the importance of capital constraints in a decision to start a business.

The shift in the production structure of the economy can be viewed as a process in which the inter-sectoral relationship of economic activities changes over time. Therefore, how sectoral activities are interacting with each other is an important question to understand the process of economic development. Chapter 4 sheds light on this point by exploring the sectoral difference in enterprise performance in urban Africa. Using micro firm-level data of manufacturing and retailing drawn from the Enterprise Surveys conducted in seven Sub-Saharan African countries: Burkina Faso (2006, 2009), Cameroon (2006, 2009), Malawi (2005, 2009), Mali (2003, 2007), Niger (2005, 2009), Senegal (2003, 2007) and Zambia (2002, 2007), chapter 4 investigates three aspects of firms' productive structure: technology, total factor productivity (TFP), and firm size. In the countries of our interest, non-manufacturing constitutes a main part of the urban economy in terms of both GDP and job creation. Nevertheless, the previous studies of African firms have disproportionally put an emphasis on manufacturing. By choosing one type of non-manufacturing business, retailing, and comparing it with manufacturing, the chapter opens the door to a new research of African firms.

Chapter 2

Migration, Educational Attainment and Welfare Growth: Evidence from Rural Tanzania¹

Abstract

Education is expected to play a major role in explaining the gains from migration because it may be more valuable at the destination. This chapter constructs a simple rural-to-urban migration model. In the model, urban jobs reward migrants for their education, and the higher return to schooling gives rural agents an incentive to migrate, encouraging the less educated to invest in further schooling before joining labour force, conditional on the initially endowed level of education. Depending upon where learning opportunities are available, some migrate to seek more education, and others migrate after obtaining it and the investment process of migration and schooling eventually leads migrants into higher earnings. The model has important implications for an empirical analysis. If the migration effect on an outcome (earnings or consumption) is entirely attributed to the higher return to schooling at the destination, the effect will disappear by controlling for the current level of education and allowing for the difference in the return between the origin and the destination in the outcome equation. This suggests that those who cannot afford the cost of schooling (as well as the cost of migration) cannot reap the benefits of migration. On the other hand, if the migration effect remains in the presence of a full set of controls, then, the evidence points to the migration effect operating other than through education. Using long-run household panel data from rural Tanzania, Kagera from 1991 to 2004, the chapter investigates the extent to which education can explain migrants' income and consumption gains by estimating those outcome equations. Empirical difficulty arises from the endogeneity of migration and schooling decisions. To address the endogeneity, the chapter takes two approaches: an IV approach (structural modelling approach) and a control function (CF) approach. Controlling for initial household fixed effects, we find that individuals are positively selected into both migration and further schooling based on their pre-migration level of education. In consumption, we provide evidence that the return to schooling is indeed higher at the destination, but that a large significant migration effect remains after controlling for education and allowing for the endogenous selection into migration and schooling. Evidence on earnings is rather ambiguous due to the difficulty in controlling for the endogeneity, but the analysis still provides some evidence to reveal a similar story to consumption. These suggest that income and consumption opportunities are higher at the destination, and that some mechanisms, operating other than through education, prevent individuals from exploiting the arbitrage opportunity. While education plays a role, it does not appear to be a major factor in limiting the internal migration as a source of raising income and consumption.

¹ A version of this chapter was presented at the 3rd International Conference: Migration & Development in Paris (10th-11th September, 2010). I am grateful to conference organizers (Paris School of Economics, the French Development Agency Research Department and the World Bank) and participants for helpful comments.

2.1 Introduction

Impact evaluation has been a rapidly growing field of economics since Lalonde (1986) used the terminology and reasoning of medical experiments to estimate the causal effect of an employment programme in the United States. More recently, economists have applied the framework to identify the gains from migration, and the literature has found that migration has a significant impact on income and consumption (McKenzie et al., 2006; Beegle et al., 2008b).

To identify the gains from migration, the impact evaluation approach normally regresses the outcomes of interest (income, consumption, etc.) on a migration dummy, as well as on several controls. In this approach, the controls need to be the attributes of a migration decision and of an outcome process unaffected by migration itself. Researchers often select pre-treatment variables for the controls, such as schooling and age at baseline, hoping that migration is statistically independent of the outcomes conditional on those controls (selection on observables or ignorability of treatment). While this approach has been emerging as a dominant way of drawing causal inference about economic relationships, identifying the gains from migration in this way leads us to the puzzling questions of why not all people move out if the return to migration is so high, and what limits the mobility of non-migrants. This is because the approach does not tell us anything about the underlying mechanisms that benefit migrants. Finding answers for these questions is very important because migration shapes the pattern of economic development by influencing welfare distribution across regions.

In the current study, we shift to an attempt to explore the underlying mechanisms away from an attempt just to identify the gains from migration. Specifically, we investigate the extent to which education can explain the gains from migration. We expect that a fundamental mechanism that drives migration is the higher return to schooling at the destination. If people move just to take advantage of the higher return, no migration effect on outcomes would exist, after controlling for education and allowing for the difference in the return between the origin and the destination. Then, this suggests that those who cannot afford the cost of schooling (as well as the cost of migration) cannot reap the benefits of migration. This hypothesis is tested in this chapter, and the idea essentially encourages us to put the current level of education in the controls in outcome equations, rather than the pre-migration level of schooling, which is almost automatically used

in the impact evaluation approach. In this sense, we fundamentally estimate the same outcome equation as in the impact evaluation approach, but use a different set of controls.

The idea that people relocate to a place in which the return to their education (more generally, human capital) is higher at the cost of movement is not new. In his paper on human capital formation, Schultz (1961) enumerated internal migration as a form of human investment that enhanced the productivity of human resources.² Similarly, Sjaastad (1962) treated migration as an investment increasing the market value of one's productive capacities and started the tradition of analyzing migration in a cost-return framework, which led into the burgeoning empirical literature in early days investigating economic gains from migration (e.g., Wertheimer, 1970; Yezer and Thurston, 1976; Grant and Vanderkamp, 1980) and an individual's response to those gains (e.g., Bowles, 1970; Laber and Chase, 1971). Despite this traditional view of human capital investment dating back to the 1960s, surprisingly, it appears that no literature has attempted to identify the difference in the return to human capital between migrants' origin and destination. Therefore, the current study can be seen as the first empirical attempt with an emphasis on the role of education.

The question of why people are willing/unwilling to migrate is not new, either. By Todaro (1969), Harris and Todaro (HT) (1970), and their followers (e.g., Johnson, 1971; Bhagwati and Srinivasan, 1974; Fields, 1975; Todaro, 1976; Cole and Sanders, 1985. See also Bhattacharya (1993) for an extensive review), the human capital theory of migration was extended to provide theoretical explanations for the significant level of urban unemployment and underemployment associated with internal migration and discussions about policy instruments to improve the social welfare. In these migration models, urban-rural differences in expected earnings primarily drive (risk-neutral) agents to relocate to urban areas. In contrast to the HT-type models, assuming migration as an activity undertaken by rural agents individually, the subsequent research moved away from the assumption with an emphasis on the role of family members in a decision to migrate (e.g., Mincer, 1978; Banerjee, 1981). For example, Mincer (1978) formalised how net family gains from migration compare with the net personal gains of both husband and wife, with several mobility outcomes including marriage dissolution. This family aspect in a mobility decision, together with rural households' income risk and risk-avoidance behavior, attracted the attention of economists

²The other major forms of investment that he enumerated included health expenditures, on-the-job training, formal education, and extension programmes in agriculture.

further and led them to view internal migration as a family's strategy to diversify income risks (e.g., Benerjee and Kanbur, 1981; Stark, 1981; Stark and Levhari, 1982; Katz and Stark, 1986; Rosenzweig and Stark, 1989; Daveri and Faini, 1999; Chen et al., 2003). The general argument in these papers is that to mitigate the risk of their variable home income, rural households unable to have access to credit and insurance markets might send their members to locations in which the income risk is negatively correlated with that in their home areas. On top of these (expected) income and risk explanations, Stark and Taylor (1991) studied the impact of relative deprivation in migrants' reference group (e.g., original village) on their migration decisions. Based on their argument, since migrants' host communities seem to replace their original communities relatively easily as their reference group (reference group substitution) in internal migration, the increased alienation arising from the reference group substitution might reduce the propensity of relatively deprived households to engage in internal migration, even though positive absolute income gains from moving out exists. In parallel with these studies, the literature has also accumulated the knowledge about the role of migrants' social network (kinship, co-villagers, etc.), discussing that it affects their decision to relocate through its impact on the cost of migration (e.g., the costs of movement, temporary accommodation, and job search) (Benerjee, 1983; Andersson, 2001; Munshi, 2003; McKenzie and Rapoport, 2007; Yamauchi and Tanabe, 2008).

Some of these studies can possibly provide alternative explanations for the question of why not all people move out in response to the higher return to migration. In the HT-type models, for example, since urban (fixed) wage institutionally set above the market-clearing level is a source of raising migrants' income and wages are equalized in terms of *expectation* between migrants' origin and destination, it is possible that a part of rural agents remains in their origin despite the *actual* higher return to movement. In relation to the way of forming the expectation, to the extent that imperfect information concerning labour market conditions in the destination leads rural agents to attach a marginal weight on earnings attainable from urban occupations, only those who have better access to the information (e.g., those who have a social network) might be able to attempt a venture of migration and reap the benefits. Rural households' risk-coping strategy may also account for the differing pattern of spatial mobility. Since the strategy essentially prevents all family members from relocating to the same place, only some members might consequently migrate to earn higher income (for some reasons, which may be due to education along the line of

the current study or institutional urban wage along the line of the HT-type models), leaving the other members in original villages. Acknowledging these alternatives, we view the current study as a good place to begin, with the examination of the role of education, to find an answer for the big question.

In contrast to the impact evaluation approach, much greater empirical difficulty arises in the identification of our study. This is because the expected future earnings (thus, more or less, realized earnings) can drive the decisions to invest in migration and schooling. In this chapter, we take two approaches to address the endogeneity: an IV approach and a control function (CF) approach.

To motivate an empirical analysis, the chapter starts with developing a simple rural-to-urban migration model, which lays its foundation on the income maximization of individual agents, rather than families, but is explicit about the role of the return to schooling as a source of raising income. In the model, migrants are rewarded for their education in an urban sector, and the higher return to schooling in cities incentivizes rural agents to migrate, encouraging the less educated to invest in further schooling before participating in labour markets, conditional on the initially endowed level of education. Depending upon where education opportunities are available, some migrate to seek more education and others migrate after obtaining it. The investment process of migration and schooling enables migrants to eventually achieve higher earnings. The model provides a direct implication for our hypothesis, making several equilibrium predictions broadly consistent with data and giving us a guide for empirical modelling in the IV approach.

The IV approach addresses the endogenous selection into migration and schooling by extending the theoretical model into an estimatable structural model of simultaneous equations and fitting it to our data. The drawback of the approach is that it does not allow the return to schooling to differ between the origin and the destination in an outcome equation. On the other hand, the CF approach can allow for it. However, it addresses the endogeneity of migration and schooling decisions non-structurally by making strong parametric assumptions about the correlation of an error term in an outcome equation with that in migration and schooling equations. Neither approach is perfect, but both approaches point towards the examination of our hypothesis: the migration effect is the schooling effect.

Data used in the analysis are drawn from a five-wave long-run household panel survey conducted in the rural region in northwest Tanzania, Kagera (Kagera Health and Development Survey,

KHDS). The first four waves were carried out between 1991 and 1994 with 6- or 7-month intervals, with the final wave carried out in 2004. As approximately 35% of initial respondents moved out of their original communities during the sample periods, the data contain a good variation of migrants and non-migrants for the purpose of the current study.³ Examining the process of migration taking place in Africa is meaningful because Africa recorded the highest annual rate of growth of urban population in the world between 1975 and 2007 (United Nations, 2008).⁴ Using data drawn from the KHDS is helpful for our purpose because we can directly follow Beegle et al. (2008b), in which the impact evaluation approach was applied to data drawn from the KHDS to identify the gains from migration, although the data used in this chapter are my construction.⁵⁶ By pooling migrants into one group and estimating a differences-in-differences (DID) specification by OLS and 2SLS, they found that moving out of the original community added 36 percentage points to consumption growth from wave 1 to 5 on average, and provided suggestive evidence on the question of how migrants achieved it by focusing on the sectoral choice of individuals (agriculture vs. non-agriculture).⁷ In their findings, the migration gain is higher for those who simultaneously moved out of agricultural jobs.

Our study makes two key contributions to the literature. As opposed to Beegle et al. (2008b), we firstly create the earnings of individuals in addition to consumption. Secondly, developing a theoretical model and, in the case of the IV approach, structurally addressing an endogenous selection into migration and schooling provides an analytical framework that helps us to understand the role that education plays in the process of migration. By contrasting the two welfare measures and applying the framework to actual data drawn from less developed economies, our research gives us a great insight into what is going on under the process of internal migration associated with economic development.

³In addition, almost half the initial respondents moved out of their original households.

⁴By 2050, Africa is projected to be second to Asia as a region to accommodate a large mass of urban inhabitants (United Nations, 2008).

⁵However, we were not aware of their study when we started our research.

⁶Other studies exploiting data collected in the KHDS also exist (e.g., Beegle, 2005; Beegle et al., 2008a). For example, Beegle (2005) used the data drawn from the first four waves to investigate how prime-age (15-50) adult mortality (past and future) affects the time allocation of surviving household members. The influence on the portfolio of farming activities (cash crops and food crops) and income diversification strategies (wage employment and non-farm self-employment) was examined, together with the impact on time engaged in farming and chore activities. In a similar vein, the impact of adult mortality shocks on the consumption growth of surviving household members was investigated, using all five-wave data (Beegle et al., 2008a). They found a 7% drop in consumption growth within the first 5 years after death, with suggestive evidence on the persistent effect of shocks, although the persistency from the distant past was debatable. The impact of female adult death and the impact on a close blood relative were found to be severe.

⁷In Appendix 2.9.3, we attempt to recover their findings by applying their strategy to our measures of consumption and earnings.

The chapter is organized in seven sections. In Section 2.2, we develop a simple rural-to-urban migration model to motivate the subsequent empirical analysis. Section 2.3 describes our empirical strategy. Section 2.4 presents an overview of data. In Section 2.5, we briefly check if our data bear several assumptions and equilibrium predictions drawn from the theoretical model. Section 2.6 reports the main estimation results for our hypothesis, with the conclusion in Section 2.7.

2.2 Conceptual Framework

This section presents a simple rural-to-urban migration model to motivate the subsequent empirical analysis. While the model is static and non-stochastic, it provides a fundamental mechanism to explain the relationship between migration, educational attainment, occupational choice, and earnings.

2.2.1 Migration for Schooling

Environment

Consider a rural economy (origin) in which a large population (a continuum) of risk-neutral agents with identical preferences exists. The population is characterized by a distribution function $G(\bar{s})$, which provides the measure of the population with the level of schooling at the origin less than $\bar{s}$.⁸ We assume that $\bar{s}$ is distributed in the interval $[0, \bar{s}_u]$. Agents decide whether to migrate into an urban area (destination) or not. If they prefer to stay in the rural sector, they simply choose occupations. On the other hand, the urban economy gives them opportunities to obtain a higher education before joining the labour force. Since the model is static, this implies migrants simultaneously choose the level of education $s \in R_+$, as well as occupations, once they migrate.

Occupational Choice and Educational Attainment

Gross income from occupations is defined as $2\alpha q + \delta$, where $q \in R_+$ describes occupations. We assume that $\alpha > 0$ and $\delta > 0$. The positive α is a crucial assumption in the model and we will see its meaning in a minute. While only one occupation ($q = 0$, self-employed farming) is available in the rural sector, heterogeneous jobs exist in the urban economy ($q \in R_{++}$). This specification

⁸In the context of wealth distribution, similar settings can be seen in Banerjee and Newman (1993) and Ghatak and Jiang (2002).

ensures that gross income from urban jobs increases as q increases. We might be able to interpret a lower q as non-farm self-employed petty trading and higher q as non-farm formal wage jobs. Alternatively, a higher q might suggest that the job pays workers more since it is a complex task that requires a higher level of skills or technologies. Any urban job requires agents to have the skills (measured by education, s) that suit it. Thus, if agents do not have the best skills for their occupations, they lose income. This monetary cost is $\alpha(q - s)^2$. With these assumptions, agents' labour income y can be written as

$$y = \begin{cases} y_r = \delta & \text{if agents do not migrate,} \\ y_u = \alpha [2q - (q - s)^2] + \delta & \text{if agents migrate.} \end{cases} \quad (2.1)$$

As already mentioned in the previous sub-section, migrants have opportunities to obtain a higher education. In that case, migrants with the initial level of schooling $\bar{s}$ have to pay the difference between the cost (e.g., tuition fees) required to achieve the target level of education (s) and the cost already incurred to obtain the current level of education ($\bar{s}$). Assuming that the cost of education is convex, we define it as $\beta(s^2 - \bar{s}^2)$, where $\beta > 0$. In this model, we assume away unemployment or withdrawal from the labour market for expositional simplicity. In addition, the opportunities for additional schooling are perfectly guaranteed as long as migrants seek for them, and there are no school dropouts in the economy.

Wealth, Cost of Migration and Credit Rationing

Each agent has his own wealth as endowments, which are expected to positively correlate with the initial level of schooling. This assumption is important because it enables the initial level of schooling to play the role of credit in the model. We define the initial level of wealth as $\gamma\bar{s} + \rho$, where $\gamma > 0$ and $\rho > 0$. We assume that the fixed cost of migration m is greater than ρ ($\rho < m$), which means that the agent with initially no schooling ($\bar{s} = 0$) is not able to migrate unless he obtains a loan.

Agents are able to afford the cost of migration and additional schooling through loan contracts with rural lenders before they migrate. However, there is the possibility that borrowers can renege on a debt because it is not easy for rural lenders to chase migrants in cities who have defaulted

on their contracts and force them to repay their debt. With the lack of alternative enforcement mechanism, this results in credit rationing: only agents with wealth greater than a certain level can obtain a loan. To incorporate credit rationing into our model, we follow Ghatak and Jiang (2002), where migrants who did not fulfil their contracts might be caught with some probability π and have to receive a nonmonetary punishment F , such as imprisonment or social sanctions. Denoting one plus the rate of interest as r , lenders will make only loans that satisfy

$$r[m + \beta(s^2 - \bar{s}^2) - \gamma\bar{s} - \rho] \leq \pi F. \quad (2.2)$$

Here, we assume that $r(m - \rho) > \pi F$. This says that no loan contract exists in the economy, which enables the agent with initially no schooling to migrate. In addition, throughout the section, we assume that the return to choosing urban jobs with additional schooling is sufficiently high in the sense that

$$\frac{m - \rho - \frac{\pi F}{r}}{\gamma} < \frac{\alpha}{\beta(1 + r)}. \quad (2.3)$$

Below, we set $\pi = 0$ without loss of generality. Then, migration will be a strategy only for those who are able to finance their migration and additional schooling with their own wealth.

Migrants' Occupation and Schooling Decision

In the urban economy, migrants choose their occupations and choose how much they acquire new skills, so their decision problem could be

$$\max_{\{q, s\}} \alpha [2q - (q - s)^2] + \delta - \beta(s^2 - \bar{s}^2) - m + r[\gamma\bar{s} + \rho - m - \beta(s^2 - \bar{s}^2)] \quad (2.4)$$

subject to $\Delta s \geq 0$ (non-negativity constraint of additional schooling, $\Delta s \equiv s - \bar{s}$) and $m + \beta(s^2 - \bar{s}^2) \leq \gamma\bar{s} + \rho$ (credit constraint).

Note that, given obtained human capital (s), an optimal occupation simply maximizes migrants' labour income, so $q^* = 1 + s^*$. Throughout Section 2.2, the superscript asterisk (*) refers to the optimal value chosen by each agent at equilibrium. Substituting this into (2.1), migrants' optimal

labour income becomes

$$y_u^* = \underbrace{\alpha \left(2 + \frac{1}{s^*} \right)}_{\text{Average return to schooling}} \times s^* + \delta. \quad (2.5)$$

Earlier, we assumed a positive α . In the model, the parameter α captures the return to schooling in an urban area, and the positive number implies that the return is higher in a city because it is assumed that the return to schooling is zero in a rural area in the model. If α is zero, migrants just obtain the same level of earnings as non-migrants, δ , and, thus, there is no incentive to migrate. In this sense, the higher return to schooling in an urban area drives rural-to-urban migration in the model.

The first order condition of migrants' schooling decision is

$$\underbrace{\frac{\partial y_u}{\partial s}}_{\text{MB of schooling}} = \underbrace{2\beta s}_{\text{MC of schooling (direct cost)}} + \underbrace{2\beta r s}_{\text{MC of schooling (opportunity cost)}} - \lambda_1 + 2\beta \lambda_2 s, \quad (2.6)$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are Lagrange multipliers for the non-negativity constraint of additional schooling and credit constraint, respectively. Marginal benefit and marginal cost are abbreviated as “MB” and “MC,” respectively. The marginal cost of schooling includes the opportunity cost of foregone interest income, as well as direct cost (e.g., tuition fees). Figure 2.1 precisely shows the entire path of the optimal value of occupation, schooling, and additional schooling of migrants, where $\hat{s} \equiv \frac{-\gamma + \sqrt{\gamma^2 - 4\beta(\rho - m) + (\frac{2\alpha}{1+r})^2}}{2\beta}$.⁹ When $\bar{s} \in [0, \frac{m-\rho}{\gamma})$, no one can afford the cost of migration, m ; therefore, the figure omits this group of people from the optimal path.

Suppose a migrant's initial level of schooling is in the interval $\bar{s} \in (\frac{m-\rho}{\gamma}, \hat{s}]$, the marginal benefit of schooling is greater than the marginal cost of schooling, $\frac{\partial y_u}{\partial s} > 2\beta s + 2\beta r s$. Thus, migrants are willing to obtain additional education, leading them to higher income jobs. However, since they do not have sufficient wealth to achieve the schooling level of their first best, they attempt to

⁹Optimal solutions of q^* , s^* , and Δs^* are

$$(q^*, s^*, \Delta s^*) = \begin{cases} (1 + \frac{m-\rho}{\gamma}, \frac{m-\rho}{\gamma}, 0) & \text{for } \bar{s} = \frac{m-\rho}{\gamma} \\ (1 + \sqrt{\frac{\beta \bar{s}^2 + \gamma \bar{s} + \rho - m}{\beta}}, \sqrt{\frac{\beta \bar{s}^2 + \gamma \bar{s} + \rho - m}{\beta}}, \sqrt{\frac{\beta \bar{s}^2 + \gamma \bar{s} + \rho - m}{\beta}} - \bar{s}) & \forall \bar{s} \in \left(\frac{m-\rho}{\gamma}, \hat{s} \right] \\ (1 + \frac{\alpha}{\beta(1+r)}, \frac{\alpha}{\beta(1+r)}, \frac{\alpha}{\beta(1+r)} - \bar{s}) & \forall \bar{s} \in \left(\hat{s}, \frac{\alpha}{\beta(1+r)} \right] \\ (1 + \bar{s}, \bar{s}, 0) & \forall \bar{s} \in \left(\frac{\alpha}{\beta(1+r)}, \infty \right) \end{cases}.$$

obtain as much additional education as possible ($\lambda_1 = 0, \lambda_2 > 0$). As a result, migrants acquire more new skills to get higher income jobs, as they are initially more educated (wealthier). In the interval $\bar{s} \in (\hat{s}, \frac{\alpha}{\beta(1+r)}]$, migrants are no longer credit constrained ($\lambda_1 = 0, \lambda_2 = 0$). Everybody whose initial level of schooling falls in this range acquires new skills to achieve the same level of education $\frac{\alpha}{\beta(1+r)}$ and to choose the same level of occupation $1 + \frac{\alpha}{\beta(1+r)}$, so that $\frac{\partial y_u}{\partial s} = 2\beta s + 2\beta r s$. Since the optimal level of schooling is the same between migrants, their educational attainment is smaller, as they are initially more educated. In the case where $\bar{s} \in (\frac{\alpha}{\beta(1+r)}, \infty)$, the marginal cost of attaining additional education is significantly large, $\frac{\partial y_u}{\partial s} < 2\beta s + 2\beta r s$ because they are already highly educated in the rural economy. Then, they optimally choose not to acquire new skills and select occupations in accordance with their initial level of schooling ($\lambda_1 > 0, \lambda_2 = 0$).

This diversity in the relationship between the initial level of schooling and the optimal behaviour avoids providing us with an unique answer for the question of whether it is the more (or less) educated in the rural economy who attain more additional education in cities, depending on where $\bar{s}_u$ is located in the distribution of the initial schooling. To the extent that $\frac{m-\rho}{\gamma} < \bar{s}_u \leq \hat{s}$, then it will be those in the upper tail of the distribution $\bar{s} \in (\frac{m-\rho}{\gamma}, \bar{s}_u]$ who acquire more new skills in the urban sector. To the contrary, if $\frac{\alpha}{\beta(1+r)} < \bar{s}_u$, those who are relatively educated at the origin $\bar{s} \in (\frac{\alpha}{\beta(1+r)}, \bar{s}_u]$ will be reluctant to attain additional education once they migrate. In the case where $\hat{s} < \bar{s}_u \leq \frac{\alpha}{\beta(1+r)}$, migrants who are initially in the intermediate level of schooling distribution are more likely to acquire new skills. All these possibilities, together with the migration decision explained in the next sub-section, potentially generate multiplicity in equilibria about who migrates and who chooses what level of occupations and education. While it might not be impossible to describe them comprehensively, doing so is out of the scope of this chapter. Thus, we restrict our attention to a certain type of equilibrium by hereinafter assuming that $\frac{m-\rho}{\gamma} < \bar{s}_u \leq \hat{s}$. However, this should not be interpreted as an artefact just for computational reasons. Rather, we should construe that this reflects the real situation in most less developed countries, where the highest level of education attainable in the rural sector is relatively low.

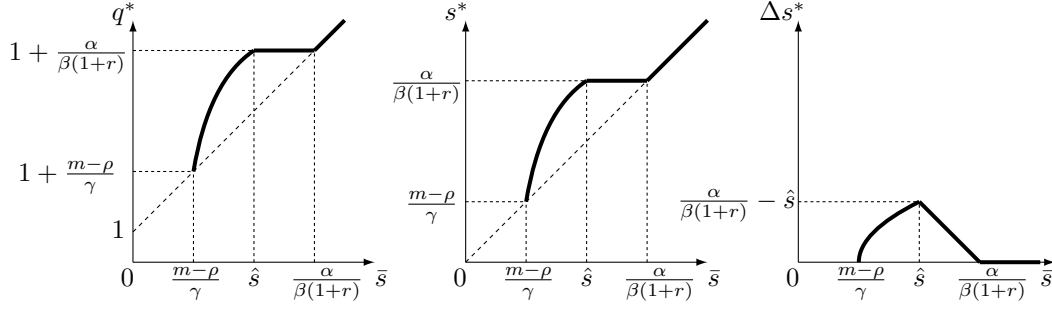


Figure 2.1: Optimal Choice of Migrants

Selection into Migration

Knowing the optimal level of occupation and schooling that they choose in cities, agents will decide to migrate if and only if the net gains from migration, $v^*(\bar{s})$, are positive:

$$v^*(\bar{s}) = y_u^* - y_r^* - \underbrace{\beta(s^{*2} - \bar{s}^2)}_{\text{Direct cost of schooling}} - \underbrace{m}_{\text{Moving cost}} - \underbrace{r\beta(s^{*2} - \bar{s}^2)}_{\text{Opportunity cost of schooling}} - \underbrace{rm}_{\text{Opportunity cost of movement}} > 0, \quad (2.7)$$

where $y_u^* = \alpha(2 + \frac{1}{s^*})s^* + \delta$ and $y_r^* = \delta$. Suppose there exists $\bar{S}_m \equiv \{\bar{s} \in R : v^*(\bar{s}) > 0\}$, an answer for the questions of who migrates and who chooses what level of occupations and education depends upon where $\bar{S}_m$ and $\bar{s}_u$ lie in the distribution of initial schooling. As we already mentioned, delineating all potential equilibria is not for the purpose of this chapter. In the next sub-section, we shall look at an illustrative example to explain the economic mobility broadly consistent with data.

Equilibrium

In this sub-section, we want to establish a particular equilibrium by imposing additional restriction on our parameters. Due to the non-linearity of the selection decision into migration in terms of the initial level of schooling in the range of our interest (i.e., $\bar{s} \in [0, \bar{s}_u]$, such that $\frac{m-\rho}{\gamma} < \bar{s}_u \leq \hat{s}$) and a large set of parameters, it might still be quite complex to even show one equilibrium. However, the following lemma could simplify our task.

Lemma 1 *For all $\bar{s} \in [\frac{m-\rho}{\gamma}, \hat{s}]$, if an agent of type $\bar{s}'$ prefers to migrate, all agents $\bar{s} > \bar{s}'$ are strictly better-off in the urban sector.*

Proof The proof is by contradiction. See Appendix 2.8.1.

Proposition 1 *If $(1+r)m < \alpha[1 + \frac{2(m-\rho)}{\gamma}]$, all agents $\bar{s} \geq \frac{m-\rho}{\gamma}$ optimally choose to migrate into the urban sector in equilibrium, whereas agents $\bar{s} < \frac{m-\rho}{\gamma}$ stay in the rural sector.*

Proof See Appendix 2.8.1.

One can note that significantly higher return to schooling (large α) satisfies the condition $(1+r)m < \alpha \left[1 + \frac{2(m-\rho)}{\gamma}\right]$. Then, we can make the following equilibrium predictions.

Prediction 1 *The initially educated ($\bar{s} \in [\frac{m-\rho}{\gamma}, \bar{s}_u]$) are willing to migrate (positive selection into migration).*

Prediction 2 *Agents acquire more additional schooling as they are initially more educated (For non-migrants, $\Delta s^* = 0$. For migrants, $\Delta s^* \geq 0$ and $d\Delta s^*/d\bar{s} > 0$) (positive selection into schooling).*

Prediction 3 *Agents choose higher income occupations as they are initially more educated (For non-migrants, $q^* = 0$ and $y_r^* = \delta$. For migrants, $q^* \geq 1 + \frac{m-\rho}{\gamma}$, $y_u^* > y_r^*$, $dq^*/d\bar{s} > 0$, and $dy_u^*/d\bar{s} > 0$).*

2.2.2 Migration of Educational Attainers

In addition to the mobility described above, another type of mobility relates the educational attainment to migration: people move to cities after they acquire skills in the rural sector to succeed at their destination. The equilibrium in this type of mobility can easily be shown by slightly changing the set-up of the model. Now, we model that opportunities for learning are supplied only in the rural sector with the remaining assumptions unchanged. So, if agents prefer to migrate, they simply choose occupations in cities given human capital imported from their origins. On the other hand, in the rural sector, they have to choose between engaging in self-employed farming and migrating after completing the schooling.

Migrants' Occupational Choice

Migrants' decision problem in the urban economy simply becomes

$$\max_{\{q\}} \alpha [2q - (q - s^*)^2] + \delta, \quad (2.8)$$

where s^* is the equilibrium level of education imported from their origins. Solving this yields $q^* = 1 + s^*$.

Schooling Decision

In the rural economy, no one invests in schooling if they do not migrate, and non-migrants simply engage in self-employed farming. This is because education is not rewarded in the rural area. On the other hand, if agents prefer to migrate, they have to decide how many new skills to acquire before migrating. The decision problem is

$$\max_{\{s\}} \alpha [2q - (q - s)^2] + \delta - \beta(s^2 - \bar{s}^2) - m + r[\gamma\bar{s} + \rho - m - \beta(s^2 - \bar{s}^2)] \quad (2.9)$$

subject to $q = 1 + s$, $\Delta s \geq 0$ and $m + \beta(s^2 - \bar{s}^2) \leq \gamma\bar{s} + \rho$.

Equilibrium

The solution of (2.8) and (2.9) is exactly the same as one in the previous mobility (i.e., migration for schooling). Given the solution, agents decide whether to migrate based upon (2.7). Thus, the equilibrium predictions are the same as before, except for migrants attain additional education before migrating.

2.2.3 Summary of the Model

This section constructed a simple rural-to-urban migration model. Urban jobs reward migrants for their education ($\alpha > 0$), and the positive return to schooling gives rural agents an incentive to migrate, encouraging the less educated to invest in further schooling before joining the labour force, conditional on the initially endowed level of education. Depending upon where learning opportunities are available, some migrate to seek more education and others migrate after obtaining it.¹⁰ The investment process of migration and schooling eventually leads migrants into higher earnings. In the model, the initial level of schooling plays a role of credit and, at equilibrium, in which the credit constraint is binding, not all agents can migrate and not all migrants can achieve the education of their first best. The model made three main predications at the equilibrium: firstly, migrants are more likely to be in the upper tail of the initial schooling distribution (positive selection into migration). Secondly, agents seek more education as they are initially more educated (positive selection into schooling). Thirdly, initially educated agents eventually achieve higher

¹⁰Strictly speaking, an agent $\bar{s} = \frac{m-\rho}{\gamma}$ will migrate without further investment in schooling.

earnings. Section 2.5 briefly checks these assumptions and predictions using our data.

The model has two closely related implications for an empirical analysis. Firstly, while higher earnings can be achieved by moving out, it is education that generates the gains. Secondly, the reason why the education generates the gains is that the return to schooling is higher in an urban area than in a rural area. Both of the implications point towards the notion that the migration effect is the schooling effect. These implications are clear to see by writing the earnings of agents in the form of

$$y^* = \underbrace{\alpha \left(2 + \frac{1}{s^*} \right)}_{\text{Migration effect}} \times s^* \times M + \delta, \quad (2.10)$$

where M is a migration dummy, equal to one for a migrant and zero otherwise. This shows that the gains from migration are small for the less educated, and that clearly, the positive α is a source of raising migrants' earnings, suggesting that those who cannot afford the cost of schooling (as well as the cost of migration) cannot reap the benefits of migration. If the return to education in an urban area is the same as in a rural area ($\alpha = 0$), there is no migration gain. In this case, no one migrates to cities.

2.3 Empirical Strategy

The implications from our theoretical model in Section 2.2 generate the following hypothesis tested in this chapter.

Hypothesis 1 *If the migration effect is the schooling effect, no direct effect of migration on outcomes (income and consumption) would exist after controlling for the current level of education and allowing for the difference in the return to schooling between the origin and the destination.*

This section outlines our empirical strategy to test this hypothesis. Throughout the section, we consider three periods: periods 0 (initial period), t , and $t + 1$, and assume that migration occurs in the period t between the periods 0 and $t + 1$ and that the effect is realized in the period $t + 1$.

The above hypothesis, motivated by equation (2.10), encourages us to estimate the following earnings function (the framework can directly be applied to consumption),

$$\log y_{it+1} = \alpha_0 + \alpha_1 M_{it} + \alpha_2 s_{it+1} + \alpha_3 (s_{it+1} - E(s_{it+1})) M_{it} + \alpha'_y \mathbf{z}_y + \epsilon_{it+1}, \quad (2.11)$$

where y_{it+1} is an individual i 's earnings in period $t + 1$; s_{it+1} is his schooling level (years); M_{it} is a migration dummy, which takes one if he moved out of his original community in the period t and zero otherwise; and a vector $\mathbf{z}_y$ contains other determinants of earnings, such as age and gender. The expectation operator $E(\cdot)$ implies the population average, but the sample average is used in an actual estimation. To the extent that the migration effect is the schooling effect, we expect that $\alpha_1 = 0$ and $\alpha_3 > 0$. However, simple OLS does not enable us to identify the parameters of interest because of the endogeneity of migration and schooling decisions. As modelled in Section 2.2, it is certainly possible that the expected future earnings (thus, more or less, realized future earnings) drive both the decisions to invest in schooling and to migrate, making M_{it} and s_{it+1} correlated with ϵ_{it+1} . Then, how can we identify the parameters in an earnings function (2.11)?

The chapter takes two approaches to address the endogeneity of migration and schooling decisions. The first (an IV approach) just ignores the interaction term between migration and schooling in (2.11) and estimates

$$\log y_{it+1} = \alpha_0 + \alpha_1 M_{it} + \alpha_2 s_{it+1} + \alpha'_y \mathbf{z}_y + \epsilon_{it+1}. \quad (2.12)$$

Obviously, this is not the best strategy. This is because the specification assumes that the return to schooling is the same across spaces (homogeneous return), contrary to our expectation that the difference in the return between the origin and the destination primarily drives migration, benefitting migrants. However, dropping the interaction term allows us structurally to address the endogeneity by explicitly modelling an individual's self-selection into migration and schooling and using a relatively straightforward econometric technique, such as 2SLS. While the homogeneous return is an unfavourable assumption, we still expect that controlling for the current level of education in an outcome equation could substantially reduce the migration effect if education is an underlying mechanism that benefits migrants ($\alpha_1 = 0$ and $\alpha_2 > 0$ in (2.12)).

The other approach is to estimate (2.11) by taking a control function (CF) approach (Garen, 1984; Card, 2001; Blundell et al., 2005). This approach is not structural and assumes a particular form of correlation of an error term in an outcome equation with that in schooling and migration equations. The assumption is strong as explained later, but the approach still enables us to estimate (2.11) as it is, allowing the return to schooling to differ between the origin and the destination.

2.3.1 IV Approach: Structural Modelling Approach

The IV approach addresses the endogenous selection into migration and schooling by extending the model presented in Section 2.2 into an estimatable stochastic three-period model.

(i) Selection into Migration

Following the discussion in Section 2.2, an individual i moves out of his original community in period t if the expected present value of net gains from migration v_{it}^* is positive.

$$M_{it} = I(v_{it}^* > 0), \quad (2.13)$$

where $I(\cdot)$ is an indicator function. Equation (2.7) contains several elements to model the net gains, and we model it as

$$v_{it}^* = V(s_{it}, E[\log y_{it+1}|\Omega_{it}], s_{it+1}^* - s_{i0}, \mathbf{z}_{\mathbf{m}}), \quad (2.14)$$

where $E[\log y_{it+1}|\Omega_{it}]$ is the expected earnings that the individual earns in period $t + 1$ (Ω_{it} is an information set at t), s_{it+1}^* is the schooling level that he wants to achieve, and a vector $\mathbf{z}_{\mathbf{m}}$ contains other variables that determine the cost and benefit of migration. In accordance with (2.7), the investment in schooling between period 0 and $t + 1$ ($s_{it+1}^* - s_{i0}$) can affect the value of migration by determining both direct and indirect (opportunity) costs of schooling in education-cum-migration. In addition, the expected future earnings determine the value of migration as long as attainable earnings systematically vary across spaces (e.g., higher income occupations are exclusively supplied in the urban sector). While the model in Section 2.2 does not take into account this thought, skilled migrants might easily be able to find jobs in destination labour markets, or more generally, adapt quickly to a new life in the urban sector and, as such, the level of schooling at

migration might increase the value of migration by lowering the cost of migration (i.e., $m = m(\bar{s})$ such that $m'(\bar{s}) < 0$). To allow for the possibility, we independently include the level of education at migration (s_{it}) in attributes determining v_{it}^* in (2.14), as well as the investment in schooling.

By assuming the linear probability model, we derive a migration equation as

$$M_{it} = \beta_0 + \beta_1 s_{it} + \beta_2 E[y_{it+1} | \Omega_{it}] + \beta_3 (s_{it+1}^* - s_{i0}) + \beta'_m \mathbf{z}_m + v_{it}. \quad (2.15)$$

To make equation (2.15) tractable in the empirical analysis, we assume that $\log y_{it+1} = E[\log y_{it+1} | \Omega_{it}] + e_{it+1}^y$ and $s_{it+1} = s_{it+1}^* + e_{it+1}^s$, where e_{it+1}^y is an expectation error between expected earnings and realized earnings, and e_{it+1}^s is the difference between the target level of schooling and the actual obtained level of schooling. Substituting these into (2.15), an estimatable version of migration equation can be obtained as

$$M_{it} = \beta_0 + \beta_1 s_{it} + \beta_2 \log y_{it+1} + \beta_3 (s_{it+1} - s_{i0}) + \beta'_m \mathbf{z}_m + \tilde{v}_{it}, \quad (2.16)$$

where $\tilde{v}_{it} \equiv v_{it} - \beta_2 e_{it+1}^y - \beta_3 e_{it+1}^s$. Equation (2.16) suggests that when modelling a migration decision, we need information on migrants' education at three points in time. As this is a potential issue in applying (2.16) to our data, we will come back to this point in Section 2.3.1 (iii).

(ii) Selection into Schooling

Following equation (2.6), individuals invest in schooling as long as the marginal benefit of an extra year of schooling is equal to or greater than the marginal cost:

$$\frac{\partial y_u}{\partial s} \geq 2\beta(1+r)s. \quad (2.17)$$

This suggests that, conditional on the initial level of schooling ($\bar{s}$), higher earnings in an urban area encourage migrants' investment in schooling, and that both the higher direct (large β) and indirect (large r) costs of additional education discourage it. While the model in Section 2.2 does not formalize this idea, we think that the marginal cost of schooling can vary with the initial level of schooling (i.e., $\beta = \beta(\bar{s})$). For example, as education is the dynamic process of accumulating human capital, the marginal (psychological) cost of an extra year of schooling might decline as

individuals are initially more educated ($\beta'(\bar{s}) < 0$). In addition, it is certainly possible that the cost of education differs between rural and urban areas due to the difference in social and economic environment (i.e., $\beta = \beta(M)$). These discussions allow us to model the target level of schooling as

$$s_{it+1}^* = \gamma_0 + \gamma_1 s_{i0} + \gamma_2 E[\log y_{it+1} | \Omega_{it}] + \gamma_3 M_{it} + \gamma'_s \mathbf{z}_s + \omega_{it+1}, \quad (2.18)$$

where a vector $\mathbf{z}_s$ contains variables other than s_{i0} , $E[\log y_{it+1} | \Omega_{it}]$, and M_{it} that determine the cost and benefit of additional schooling. As before, we can rewrite (2.18) in the form of an estimatable equation as

$$s_{it+1} = \gamma_0 + \gamma_1 s_{i0} + \gamma_2 \log y_{it+1} + \gamma_3 M_{it} + \gamma'_s \mathbf{z}_s + \tilde{\omega}_{it+1}, \quad (2.19)$$

where $\tilde{\omega}_{it+1} = \omega_{it+1} - \gamma_2 e_{it+1}^y + e_{it+1}^s$.

(iii) Fitting the Framework to Data

The above structural framework yields three simultaneous equations: (2.12), (2.16) and (2.19). The straightforward way of fitting them to our data is to treat waves 1 (year 1991) and 5 (year 2004) as periods 0 and $t + 1$, respectively. However, there is one potential problem in doing so, as equation (2.16) requires information on the level of education at migration (s_{it}), which is not available in our data.

As we will see in Section 2.4.1, the data contain a sufficient number of migrants with their leave from original communities distributed between waves 1 and 5 with an adequate variation. This characteristic of data generates three possible relationships among the level of s_{i0} , s_{it} and s_{it+1} of those who migrated between waves 1 and 5. Suppose that the migrants completed all their education by wave 1 (e.g., the old in wave 1), we expect that $s_{i0} = s_{it} = s_{it+1}$. On the other hand, for the migrants who would start education in due course or in the middle of their education in wave 1 (e.g., the children in wave 1), it will be the case that $s_{i0} \leq s_{it}$, where the difference between s_{i0} and s_{it} is the migrants' investment in rural education. In particular, if the migrants moved for the purpose of schooling (as in the case of Section 2.2.1), $s_{i0} \leq s_{it} < s_{it+1}$, where the migrants' investment in urban education can be captured by the difference between s_{it} and s_{it+1} . Alternatively, if they migrated after completing their education in rural areas (as in the case of

Section 2.2.2), it is likely that $s_{i0} < s_{it} = s_{it+1}$.

While the level of migrants' education when they relocated is not available in data, it is highly likely that s_{it} is a function of s_{i0} , provided that the cost of rural education varies with the initial level of schooling. Assuming the linear relationship between s_{it} and s_{i0} as $s_{it} = \frac{\beta_4}{\beta_1}s_{i0} + \beta_5$ for tractability's sake and substituting this into (2.16), we can obtain

$$M_{it} = \tilde{\beta}_0 + (\beta_4 - \beta_3)s_{i0} + \beta_2 \log y_{it+1} + \beta_3 s_{it+1} + \beta'_m \mathbf{z}_m + \tilde{v}_{it}, \quad (2.20)$$

where $\tilde{\beta}_0 = \beta_0 + \beta_1\beta_5$. This is a migration equation exactly estimated in the subsequent empirical analysis. While (2.20) is not the first best of modelling a migration decision, estimating a migration equation is not in the first-order of importance in this chapter. As long as our main interest lies in estimating (2.12), it does not appear that modelling a migration decision in the form of (2.20) is a serious problem in this study.

(iv) Identification and Estimation

Now, we have three simultaneous equations: (2.12), (2.19), and (2.20). To identify the structural parameters in equation (2.12), we need to make the following identification assumptions.

IV assumption 1 *All $\mathbf{z}$ variables are orthogonal to all structural disturbances $(\epsilon_{it+1}, \tilde{v}_{it}, \tilde{\omega}_{it+1})$.*

In Section 2.6, we will explain the $\mathbf{z}$ variables actually exploited in the estimation.

IV assumption 2 *The initial level of schooling s_{i0} is predetermined.*

The initial level of schooling is not obviously exogenous, but it could be regarded as having already been determined with respect to the current values of endogenous variables $(y_{it+1}, M_{it}, s_{it+1})$.¹¹

IV assumption 3 *Both the order and rank conditions for equation (2.12) hold.*

The order condition requires that either $\mathbf{z}_m$ or $\mathbf{z}_s$ (or both of them) contain at least two exogenous variables that do not appear in equation (2.12). Potential candidates for those variables will be discussed in Section 2.6. With respect to the rank condition, we simply assume that it is satisfied.

¹¹This assumption requires that an expectation error in earnings (e_{it+1}^y) also be uncorrelated with the initial level of schooling s_{i0} . However, the assumption might be too strong if the precision of expectation varies with s_{i0} (e.g., the expectation of the educated is more precise). Related to the point, Attanasio and Kaufmann (2009) provide evidence that expected earnings could depend upon the current level of academic achievements measured by the Grade Point Average (GPA).

If the rank condition fails, however, the structural parameters will be poorly identified, with very large estimated standard errors in finite samples.

The estimation procedure is straightforward (see Wooldridge (2002) and Greene (2003), for examples). With all the above maintained assumptions, we firstly estimate the reduced form of all endogenous variables by OLS¹² and then estimate (2.12) by using the least squares predictions from the first stage. We mainly report the estimation results of equation-by-equation 2SLS in the chapter.

2.3.2 Control Function (CF) Approach

In contrast to the IV approach, the CF approach addresses the endogeneity of migration and schooling in a non-structural way by directly controlling for the correlation of an error term in an outcome equation with that in the migration and schooling equations. The advantage is to allow for the heterogeneity in the return to schooling across spaces. However, it needs strong assumptions about a form of the correlation. Now, we assume that a stochastic error in (2.11) consists of two additively separable parts:

$$\epsilon_{it+1} = \epsilon_{it+1}^m + \epsilon_{it+1}^s. \quad (2.21)$$

In addition, let us model migration and schooling equations as

$$M_{it} = I[g_m(s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s) + u_{it}^m > 0], \quad (2.22)$$

$$s_{it+1} = g_s(s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s) + u_{it+1}^s, \quad (2.23)$$

where g_j , $j \in (m, s)$, is a linear function of attributes. Taking (2.11), (2.21), (2.22), and (2.23) together, we make the following identification assumptions.

CF assumption 1 $E(u_{it}^m | s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s) = 0$.

¹²The reduced form expresses all endogenous variables as a linear function of a predetermined variable, exogenous variables, and stochastic disturbances:

$$\begin{aligned} y_{it+1} &= f_1(s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s, \epsilon_{it+1}, \tilde{v}_{it}, \tilde{\omega}_{it+1}), \\ M_{it} &= f_2(s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s, \epsilon_{it+1}, \tilde{v}_{it}, \tilde{\omega}_{it+1}), \\ s_{it+1} &= f_3(s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s, \epsilon_{it+1}, \tilde{v}_{it}, \tilde{\omega}_{it+1}). \end{aligned}$$

The precisely derived version of these equations is available from the author if required.

CF assumption 2 $E(u_{it+1}^s | s_{it}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s) = 0$.

If we additionally assume a linear probability model for (2.22), equations (2.22) and (2.23) are nothing but a reduced form of the migration and schooling equations in the IV approach. Then, these two assumptions are, in essence, equivalent to the IV assumptions 1 and 2.

CF assumption 3 $E(\epsilon_{it+1}^m | M_{it}, s_{it+1}, \mathbf{z}_y) = E(\epsilon_{it+1}^m | M_{it}, \mathbf{z}_y)$.

CF assumption 4 $E(\epsilon_{it+1}^s | M_{it}, s_{it+1}, \mathbf{z}_y) = E(\epsilon_{it+1}^s | s_{it+1}, \mathbf{z}_y)$.

According to these assumptions, s_{it+1} is a statistically mean independent of ϵ_{it+1}^m , conditional on M_{it} and $\mathbf{z}_y$, and M_{it} is a statistically mean independent of ϵ_{it+1}^s , conditional on s_{it+1} and $\mathbf{z}_y$. Roughly speaking, these assumptions suggest that an unobserved component in an outcome equation that has correlation with a migration decision can be separated from the other unobserved component that has correlation with a schooling decision. These can be challenged, provided that an individual's identical unobserved traits determining the outcome simultaneously affect the decisions to invest in both migration and schooling. For example, it is certainly possible that relocating to new places and investing in further education require the same type of reasoning ability and cognitive skills (e.g., language skills). Alternatively, it is also likely that the same type of non-reasoning and non-cognitive attributes, such as motivation, enable individuals jointly to attempt a venture of moving from their original communities and to be ambitious enough to seek higher education. In these cases, the CF assumptions 3 and 4 will be untenable, although they allow us to econometrically move ahead.

CF assumption 5

$$\left(\begin{array}{c} \epsilon_{it+1}^m \\ u_{it}^m \end{array} \middle| \mathbf{z}_y \right) \sim N \left[\begin{array}{c} \left(\begin{array}{c} 0 \\ 0 \end{array} \right), \left(\begin{array}{cc} \sigma_1^2 & \rho_m \sigma_1 \\ \rho_m \sigma_1 & 1 \end{array} \right) \end{array} \middle| \mathbf{z}_y \right], \quad (2.24)$$

$$\left(\begin{array}{c} \epsilon_{it+1}^s \\ u_{it+1}^s \end{array} \middle| \mathbf{z}_y \right) \sim N \left[\begin{array}{c} \left(\begin{array}{c} 0 \\ 0 \end{array} \right), \left(\begin{array}{cc} \sigma_2^2 & \rho_s \sigma_2 \sigma_s \\ \rho_s \sigma_2 \sigma_s & \sigma_s^2 \end{array} \right) \end{array} \middle| \mathbf{z}_y \right]. \quad (2.25)$$

We assume that ϵ_{it+1}^m and u_{it}^m are jointly normally distributed, conditional on $\mathbf{z}_y$, and that ϵ_{it+1}^s and u_{it+1}^s are too. With all the above maintained assumptions, we can express earnings as

$$\begin{aligned} \log y_{it+1} = & \alpha_0 + \alpha_1 M_{it} + \alpha_2 s_{it+1} + \alpha_3 (s_{it+1} - E(s_{it+1})) M_{it} + \alpha'_y \mathbf{z}_y \\ & + \rho_m \sigma_1 \left[\frac{\phi[g_m(\cdot)]}{\Phi[g_m(\cdot)]} M_{it} - \frac{\phi[g_m(\cdot)]}{1 - \Phi[g_m(\cdot)]} (1 - M_{it}) \right] + \frac{\rho_s \sigma_2}{\sigma_s} u_{i+1}^s + \tilde{\epsilon}_{it+1}, \end{aligned} \quad (2.26)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are standard normal density and distribution functions, respectively, and $E[\tilde{\epsilon}_{it+1} | M_{it}, s_{it+1}, s_{i0}, \mathbf{z}_y, \mathbf{z}_m, \mathbf{z}_s] = 0$. The square bracket term containing the inverse Mills ratio can be derived by estimating a probit model of the migration equation (2.22), whereas u_{i+1}^s can be consistently estimated as a predicted residual from the schooling equation (2.23). Since these are estimates, we estimate the standard errors by bootstrapping with 100 replications in estimating (2.26). As in the case of the IV approach, we fit the framework to waves 1 (period 0) and 5 (period $t + 1$) in our data.

2.4 Data

The data used in this chapter are drawn from the Kagera Health and Development Survey (KHDS). The KHDS is a longitudinal household panel survey that consists of five waves¹³ and started in the rural region in northwest Tanzania, Kagera, as a part of a research project on adult mortality in Sub-Saharan Africa launched by the World Bank in 1991.

The KHDS sample households were, with stratification based on geography and mortality risk, randomly drawn through two stages: selection of clusters and selection of households. In the first stage, 550 geographical areas (communities) delineated by the 1988 Tanzanian Census were initially classified into eight strata defined over four agronomic zones and the level of adult mortality (high and low) in each zone. Then, six or seven communities were selected from each stratum, which generated 51 clusters (of 49 communities). In the second stage, 16 sample households were drawn from each of the 51 clusters. In order to draw the sample, households in each selected cluster were enumerated, which amounted to more than 29,000 in all selected clusters, and were sorted into two groups according to information on adult mortality and morbidity within households: sick or

¹³Wave 1, September 1991 to May 1992; wave 2, April 1992 to November 1992; wave 3, November 1992 to May 1993; wave 4, June 1993 to January 1994; wave 5, January 2004 to August 2004.

well. In every cluster, 14 households were finally selected at random from the sick group and 2 from the well group (See *User's Guide to the Kagera Health and Development Survey Datasets* (2004) and *Kagera Health and Development Survey 2004 – Basic Information Document* (2006) for details of sampling design). These sampling procedures as a whole generated a sample where households with a high risk of an adult death were overweighted; therefore, our findings in the subsequent sections should be treated with caution in generalizing them when those health characteristics affect the decision of migration, the choice of occupations, and the attainment of schooling.

The first four surveys were conducted between 1991 and 1994 with a 6- or 7-month interval. Based upon the sampling strategy explained in the previous paragraph, a total of 840 households were interviewed in the first wave,¹⁴ with the 759 households continuously contacted in the subsequent three waves. Since the subsequent three waves replaced households that dropped from the sample,¹⁵ the first four waves eventually covered 912 households in total. By the number of interviews conducted, the 912 households can break down into: 39 households interviewed once, 45 households twice, 69 households three times, and 759 households interviewed throughout the waves.

After a 10-year gap from wave 4, approximately 91% of the 912 baseline households were recontacted in 2004 (wave 5), even if located outside their original villages. The significantly high recontact rate is one innovation in such a long-run panel survey. When previous members moved out of their original households, their new households were traced. This exercise generated 2719 household surveys in wave 5, split from the 832 recontacted original households. Of the 2719 households, only half stayed in the same village as 10 years previously. This suggests a substantial demographic mobility in this region for those 10 years.

While the KHDS is a household panel survey, it enabled us to construct unbalanced panel data from wave 1 to 5 on the basis of individuals, as the survey collected the information of all household members in all waves. The rate of sample attrition in the KHDS is very low even on an individual basis. Based upon the careful examination of sample attribution by Beegle et al.

¹⁴The 840 households consist of 816 original households selected from the enumeration based upon the sampling strategy explained in the previous paragraph (16 households $\times$ 51 clusters) and extra 24 households taken from the list of replacement. The extra sample was added when the field team sensed that some original households were likely to drop out or to provide poor information.

¹⁵Approximately 80% of the attrition was due to the relocation of households.

(2008b), excluding people who died, approximately 82% of 5394 original respondents interviewed in the first four waves were successfully recontacted in wave 5.

Throughout the waves, a standardized survey questionnaire was used, although several changes were made in wave 5. This allows collected information to be highly comparable across waves. In addition, the data contain a variety of information related to a household and its members, such as education, health conditions, income generating activities (e.g., hours worked, earnings), economic shocks, and expenditures, which makes the KHDS highly resourceful for an empirical study.

2.4.1 Migration

Our identification about whether individuals are migrants or not is based upon the year of migration and the place of living in 2004. This information was collected in wave 5. Since the strategy does not enable us to identify whether earlier respondents not in wave 5 migrated out during the sample periods,¹⁶ we restrict our sample only to those who are in both wave 5 and one of the earlier four waves, and call them panel individuals in this chapter. This means that we exclude from the analysis both respondents interviewed only in the earlier four waves and new respondents interviewed only in wave 5,¹⁷ which leads us to exploit 4476 panel individuals in this study. Of those panel individuals, we define those who moved out of their original communities between 1991 and 2004 as migrants, resulting in 1644 migrants. Migrants can be split into those who moved out but remained inside Kagera (1289 inside migrants) and outside (355 outside migrants), where “outside” includes elsewhere in Tanzania and Uganda.¹⁸

Firstly, we grouped panel individuals in wave 1 by original location together with the breakdown of the subsequent migration status (non-migrants, inside migrants, and outside migrants) in Table 2.1. While the slightly larger proportion of migrants (in particular, outside migrants) resulted from Biharamulo, Bukoba Rural, and Bukoba Urban than the other districts, on average, all locations supplied 30 to 40% of panel individuals as migrants into elsewhere at some point within the 13 years following wave 1. Next, in Table 2.2, migrants were classified by age at migration, and data show that most of them moved out of their original communities between the ages of 7 to 30 years.

¹⁶Attrition could happen for several reasons, such as not only the difficulty in tracing individuals but also the death of the previous elderly. See also Beegle et al. (2008b) for the issue of sample attrition in the KHDS.

¹⁷Excluding new respondents in wave 5, we dropped approximately 22% of a total of 23,915 observations of person-wave units in all five waves.

¹⁸Those who migrated to other countries were not traced, although there were not too many.

In addition, the distribution of migrants was checked by year of migration, and this is shown in Table 2.3. They are distributed over the entire sample periods with reasonable variation, although 1994 is seen as the peak year of migration in our data, which may be related to an economic trough in that year. As our data are panel tracing the same individuals over 13 years, these suggest that the migrants in our data were mostly the young at the start of the KHDS.

Table 2.4 groups migrants by destination to see the reason for migration. In both destinations, marriage is the primary reason for the migration of the female. On the other hand, economic motives largely appear to drive the migration of the male. Seeking for business opportunities and better jobs takes up a large proportion of the reason for outside migration of the male. The lack of arable land is also an important reason for inside male migrants. The recent increase in population in this region might have caused this (Mitti and Rweyemamu, 2001). Seeking a higher level of schooling is another common reason for outside migrants. About 15% of the movement of outside migrants is attributed to the attainment of education. Apart from natural disasters, it seems that marriage, economic motives, and schooling are key factors driving an individual's self-selection into migration. In pooling inside and outside migrants, about 48% of male migrants moved for economic or schooling reasons, whereas about 47% of female migrants moved for marriage.

While our analytical framework is modelled based upon economic or schooling migration, those migrants constitute just 30% of total migrants in our data. Then, one natural question is whether to exclude non-economic migrants from the analysis. There are several reasons not to do so. First of all, the primary interest of the chapter is the impact of education on migrants' outcomes. The fact that migrants said that they migrated for non-economic reasons does not mean that education does not impact on outcomes. The migration reason is important in modelling a migration decision. But, estimating the migration equation is not in the first-order of importance in the chapter. Secondly, keeping all types of migrants allows us to model the migration decision more flexibly in empirics, and it indeed helps us find good instruments that explain migration but not outcomes. Moreover, confining our analysis only to economic or schooling migrants can artificially generate a sample selection problem. This is why we keep all types of migrants in this chapter.

2.4.2 Occupational Mobility

While Kagera is a rural area, it has seemingly been difficult for people only to survive on self-employed farming. Figure 2.2 plots each income generating activity's share of weekly hours worked in total by migration status. In all groups of panel individuals, farming self-employment gradually lost the leading position, whereas non-farm wage employment and non-farm self-employed businesses started to become more important.¹⁹ In particular, outside migrants spend more than 90% of total hours worked on non-farm wage employment and non-farm self-employment in wave 5 and, as a result, they replace farming self-employment as a primary income source.

However, Figure 2.2 does not allow us to distinguish between two potential mechanisms driving the picture. For example, all panel workers in all age groups may uniformly be moving away from farming self-employment to non-farm activities. Alternatively, only new participants in labour markets (i.e., children in wave 1) might exclusively be entering into non-farm activities once they become working-age adults, whereas the previous elderly did not change the type of jobs that they engaged in. More careful examination of data in Appendix 2.11.1 supports the latter story. The activity selected as an alternative depends upon the level of schooling, and data suggest that the (young) educated are more likely to engage in non-farm wage jobs than non-farm self-employed activities.

2.4.3 Income Measurement Issues

While measurement errors are always an inevitable concern for researchers who conduct empirical studies, they might not necessarily detract from the comparability of measured variables between individuals and over time, when there is little heterogeneity in the behaviour of individuals. This is likely to be the case in waves 1 to 4 in our data. However, taking into account the significant difference in the pattern of individual labour allocation between migrants and non-migrants in wave 5 and the occupational mobility towards non-farm activities in all groups; it seems that small measurement errors in earnings from a particular activity in wave 5 make us lose the comparability of total labour income between individuals and waves. As this forces us to be very cautious about

¹⁹One caution about the trend in the early 1990s is that part of the occupational mobility may be generated by an individual's response to smooth income against the economic recession (Kochar, 1999). In fact, we found that the proportion of non-farm businesses aged 3 months below relative to the total number of them was larger in waves 2 to 4 than in the other two waves.

correctly measuring earnings in wave 5, this sub-section explains potential income measurement issues in our data.

The survey questionnaire allows us to construct two possible income measures for each individual: per worker weekly income and per capita weekly income (average of the total labour income of household), both of which are computed as the sum of wage earnings, farming self-employment income, and non-farm self-employment income. As the reference periods of recorded income were different between income generating activities in the questionnaire: 1 week for wage employment, 1 year (or 6 months in waves 2 to 4) for farming self-employment, and 2 weeks for non-farm self-employment, we simply took the time average to compute weekly income. As opposed to wage earnings that respondents aged 7 years or older individually reported, both farming and non-farm self-employment incomes were collected on a household basis. Thus, in per worker income measure, they were assigned to each household member in proportion to hours worked on those activities in the past 7 days, which were individually reported by respondents aged 7 years or older in a different section from sections collecting household-level self-employment income.²⁰ As a result, per worker income measure is constructed only for panel individuals aged 7 years or older.

Table 2.5 (rows (2) and (3)) summarize the two income measures and a substantial difference can be seen between them. After the exhaustive examination of the accuracy of our income measures in Appendix 2.10, we found that we were seriously underestimating the income of migrants' *household* (in particular, outside migrants) in wave 5 relative to its actual amount. In other words, household income (latent income) that is not contained in our income measures exists in wave 5 and this problem is more serious for migrants. Two main reasons are detected: firstly, given the way of constructing income measures, we are systematically failing to capture household workers engaged in wage employment and non-farm self-employment due to the difference in the periodicity between income generating activities in the survey questionnaire. For example, if one does not happen to engage in those activities at the point of survey, we lose his contribution to household income, even though he usually makes a living from wage jobs or non-farm self-employed activities. If this is the case, we will underestimate household income as the household's main economic activities concentrate more on wage employment and non-farm self-employment and as the access to such activities becomes more irregular. Taking into account the occupational mobility towards non-farm

²⁰Thus, adult and child workers are equally weighted in terms of their efficiency in the analysis.

activities and generally episodic urban occupations, it is highly likely that we are failing to capture the full extent of workers available to (in particular, outside) migrants' households in wave 5.

Secondly, a relatively large number of households did not report the fortnightly income from non-farm self-employed businesses in wave 5 despite some of their members reporting positive hours worked on those activities in the past 7 days. This is clear from comparing the number of non-farm self-employed observations in wave 5 between rows (2) and (3) in Table 2.6, which present descriptive statistics at the household level. In the table, households are defined as migrants in waves 1 to 4 if at least one member moved out of his original community between 1991 and 2004 and in wave 5 if they contained at least one panel migrant as a member. For example, it is critical that only 11 households of outside migrants reported income from non-farm self-employment, whereas 83 households actually engaged in it.²¹

After recognizing these problems, we decided to exploit per worker weekly income as a primary income measure in the analysis. This preference results from it being more accurate than per capita income for two reasons: firstly, the latent income problem conceptually unaffected per worker income and, as such, it is accurate as long as the reported income is correct. This is because the way of constructing per worker measure ensures that collected household income is allocated only to those who actually earned it in accordance with their contribution. On the other hand, per capita income is very vulnerable to the problem. Since we are not correctly measuring earnings at the household level, neither are we at the per capita level. Secondly, we expect that per capita income measure is generally very sensitive to the demographic structure of the household. Our sample shows that the proportion of workers is smaller in outside migrants' households than in the other groups in wave 5 (row (8) in Table 2.6), which implies lower per capita income. On the other hand, one can note from row (9) in Table 2.6 that the size of migrants' households is smaller than that of non-migrants in wave 5 and that it is likely to increase per capita income. As a result, it seems that how the demographic structure of household influences per capita income is more complicated than we think and that it more easily detracts from the comparability between individuals and waves when there is a large heterogeneity in the household structure and the pattern of individual labour allocation between activities. In fact, we found a large number of unreasonable observations

²¹This problem is also reflected in the difference in the number of observations between migrants' per worker weekly income and per worker weekly hours worked in wave 5 in Table 2.5.

in wave 5: many migrants (in particular, outside migrants) with lower per capita weekly income but higher per capita annual consumption. All these push us to exploit per worker income more reliably.

For per worker income to be used in the empirical analysis, it still calls attention to several aspects. Firstly, per worker income does not contain many outside migrants in wave 5 (approximately 3.8% of total panel workers earning income), and even when it does, outside migrants' per worker income is mainly calculated from earnings information on non-farm wage activities (approximately 67% of total outside migrant workers earning income), not on non-farm self-employed activities. In addition, the accuracy of measured per worker income is not guaranteed either. In Appendix 2.10.3, we estimated per worker earnings functions by activity (wage employment, farming self-employment, and non-farm self-employment), group, and wave (pooled waves 1 to 4 and wave 5) to test the accuracy of reported income. While we do not describe the details of the estimation results here to save space, it seems that migrants' earnings functions in both wage employment and non-farm self-employment are poorly specified relative to those of non-migrants in wave 5. We thus expect that in wave 5, measurement errors (if any) are more serious in earnings than in consumption, and that migrants' earnings contain larger noise than non-migrants' earnings do.

2.4.4 Welfare Changes

Since the economic recession in the early 1990s, which followed the second oil shock and the subsequent world recession, the Tanzanian economy has been recording a consolidating economic improvement. While Kagera is one of the remotest regions from the administrative centre of Dar es Salaam, a similar economic trend can be seen in our data in Table 2.5. Firstly, both mean and median of per capita annual consumption (the average of the total annual consumption of household) decline from wave 1 to 4 and then rise in wave 5. The pattern is consistent with the movement of national GDP.²² While there is almost no change in non-migrants' consumption between waves 1 and 5, migrants have achieved positive growth. In particular, the consumption growth of outside migrants is outstanding and the level has doubled from wave 1 to 5. As a result, outside migrants enjoy a higher level of consumption than inside migrants in wave 5, who, in turn,

²²According to World Economic Outlook Database (IMF, September 2004), GDP per capita gradually declined from 1990 to 1994 and has been recording a continuous growth since then (as of 2005) - <http://www.imf.org/external/pubs/ft/weo/2004/02/data/>.

enjoy a higher level than non-migrants do. Lower per worker income of migrants in wave 1 than that of non-migrants appears to reflect that most migrants are school children in wave 1. As in per capita consumption, however, outside migrants are earning more than inside migrants in wave 5, and inside migrants earn more than non-migrants.

2.5 Validating the Theoretical Model

In Section 2.3, we built our analytical framework for an empirical analysis on the theoretical model presented in Section 2.2. Thus, there is danger in estimating completely meaningless empirical equations if the model cannot account for data. While the purpose of this chapter is not to prove the model, this section checks if our data bear several assumptions and predictions explained in the theory. The model assumed that (1) the return to schooling was higher in the destination than in the origin; (2) the initial level of schooling was positively correlated with the pre-migration level of wealth, and showed that at equilibrium, in which a credit constraint was binding; (3) those who were initially educated were more likely to migrate (positive selection into migration); (4) those who were initially educated were more likely to invest in further schooling (positive selection into schooling); (5) some migrate to seek more education and others migrate after obtaining it; and (6) those who were initially educated would be better-off in the end. The assumption (1) is exactly a part of our hypothesis tested in this chapter, and we thus leave probing it to Section 2.6. The other assumptions and predictions are briefly checked in this section.

Since data contain several indicators of wealth, such as non-food expenditures, land area cultivated by households, household livestock, and household capital for farming (e.g., farm buildings, hand mills, tractors) and non-farming (e.g., buildings, vehicles like bicycles, tools), a simple way of checking the positive correlation between the pre-migration level of schooling and wealth is to regress these wealth variables on either individual or household average education by exploiting data in wave 1. However, there is a large number of missing observations for the livestock and capital stocks, we conducted the exercise only for the non-food expenditures and cultivated land. The log of per capita non-food consumption and land area cultivated by households, measured by acres, were estimated in Appendix Table 2.14, where an individual's education is measured by years based upon the highest grade that he/she completed by the time of each wave. Firstly,

the significantly positive correlation between the pre-migration level of non-food consumption and education was identified in columns (a) and (b). Using the log of per capita consumption as a dependent variable also identified the strong positive correlation in columns (c) and (d). Turning to the cultivated land area, columns (e) and (f) show that it positively correlates with the pre-migration level of education with the conventional level of significance.

In Table 2.7, panel respondents who were interviewed in both waves 1 and 5 are categorized by age in wave 1 and migration status to see how both the completed level and the attainment of schooling (years) are related to migration. It can be seen from the table that migrants in wave 5 are more educated than non-migrants and that this is true even before they migrate. With the exception of the youngest cohort in wave 5, a clear ordering in the level of schooling from outside migrants to non-migrants exists in both waves and in all age categories. In addition, it seems that outside migrants aged below 20 years in wave 1 have acquired a larger amount of new skills and knowledge over time than non-migrants and inside migrants. For example, the mean attainment of schooling was 5.7 years for outside migrants aged 7 to 15 in wave 1, which is more than one year longer than that of non-migrants and inside migrants. Similarly, outside migrants' educational attainment was 1.4 years in the age category of 16 to 20 years, whereas that of non-migrants and inside migrants was less than one year. These observations are broadly compatible with the above predictions (3) and (4). We explore the patterns of the positive selection into migration and schooling more formally in Section 2.6 when estimating the reduced form of migration and schooling equations.

One cannot identify from Table 2.7 when (before or after migration) and, hence, where (origin or destination) migrants have acquired the additional education. To see this, we compared the reasons for migration between those who have or not acquired a larger amount of education than the mean attainment of non-migrants, and this is shown in Table 2.8. About a quarter of outside migrants in the group of larger educational attainment migrated for schooling and this is the most important reason for this group of migrants. On the other hand, migration for schooling is less important for inside migrants. The table suggests that most inside migrants who acquired a larger amount of education during the sample periods moved out after completing schooling in their origins. This difference between inside and outside migrants can broadly be mirrored in the two types of mobility described in the above prediction (5).

We saw in Section 2.4.4 that outside migrants were better-off than inside migrants in wave 5 and inside migrants were better-off than non-migrants in terms of both income and consumption. Table 2.7 also shows that there was a clear ordering in the pre-migration level of education from outside migrants to non-migrants in each age cohort. These observations point to the above prediction (6).

2.6 Estimation Results

2.6.1 IV Approach: Structural Modelling Approach

We now go back to the system of simultaneous equations of (2.12), (2.19), and (2.20) to estimate the returns to migration and schooling in (2.12) separately. Age (and its squared term) and education (in wave 1 for (2.19) and (2.20) and wave 5 for (2.12)), a male dummy, and initial location dummies were included in all $\mathbf{z}$ variables.²³ Following Beegle et al. (2008b), we also controlled for original household heterogeneity by putting in dummies for initial households in all $\mathbf{z}$. Controlling for the initial household fixed effects (IHHFE) means that the impact of migration in (2.12) is identified based on a sample of households that have split since 1991 (within initial household variation).²⁴ In equation (2.19) and (2.20), the initial household fixed effects are expected to capture the household-specific determinants of migration and schooling, such as wealth and demographic structure within an original household.

The discussion in Section 2.3 indicated that to satisfy the order condition, we needed either the $\mathbf{z}_m$ or $\mathbf{z}_s$ (or both of them) variables that did not appear in (2.12). Then, four instruments to be included in $\mathbf{z}_m$ were constructed to explain migration. As shown in Table 2.4, in our sample, a relatively large number of migrants moved out because of natural disasters. While the type of natural disasters that were critical push factors into migration is unknown from the data, it suggests that individuals migrated due to the economic shock. Then, we constructed a measure of rainfall shock for each cluster, which is the deviation of the mean rainfall at the origin between

²³Initial location has six categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe.

²⁴Approximately 58% of original households have split since 1991.

1986 and 2000 from the average of all sample clusters during the same periods.²⁵²⁶ Similar to Beegle et al. (2008b), we interacted the measure with a dummy for being 5 to 15 years old in wave 1. Interacting between them is meaningful in two senses. Firstly, since migrants in our data largely consist of the young in wave 1, we might be able to reinforce the explanatory power of rainfall shock by interacting. Another merit is associated with one potential concern for using the shock as an instrument, which is that the instrument might be independently driving earnings in rural areas. If the rainfall shock is serially correlated, the exogeneity of the instrument will be violated in an earnings function in wave 5. However, since there is no strong reason to believe that the young who experienced the shock earned more or less than the old who experienced the same shock, we might be able to ensure the exogeneity by interacting the shock with the young age dummy. The second constructed instrument is a dummy for the female aged 5 to 15 years in wave 1. Since it is known from Table 2.4 that marriage is an important migration reason for females, this will have a strong explanatory power for migration, in particular, inside migration. The other two instruments - dummies for the household head and its spouse in wave 1 - again follow Beegle et al. (2008b). Since it seems that these people have a strong responsibility to maintain their house and social position in their rural communities, we expect that they are less likely to migrate than the other members of the household are.

On top of these instruments for migration, we included parents' education (years) in $\mathbf{z}_s$, which is often used as an instrument for education when estimating earnings functions. The reasoning is that parents' education affects the earnings of their children because it determines the schooling levels of their children. In practice, we again interacted it with a dummy for being 5 to 15 years old in wave 1. Since Table 2.7 shows that the young in wave 1 have acquired a large amount of education over time, the interaction between them is likely to increase the explanatory power of parents' education for their children's subsequent schooling.

²⁵Monthly rainfall information between 1980 and 2004, sourced from Tanzania Meteorological Agency, is publicly available from EDI Ltd. (Economic Development Initiatives) - <http://www.edi-africa.com/research/khds/introduction.htm#data>. For each cluster, we exploited rainfall information at either the nearest or (if not available) the 2nd nearest rainfall station. In the case where the rainfall information is not available for some cluster in some year, it was interpolated in this study by using the rainfall data in the nearest past or future.

²⁶The mean deviation of rainfall between 1991 and 2001 from the local average is used in Beegle et al. (2008b). While we do not have information on the local average, we also constructed our measure of rainfall shock by using the same reference periods. However, the alternative measure did not affect the subsequent empirical analysis.

(i) Checking the Validity of Instruments

Firstly, we informally checked the exclusion restriction of these six instruments in Table 2.9 by simply putting them into (2.12). We attempted several OLS specifications, depending upon whether outcome measures are income or consumption, whether migrants are pooled into one group or not, and whether education is included in the linear or quadratic form. Throughout an empirical analysis, we excluded remittance expenditures (urban-to-rural, rural-to-urban, rural-to-rural, and urban-to-urban) from consumption because they are just a monetary transfer among households.²⁷ The results show that all instruments are insignificantly different from zero with the exception of the dummies for household head and its spouse in earnings functions.

Next, Table 2.10 reports the estimation results of the reduced form equations of migration and schooling. Pooling migrants into one group, column (a), shows that the young in wave 1 who experienced the rainfall shock are more likely to migrate. As anticipated, being the household head or spouse in wave 1 reduces the probability of migration, and the young female in wave 1 show a greater propensity to migrate. These are broadly consistent with the first-stage estimation results of 2SLS conducted by Beegle et al. (2008b) to identify the average treatment effect (ATE) of migration on consumption growth. Dummies for inside and outside migrants were used as a dependent variable in columns (b) and (c), respectively. The results show that our instruments do not have strong explanatory power for outside migration, which disappointingly forces us to pool migrants into one group in the subsequent analysis.²⁸ Another point to note from columns (a) to (c) is that individuals are more likely to migrate if they are initially more educated (positive selection into migration). This is consonant with the equilibrium prediction 1 in the model in Section 2.2. A negative coefficient on a male dummy is somewhat surprising and seems to contradict the model presented in Section 2.2 if the access to higher income occupations is against the female. However, since the marriage of the female is an important migration reason in both inside and outside migration in our data, the negative coefficient is not necessarily inconsistent with our model. Also, the young in wave 1 are more likely to migrate.

The estimation result of schooling regression in column (d) is quite intuitive. The young in

²⁷However, key implications from the current study remained unchanged even though we include remittances.

²⁸This corresponds to a problem we encountered in Appendix 2.9.3, when the average treatment effects of inside and outside migration are separately estimated by 2SLS in Appendix Table 2.16 by taking the impact evaluation approach.

wave 1 who have educated parents have attained higher education in wave 5. In addition, the young and educated male in wave 1 has achieved a higher level of schooling in wave 5 and the effect is outstanding. The schooling effect implies that the change in education is higher for the initially educated (positive selection into schooling), which supports the equilibrium prediction 2 in the model in Section 2.2.

To sum up the estimation results in Tables 2.9 and 2.10, it seems that our instruments satisfy the exclusion restrictions relatively well (order condition) and have reasonable explanatory power for the variables of interest, as long as we pool migrants in one group (rank condition). All these encourage us to use them as instruments for identifying the structural parameters in an outcome equation.

(ii) Testing the Hypothesis

Table 2.11 reports the structural estimation results of (2.12), where migrants are pooled in one group because we do not have valid instruments that can convincingly account for outside migration (the rank condition fails). The first four columns present the estimation results of income regressions, whereas those of consumption are in the latter four columns. Since we found in Table 2.9 that being a household head or its spouse in wave 1 had a significantly negative impact on earnings in wave 5, those dummies were included in $\mathbf{z}_y$ in earnings functions.

Earnings

We firstly estimated an earnings function without education by OLS in column (a), and the result shows that migration results in an increase in income of 36 percentage points. Including education in column (b) reduced the return to migration and estimated the return to an extra year of schooling to be 4%. The reduction in the migration effect is consistent with the story that education partly explains the gains from migration.

However, a simple OLS in column (b) can be misleading when schooling and migration decisions are driven by earnings. Then, equation (2.12) is estimated by 2SLS in column (c). Controlling for the endogeneity further reduced the migration effect and raised the return to schooling. Now, 1 year of educational attainment adds about 9 percentage points to income, which is quite similar to the number often estimated in other studies (See Card (2001) for an extensive summary). In

addition, the observation that attempting to control for the endogeneity of schooling in earnings functions leads into a higher return than that in OLS is a common finding in the literature (Card, 2001). Regarding the migration effect, it seems that instruments for migration are somewhat weak, and the standard errors of the effect are very large. As a result, we are not able to say very much about the migration effect. The Sargan test disappointingly rejects the exogeneity of instruments.

Column (d) reports another 2SLS estimation result by using a different set of instruments, in which the pre-migration level of education is excluded from instruments. The result now easily satisfies the exclusion restriction, identifying the much larger return to schooling than that in column (c).²⁹ The big rise in the return to schooling might partly result from the declining power of instruments for education. Instruments for migration again appear not strong, and the migration effect is estimated with very large standard errors. As in column (c), unfortunately, we do not have much to say about the migration effect.

Consumption

In the absence of education, migration adds about 34 percentage points to consumption in column (e). Column (f) additionally includes education, but we now exploited the average level of household education, rather than the individual level of education. This is because our measure of individual consumption is the household average, and we expect that the per capita consumption correlates with the average level of household education more strongly.³⁰ Including the education reduced the migration effect from column (e), suggesting that migration in part operates through education.

Column (g) reports the 2SLS estimation result by instrumenting migration and schooling.³¹ In contrast to earnings functions, our benchmark set of instruments satisfies the exclusion restriction in consumption, and those instruments appear strong enough to explain migration and schooling. The estimated returns to migration and schooling are insignificantly different from those in OLS. This suggests that the endogeneity of migration and schooling is not a serious problem in consump-

²⁹Appendix Table 2.15 reports a range of 2SLS estimation results by using a different set of instruments, including the sets used in columns (c) and (d) in Table 2.11. The results satisfying the exclusion restriction identified almost identical schooling and migration effects with those in column (d) in Table 2.11.

³⁰When we use the individual level of education in consumption regressions, we find a significant but tiny schooling effect.

³¹While the average level of household education was used in a consumption equation, we exploited the same set of instruments as used in column (c) of earnings.

tion given the current controls. One implication from this finding is that initial household fixed effects, capturing characteristics specific to original households, have a strong power to control for the endogenous selection into migration and schooling. Similar to us, controlling for the initial household fixed effects, Beegle et al. (2008b) did not find a significant difference in the estimated migration effect between OLS and 2SLS, although they are estimating not the level but the growth of consumption (DID specification). If we are, then, willing to believe the OLS estimates in column (f), supposed to be unbiased and more efficient, the returns to migration and schooling are approximately 27% and 9%, respectively. If we accept that the migration effect unconditional on schooling is about 36%, as estimated in Beegle et al. (2008b), our point estimates suggest that about 25% of the effect is the schooling effect.

Finally, we jointly estimated the simultaneous equations of (2.12), (2.19), and (2.20) by 3SLS in column (h), which is expected to improve the efficiency of point estimates relative to 2SLS.³² The estimated returns to migration and schooling are quite similar to those in the OLS.

2.6.2 CF Approach

Consumption regressions in the IV approach showed that while education partly explained the gains from migration, a large significant migration effect still remained after controlling for the current level of education and addressing the endogenous selection into migration and schooling. Evidence on earnings was rather ambiguous due to the difficulty in controlling for the endogeneity, but the OLS still found a minor reduction in the migration effect in the presence of the current level of education. These findings suggest that education is important for migrants' outcomes but not very much.

However, the homogeneous return to schooling between the origin and the destination, an implicit assumption made in the IV approach might have prevented us from identifying the importance of education in determining migrants' outcomes. Then, we relax the assumption in this sub-section by estimating equation (2.11). Allowing for the heterogeneity in the return to schooling between places is very important. This is because the higher return at the destination is expected to generate the gains from migration if the migration effect is the schooling effect. The estimation

³²The estimation of 3SLS was conducted by using the command “*reg3*” in Stata. In the system, we used the average level of household education as an independent variable in an outcome equation and as a dependent variable in a schooling equation.

results of (2.11) are reported in Table 2.12 for earnings and in Table 2.13 for consumption, respectively. In both tables, panel (A) pools migrants into one group, whereas they are split into two groups in panel (B). The first column in each panel reports the migration effect, unconditional on education, for reference.

Earnings

Column (b) in Table 2.12 shows the OLS estimation result of (2.11). The interaction effect between migration and schooling is insignificantly different from zero, suggesting no difference in the return to schooling between the origin and the destination. The estimated returns to migration and schooling are almost identical with column (b) in Table 2.11. We addressed the endogeneity of migration and schooling decisions in column (c) by estimating equation (2.26).³³ We exploited the same set of instruments as used in column (c) in Table 2.11. Controlling for the endogeneity reduced the migration effect close to zero, but it seems that the control term for migration absorbed the effect. The 2SLS estimation results in Table 2.11 suggested that our instruments for migration were somewhat weak in earnings. The weak instrument problem might have generated the imprecise migration effect in column (c) in Table 2.12. The control term for schooling is significantly different from zero, suggesting the importance of controlling for the endogeneity, and pushed up the return to schooling from column (b) (0.041 to 0.108). But, the result does not support that migrants' education is more valuable than non-migrants' education. In column (d), we addressed the endogeneity only for schooling. The return to schooling rose from column (b) to the similar level in column (c), but the result again rejects the higher return to migrants' schooling.

Column (i) estimated (2.11) after splitting migrants into two groups. Since our instruments for migration do not have strong explanatory power for outside migration, we addressed the endogeneity only for schooling in column (j). As in panel (A), both results do not provide evidence to support that education is more valuable at the destination.

So far, all our attempts in earnings regressions have failed to find evidence to support that the return to schooling is higher at the destination. However, the finding might have some bearing on the linearity of education included and sample size. To the extent that the difference in the

³³The sample size is reduced in column (c). This is because the probit estimation for migration drops the observations of those who belonged to households in which all initial members did or did not migrate during the sample periods. The reduction in the sample size in column (c) in Table 2.13 is for the same reason.

schooling return between the origin and the destination is larger at the higher level of schooling and that data do not have a sufficient number of migrants in the upper part of schooling distribution, we might have failed to identify the difference in the return between places by assuming the constant marginal return.

Then, we assumed the linear marginal return to schooling by putting in the squared term of education. Pooling migrants, we estimated an earnings function in column (e). Now, the interaction effect is significantly different from zero, supporting the higher return to schooling at the destination at the mean level of migrants' education.³⁴ Point estimates suggest that migrants' 6-year education is 19% more valuable at the destination than at the origin, for example. The estimated migration effect is slightly reduced from column (b) in Table 2.11 (0.339 to 0.327), suggesting the importance of allowing for the difference in the schooling return between places in a non-linear form of education. Column (f) estimated equation (2.26). As in column (c), it seems that the migration effect is absorbed by the control term and, as such, that the point estimate is very doubtful. However, the result again supports the higher return to education at the destination. Addressing the endogeneity only for schooling reveals a similar picture in column (g). Education is more valuable at the destination, and the result implies that about 12% of the unconditional migration effect (0.361, column (a)) is the schooling effect.

Panel (B) conducted similar exercises after splitting migrants. Column (k) reports the OLS estimation result, whereas column (l) addresses the endogeneity of schooling by taking the CF approach. Both results suggest that the return to inside migrants' schooling is significantly higher than that to non-migrants' schooling at the mean level of inside migrants' education. While insignificantly different from zero, the point estimates on the interaction terms between education and outside migrants are quite similar to those on inside migrants' interaction terms. We might have lost the efficiency of the point estimates for outside migrants due to a small number of their observations in an earnings function.³⁵ In contrast to the unconditional migration effects in column (h), column (l) suggests that education explains approximately 10% (0.332 vs. 0.297) and 20% (0.547 vs. 0.435) of those effects for inside migration and outside migration, respectively.

³⁴The mean level of migrants' education is 5.75 years, with a standard error of 3.42 in wave 5.

³⁵Only 3.7% of 2562 observations is outside migrants.

Consumption

Pooling migrants into one group, we estimated (2.11) for consumption by OLS in column (b) in Table 2.13. In contrast to column (f) in Table 2.11, including the interaction between migration and schooling reduced the migration effect (0.269 to 0.198). The interaction effect is different from zero with 1% significance, suggesting that the return of migrants' extra year of schooling is higher than that of non-migrants by about 7%. These findings imply the fundamental importance of allowing for the difference in the return to schooling between the origin and the destination when estimating a consumption function. Addressing the endogeneity of migration and schooling, column (c) estimated equation (2.26). The result reports the migration and schooling effects insignificantly different from those in column (b), together with the estimates on control terms insignificantly different from zero. As in the case of the IV approach, these suggest that the endogeneity does not matter very much in consumption given our controls. Next, we addressed the endogeneity only for schooling in column (d). The estimated returns to migration and schooling are almost identical with those in column (b).

In column (e), we estimated the growth of consumption with the growth of education (from wave 1 to 5). If one source of the endogeneity of migration and schooling is attributed to time-invariant heterogeneity across individuals, we can eliminate it by taking the difference of equation (2.11). While the estimated migration effect is somewhat larger than that in column (b) (0.198 vs. 0.251), the result confirms that the return to migrants' schooling is significantly higher than that of non-migrants.

Migrants are split into two groups in panel (B). The OLS estimation result in column (g) shows that 1 year of both inside and outside migrants' educational attainment results in 6% higher consumption than non-migrants' attainment. Since we do not have good instruments to explain outside migration, we did not estimate (2.26) in panel (B). Instead, we addressed the endogeneity only for schooling in column (h). Controlling for the endogeneity of schooling hardly affects the result in column (g), identifying the schooling control term insignificantly different from zero. Finally, column (i) regressed the growth of consumption on the growth of education. The estimated returns to migration and schooling are insignificantly different from those in column (g). The result provides evidence to support that the return to migrants' schooling is higher than that

to non-migrants' schooling.

If we believe the OLS estimation results in a level equation, education accounts for approximately 40% of the unconditional migration effect for both inside and outside migrants, after allowing for the difference in the return to schooling between the origin and destination. This number is larger than the number (25%) estimated in the IV approach, which assumes the homogenous return. It seems that the difference in the return to education is an important component of the schooling effect.

2.6.3 Summary of the Estimation Results

In consumption, we provided much evidence to suggest that although not all, education was a part of the story to explain the gains from migration, and that education was more valuable at the destination. Assuming the homogeneous return between places, education is likely to explain about 25% of the migration effect. The number is approximately 40% in level OLS, after allowing for the difference in the return to schooling between places. These findings point to the notion that education indeed plays a role in explaining migrants' consumption gains, but it is not the main part of the mechanisms that benefit migrants. In addition, we did not find that the endogeneity mattered in consumption, suggesting the importance of controlling for original household heterogeneity.

Evidence on earnings was rather ambiguous for two main reasons: the seemingly existent endogeneity and the difficulty in controlling for it. The endogeneity is a problem in earnings, even though the current observables are an adequate control for the endogenous selection into migration and schooling. This is because our income measure contains large noise. Normally, measurement errors in a dependent variable just reduce the precision of point estimates. However, it seems that measurement errors are more serious in migrants' earnings than non-migrants' earnings (see Appendix 2.10.3 for the accuracy of per worker income) and, as such, those errors might have generated biases in earnings regressions because of their correlation with a migration dummy and the level of education.

Nevertheless, after allowing for the difference in the marginal return to schooling between places in a linear form (i.e., quadratic education) and controlling as far as possible for the endogeneity of schooling decision, education explained about 12% of the migration effect in earnings, supporting that migrants' education is more valuable than non-migrants' education. The explanatory power

of education is smaller than that for consumption. This might be suggesting that education affects migrants' consumption, other than through their earnings.³⁶ Or, it might be attributed to measurement errors in earnings. In any case, however, it appears true that education is not the main part of the story for explaining migrants' income gains.

2.7 Conclusion

While the impact evaluation can give us a reduced-form interpretation of migration effect, it leaves the important question of why not all people move out if the return to migration is so high. To improve our understanding of the mechanisms that benefit migrants (or limit mobility), the chapter investigated the extent to which education could explain migrants' income and consumption gains by using long-run household panel data from rural Tanzania. Measuring earnings, as well as consumption, and contrasting the two welfare measures are one important contribution of the chapter.

Identifying jointly the returns to migration and schooling in an outcome equation is a hard empirical task. This is because the outcomes of interest can affect both decisions to invest in migration and schooling. To address the endogeneity, the chapter took two approaches: the IV and CF approaches. At the cost of giving up identifying the difference in the return to schooling between the origin and the destination, an important component of our hypothesis, the IV approach provided evidence on the general importance of education for migrants' income and consumption by structurally addressing an individual's endogenous selection into migration and schooling. On the other hand, the CF approach identified the difference in the return to schooling between places by addressing the endogeneity in a non-structural way with strong parametric assumptions. Neither approach was perfect, but both approaches pointed to testing our hypothesis that education plays a major role in explaining the gains from migration. We found that individuals were positively selected into not only migration, but also further schooling based on their pre-migration level of education. The chapter provided some evidence that the return to migrants' education was higher than that to non-migrants' education. However, after controlling for the current level of education in outcome equations and allowing as far as possible for the endogenous selection into

³⁶For example, migrants' education might change their consumption bundles. If educated migrants prefer to consume high value-added goods, their education might independently drive high consumption, not earnings.

migration and schooling, we still found a large migration effect. This suggests that income and consumption opportunities are higher at the destination, and that some mechanisms, operating other than through education, prevent individuals from exploiting the arbitrage opportunity. While education plays a role, it does not appear to be a major factor in limiting the internal migration as a source of raising income and consumption.

The remaining migration effect might be explained by several possibilities. Simply, education might not be an appropriate measure of migrants' skills (more generally, human capital) valued in destination labour markets. Alternatively, as briefly reviewed in Section 2.1, the argument along the line of the HT-type models (urban wage institutionally set above the market-clearing and the equalization of wages across spaces in terms of expectation) is also one possibility. On the other hand, the risk-diversification argument might not provide a convincing account of the mobility observed in our data. One interesting hypothesis for the risk-sharing story is that migration does not necessarily mean the increase in migrants' consumption at the destination, even though they have achieved higher earnings. If income gains from migration are shared with their original families through the form of remittances, it is likely that migration leads into a smaller gain in consumption (exclusive of remittance expenditures) than in earnings. The relatively similar magnitude of migration effects identified in this study on both earnings and consumption might not provide a strong support for the view.

The role of social network might also be able to account for the remaining migration effect identified after controlling for education, as it seems that the network primarily benefits the less educated. Essentially, participating in such a nonmarket institution as the social network is a costly process, because the members are expected to make a commitment to assisting the others in the future (e.g., providing the subsequent migrants with job referrals, for example), when (currently) receiving a support from the network. Given the high likelihood that the educated (more generally, the talented) benefit less from the network but have to contribute more to it, therefore, it is certainly possible that they are reluctant to belong to the network due to the high cost of membership. In urban labour markets in Kenya, for example, Luke and Munshi (2006) indeed found that the high-ability individuals were more likely to defer entry into marriage-oriented kinship network. Since these discussions imply that the majority of network members is relatively low-ability individuals, the network might exclusively benefit the less educated, which can be consistent with the remaining

migration effect in this study.

Moreover, the remaining migration effect might have some bearing on an assortative matching in a marriage market because our data contain a relatively large number of female migrants for marriage. In the region of our analysis, women usually have a right to marry a man of their choice, as long as the parents approve of the marriage, and while the payment of bride price is still practiced, the amount is quite small, and the fear of refunding does not bind them by marriage when their husband mistreats them (Mitt and Rweyemamu, 2001). Since this implies that women are relatively free to find a better partner, the process might explain the return to migration.³⁷ All these alternatives are intriguing possibilities, and we leave probing them for future research.

We conclude the chapter with a few suggestions on the survey practice. At first, it might be better to have relatively longer reference periods to collect reasonable income data from urban occupations because access to those jobs seems to be generally episodic. Therefore, instead of just one or two weeks, one month might be better. In addition, we reconfirmed the common view of the difficulty in measuring income from non-farm self-employed activities in household surveys in less developed economies. This should continue to be a major concern in empirical research.

³⁷In the literature, marriage is a popular mechanism to explain individuals' mobility, which is associated with a risk coping strategy. For instance, women's marriage cum migration, which accounts for a significant proportion of migration in India, can be seen as an interhousehold contractual arrangement to mitigate income risks and to facilitate consumption smoothing (Rosenzweig and Stark, 1989).

2.8 Appendix: Model in Section 2.2

2.8.1 Proof

Proof of Lemma 1: Suppose the contrary. If an agent of type $\bar{s}'$ prefers to migrate and all agents $\bar{s} > \bar{s}'$ are weakly better-off in the rural sector, it must be that $v^*(\bar{s}') > 0$ and $v^*(\bar{s}) \leq 0$. This means that $v^*(\bar{s}') > v^*(\bar{s})$. Denoting $s^*(\bar{s}')$ as $s^{*'}$ and $s^*(\bar{s})$ as s^* , it can be shown that for all $\bar{s} \in \left[\frac{m-\rho}{\gamma}, \hat{s}\right]$,

$$v^*(\bar{s}') - v^*(\bar{s}) = (s^{*'} - s^*)[2\alpha - \beta(1+r)(s^{*'} + s^*)] + \beta(1+r)(\bar{s}' + \bar{s})(\bar{s}' - \bar{s}). \quad (2.27)$$

From Figure 2.1, it is clear that $s^{*'} < s^*$ and $\frac{s^{*'} + s^*}{2} < \frac{\alpha}{\beta(1+r)}$. Hence, this results in $v^*(\bar{s}') < v^*(\bar{s})$, which is a contradiction.

Proof of Proposition 1: In order to show that all agents $\bar{s} \geq \frac{m-\rho}{\gamma}$ optimally choose to migrate, it is enough to show that an agent $\bar{s} = \frac{m-\rho}{\gamma}$ is willing to migrate by lemma 1.

$$v^*\left(\frac{m-\rho}{\gamma}\right) = \alpha \left[1 + \frac{2(m-\rho)}{\gamma}\right] - (1+r)m. \quad (2.28)$$

Under our assumption, this is positive. Hence, all agents $\bar{s} \geq \frac{m-\rho}{\gamma}$ prefer to migrate. Also, it is straightforward that agents $\bar{s} < \frac{m-\rho}{\gamma}$ stay in the rural sector because they cannot afford the cost of migration.

To claim that this is the mobility in equilibrium, we have to ensure that parameters exist to satisfy all preceding restrictions: (a) $\frac{m-\rho}{\gamma} < \frac{\alpha}{\beta(1+r)}$, (b) $\frac{m-\rho}{\gamma} < \bar{s}_u \leq \frac{-\gamma + \sqrt{\gamma^2 - 4\beta(\rho-m) + (\frac{2\alpha}{1+r})^2}}{2\beta}$, and (c) $(1+r)m < \alpha \left[1 + \frac{2(m-\rho)}{\gamma}\right]$. Since we have seven unknown parameters $(\alpha, \beta, \gamma, \rho, r, m, \bar{s}_u)$, it is possible to find parameters that simultaneously satisfy these restrictions.

2.8.2 Comparative Statics of the Model

This sub-section conducts standard comparative statics in equilibrium presented in Section 2.2.

Firstly, the increase in the return to schooling allows migrants to obtain higher earnings ($dy_u^*/d\alpha > 0$), although it does not change their educational attainment and, thus, occupations ($d\Delta s^*/d\alpha = 0$, $dq^*/d\alpha = 0$) because it does not relax their credit constraint in this model. Both the decrease

in the marginal educational cost and the increase in the initial level of wealth enable migrants to acquire more additional skills and to choose higher income jobs by relaxing their credit constraint ($d\Delta s^*/d\beta < 0$, $dq^*/d\beta < 0$, $dy_u^*/d\beta < 0$, $d\Delta s^*/d\rho > 0$, $dq^*/d\rho > 0$, $dy_u^*/d\rho > 0$). Another interesting exercise is to examine how the change in the underlying parameters affects the population of migrants, $M_p = 1 - G(m - \rho/\gamma)$. For instance, the increase in the migration costs reduces the population of migrants because it requires them to hold the higher level of initial wealth to move out ($dM_p/dm < 0$). We might also want to investigate how the total labour income Y in the economy responds to the change in the parameters. In the equilibrium discussed in Section 2.2, it can be shown that

$$Y = \delta \bar{s}_u + \alpha \left[\bar{s}_u - \frac{m - \rho}{\gamma} \right] + 2\alpha \int_{\frac{m-\rho}{\gamma}}^{\bar{s}_u} \sqrt{\frac{\beta \bar{s}^2 + \gamma \bar{s} + \rho - m}{\beta}} d\bar{s}. \quad (2.29)$$

Because of the complexity of the last term, it is difficult to solve this by hand. The most that can be said is that the total labour income will increase as the returns to schooling and self-employed farming increase ($dY/d\alpha > 0$, $dY/d\delta > 0$). For these discussions to hold, note that the shift of parameters should be marginal. If the external shocks that hit the economy drastically change the parameters, the economy is likely to move away from the current equilibrium to another.

In addition, the model shows that the underlying parameters in the economy have not only direct influence on migrants' occupational choice and earnings, but also indirect influence by stimulating migrants' investment in human capital. To see this formally, we can write them as

$$dq^* = \sum_{\Phi_i \in \Phi} \frac{ds^*}{d\Phi_i} d\Phi_i, \quad (2.30)$$

$$dy_u^* = d\delta + (2s^* + 1)d\alpha + 2\alpha \sum_{\Phi_i \in \Phi} \frac{ds^*}{d\Phi_i} d\Phi_i, \quad (2.31)$$

where $\Phi \equiv (\alpha, \beta, \gamma, \delta, \rho, r, m)$. The right hand side of (2.30) captures the indirect impact on migrants' occupational choice through educational attainment. The first two terms in equation (2.31) document the direct effect on migrants' earnings and the last term captures income change attributed to human capital accumulation.

2.8.3 Some Thoughts on Modelling Migration

Our model has several general caveats. Firstly, it ignores the general equilibrium effect arising from migration. Suppose that labour markets are perfectly competitive, the increase in rural-to-urban migration is likely to decrease earnings from urban jobs and to increase them from rural jobs. Carrington et al. (1996) incorporated this effect into their migration model with endogenous moving costs, where the population of migrants in the destination from the same region or ethnic group reduces the cost of migration.

In addition, it is obviously unrealistic to assume away unemployment (Todaro, 1969; Harris and Todaro, 1970). It is not so difficult just to introduce the stochastic component into the model, but doing so in a reasonable way would be a relatively hard task because the probability of obtaining urban jobs must be endogenous in reality. The increase in the population of migrants might reduce the likelihood of each migrant being employed in cities through intense competition between them. Alternatively, the expansion of migrant networks in the urban economy might help each migrant easily find a job (Benerjee, 1983; Andersson, 2001; Munshi, 2003; McKenzie and Rapoport, 2007; Yamauchi and Tanabe, 2008). Moreover, the schooling level of migrants will also affect the probability of employment if education works as a signal of their ability.

While our model assumed that the returns to schooling (α) and self-employed farming (δ) were constant, this could easily be amended. For instance, the return to self-employed farming might be an increasing function of the level of wealth (e.g., land) as assumed in McKenzie and Rapoport (2007). A significantly higher return to self-employed farming is likely to make rural agents in the upper tail of the wealth distribution more strongly hesitant to migration by increasing their opportunity costs of moving out. On the other hand, the return to schooling might increase with the level of agents' schooling.

Further considerations are possible. While our model perfectly regarded individuals as a decision-making unit, viewing families as one will make the decision to migrate more complicated. We do not consider agents' risk-preferences, income risk, and their response to it. Instead of a static model, dynamically treating the economic mobility associated with migration, together with uncertainty about the urban environment might be able to explain the adjustment process of migration (move out and return) through which the economy eventually reaches the stable equi-

librium. All the above and possibly more ingredients in the process of migration are missing in our model. However, our model is not against them and, rather should be regarded as a complement that sheds much light on the role that educational attainment plays in the process of migration.

2.9 Appendix: Impact Evaluation Approach

To the extent that our interest just lies in identifying the gains from migration, we can take an impact evaluation approach that has been applied to many important questions in labour economics since Lalonde's (1986) influential work and has recently been applied to migration studies (McKenzie et al., 2006; Beegle et al., 2008b). This section briefly reviews the literature along the line and identifies the gains from migration by following the strategy adopted by Beegle et al. (2008b).

2.9.1 Literature Review

One traditional view of migration is that it is an investment in human capital (Shultz, 1961; Sjaastad, 1962; Bowles, 1970; Laber and Chase, 1971). Workers move from one place to another, where the return to their human capital is higher at the cost of movement. In this view, the cost can be seen as a form of human investment to increase the productivity of human resources that each person is endowed with. According to the framework, earlier empirical studies attempted to examine if migrants were gaining from migration (Wertheimer, 1970; Yezer and Thurston, 1976; Grant and Vanderkampt, 1980). However, the findings from those studies, which usually estimate the effect of migration by using cross-sectional data and resorting to a simple OLS technique, are unconvincing.

The key difficulty in identifying the gains from migration is to estimate counterfactual earnings that migrants would have achieved had they not moved out. While they are often proxied by the earnings of non-migrants, such proxy is improper if migrants are self-selected from their origins (Robinson and Tomes, 1982; Borjas, 1987, 1994; Barham and Boucher, 1998; Lanzona, 1998; Hartog and Winkelmann, 2003; Chiquiar and Hanson, 2005; McKenzie et al., 2006). Suppose that migrants are, for example, more able, more skilled, and more motivated than non-migrants are, the difference in earnings between migrants and non-migrants might just be reflecting the difference

in such unobserved quality rather than the gains from migration and, therefore, simply comparing the earnings of migrants with those of non-migrants overestimates the effect of migration.

Indeed, such positive self-selection is often supported by empirical evidence from the study of internal migration. For instance, Barnum and Sabot (1977) used population data in Tanzania to argue that educated workers have a greater propensity to migrate from rural to urban areas, and that part of the reason could be attributed to the economic returns to migration, which increases with the level of schooling. Lanzona (1998) found that the (wage) returns to schooling and work experience in rural areas were biased downwards if self-selected out-migration was not controlled for by exploiting panel data in the Bicol region of the Philippines. The positive relationship between migrants' educational background and mobility, together with the important role of language skills in making a decision to migrate, was observed in many provinces in Canada (Robinson and Tones, 1982).

In the study of international migration, the literature has also found another type of selection. The pioneering work of Borjas (1987, 1994) presented a theoretical mechanism that can explain whether it is the skilled or less skilled in source countries that migrate. The answer depends upon the correlation of earnings (hence, skills) between source and host countries and the earnings dispersion (not, mean earnings) in the two countries. As one likely case to explain immigrants flow to the United States, he argued that where skills were transferable across countries and the source country had higher earnings inequality than the host country, immigrants were self-selected from the lower tail of skills distribution in the source country (negative selection). Chiquiar and Hanson (2005) took Borjas's discussion a step further by assuming that the cost of migration decreases with the initial level of schooling. Not just positive or negative selection, they showed the possibility that those who had an intermediate level of schooling at the origin were more willing to migrate. Their empirical analysis of Mexican migrants in the United States provided evidence consistent with the intermediate selection.

All these studies suggest that migrants' self-selection is such a critical problem that we are not able to identify the gains from migration without controlling for it. To our knowledge, McKenzie et al. (2006) is one study that successfully identifies the (income) gains from migration and the most instructive paper to understanding the extent to which standard econometric techniques can adequately control for migrants' self-selection. They exploited a natural experiment in which only

those who won the lottery, which was used by the New Zealand Department of Labour (DoL) in order to randomly select Tongan migrants from registered applicants that outnumbered quotas, were allowed to migrate into New Zealand. The data contain three groups of people: (i) lottery winners (treatment group), which are further disaggregated into those who actually migrated (compliance group) and did not (noncompliance group); (ii) lottery losers (control group); and (iii) the representative sample of non-applicants. They estimated the gains from migration (average treatment effect on the treated or local average treatment effect) by instrumenting a migrant dummy with a lottery-winning dummy in data containing groups (i) and (ii), and then attempted to identify the effect by using several standard econometric techniques in data containing migrants and the group (iii), which resembled non-experimental settings that were usually seen in other studies. They found that registered applicants (hence, migrants as well) were positively selected in terms of both observed and unobserved skills, which hence generated a significantly large upward bias in a simple OLS estimate, and that IV estimation using good instruments, such as the distance to the nearest office of the DoL, worked best to identify the gains from migration. In addition, differences-in-differences (DID) and bias-adjusted propensity-score matching performed so well that they produced a migration effect not significantly different from an experimental point estimate.

Beegle et al. (2008b) also took the impact evaluation approach to identify the gains from migration by exploiting data drawn from the same information source as ours. By pooling migrants into one group and estimating a DID specification by OLS and 2SLS, they found that moving out of original communities on average added 36 percentage points to consumption growth from waves 1 to 5.

2.9.2 Structural Interpretation of Impact Evaluation Approach

The basic specification of impact evaluation approach can be derived by using our structural model of simultaneous equations (2.12), (2.19), and (2.20). Given our system, a simple algebra yields³⁸

$$\begin{aligned} \log y_{it+1} = & \left(\frac{\alpha_0 \delta_1 + \alpha_2 (\gamma_0 + \tilde{\beta}_0 \gamma_3)}{\delta_1 - \alpha_2 \delta_2} \right) + \left(\frac{\alpha_1}{1 - \alpha_2 \left(\frac{\delta_2}{\delta_1} \right)} \right) M_{it} + \left(\frac{\alpha_2 [\gamma_1 + \gamma_3 (\beta_4 - \beta_3)]}{\delta_1 - \alpha_2 \delta_2} \right) s_{i0} \\ & + \left(\frac{1}{1 - \alpha_2 \left(\frac{\delta_2}{\delta_1} \right)} \right) \alpha'_y \mathbf{z}_y + \left(\frac{\alpha_2 \gamma_3}{\delta_1 - \alpha_2 \delta_2} \right) \beta'_m \mathbf{z}_m + \left(\frac{\alpha_2}{\delta_1 - \alpha_2 \delta_2} \right) \gamma'_s \mathbf{z}_s + \left(\frac{\alpha_2 (\gamma_3 \tilde{v}_{it} + \tilde{\omega}_{it+1}) + \delta_1 \epsilon_{it+1}}{\delta_1 - \alpha_2 \delta_2} \right), \end{aligned} \quad (2.32)$$

³⁸We firstly substituted (2.20) into (2.19) to solve for s_{it+1} . Then, we put the equation into (2.12) to solve for $\log y_{it+1}$.

where $\delta_1 \equiv 1 - \beta_3\gamma_3$ and $\delta_2 \equiv \gamma_2 + \beta_2\gamma_3$. Assuming that $\mathbf{z}_y$ is time-invariant for expositional simplicity, we can also derive by subtracting $\log y_{i0}$ ($= \alpha_0 + \alpha_1 s_{i0} + \alpha'_y \mathbf{z}_y + \epsilon_{i0}$ because the period 0 is a pre-treatment period) from (2.32),

$$\begin{aligned} \Delta \log y_{it+1} &= \left(\frac{\alpha_2(\gamma_0 + \beta_0\gamma_3 + \alpha_0\delta_2)}{\delta_1 - \alpha_2\delta_2} \right) + \left(\frac{\alpha_1}{1 - \alpha_2\left(\frac{\delta_2}{\delta_1}\right)} \right) M_{it} + \left(\frac{\alpha_2(\gamma_1 + \beta_1\gamma_3 - 1) + \alpha_2^2\delta_2}{\delta_1 - \alpha_2\delta_2} \right) s_{i0} \\ &\quad + \left(\frac{\alpha_2\left(\frac{\delta_2}{\delta_1}\right)}{1 - \alpha_2\left(\frac{\delta_2}{\delta_1}\right)} \right) \alpha'_y \mathbf{z}_y + \left(\frac{\alpha_2\gamma_3}{\delta_1 - \alpha_2\delta_2} \right) \beta'_m \mathbf{z}_m + \left(\frac{\alpha_2}{\delta_1 - \alpha_2\delta_2} \right) \gamma'_s \mathbf{z}_s + \left(\frac{\alpha_2(\gamma_3 \bar{v}_{it} + \bar{\omega}_{it+1} + \delta_2 \epsilon_{i0}) + \delta_1 \Delta \epsilon_{it+1}}{\delta_1 - \alpha_2\delta_2} \right), \end{aligned} \quad (2.33)$$

where Δ denotes the difference between period 0 and t . These are the standard specifications often used in the impact evaluation approach that attempt to identify the average treatment effect (ATE) of migration conditional on several exogenous and pre-treatment variables. Equation (2.33) is sometimes called a DID specification (McKenzie et al., 2006; Beegle et al., 2008b). The theoretical ground of DID is as follows. As described above, it is expected that whether one moves out or not is driven by some unobserved difference between migrants and non-migrants, which also makes the distinction between their earnings. If this is the case, simply comparing the earnings of migrants with those of non-migrants in equation (2.32) fails to identify the true impact of migration. However, if we assume that the unobserved difference (in other words, selection bias) is constant over time, the DID estimation, which exploits the panel dimension of data, can control for such a problem because the observed growth of non-migrants' earnings provides us with the counterfactual earnings growth that migrants would have achieved had they not moved out.

Equations (2.32) and (2.33) help us to understand an empirical problem associated with the impact evaluation approach under our structural model. Given the simultaneity of endogenous variables, the ATE (a coefficient on a migration dummy) cannot be identified in the absence of good instruments because M_{it} is obviously correlated with error terms. This is true, even in DID.³⁹ Then, what we need to identify the average treatment effect is $\mathbf{z}_m$ variables, which can be excluded from an earnings function. This corresponds to the order and rank conditions in our IV approach.

³⁹One theoretically possible exception is the case where $\alpha_1 = 0$. In this case, we obtain $\Delta \log y_{it+1} = \alpha_2 M_{it} + \Delta \epsilon_{it+1}$ as a DID specification. If $\epsilon_{it+1} \equiv \eta_i + \bar{\epsilon}_{it+1}$, the DID can eliminate the individual fixed effect (η_i) that might simultaneously be driving M_{it} and $\log y_{it+1}$. However, it is very unlikely that there is no impact of schooling on earnings in practice.

2.9.3 Estimation Results

Specifications (2.32) (level in wave 5) and (2.33) (growth from wave 1 to 5) are applied to our measures of income and consumption in Appendix Table 2.16, where migrants are pooled in one group in panel (A) and split into two groups in panel (B). Following Beegle et al. (2008b), we explicitly controlled for original household heterogeneity by putting in dummies for initial households, together with the general controls of age, education, and a male dummy at baseline (wave 1). Controlling for the initial household fixed effects (IHHFE) means that the impact of migration is identified based on a sample of households that have split since 1991 (within initial household variation). Columns (a), (d), (g), and (j) report the OLS estimation results of (2.32), whereas a DID specification of (2.33) is estimated in columns (b), (e), (h), and (k) by OLS, and in columns (c), (f), (i), and (l) by 2SLS instrumenting a migration dummy. In 2SLS, we used the four instruments explained in Section 2.6: rainfall shock in an original cluster (interacted with a young age dummy in wave 1), dummies for the household head and its spouse in wave 1, and a young female dummy in wave 1. These are similar to the set of instruments exploited by Beegle et al. (2008b).

Let us start with reporting the results for consumption. In order to estimate the impact of migration on the growth of consumption more accurately, we excluded remittance expenditures (urban-to-rural, rural-to-urban, rural-to-rural, and urban-to-urban) from consumption because they are just a monetary transfer among households.⁴⁰ By pooling migrants in one group, column (d) shows that migration has added 34 percentage points to consumption growth over 13 years with 1% significance. The OLS estimation result of DID in column (e) produced a surprisingly (and somewhat weirdly) similar coefficient.⁴¹ As discussed in Section 2.9.2, these estimation results could, however, be biased given the simultaneous choice of schooling, migration, and occupations (hence, income or consumption). Then, we instrumented the migration dummy in column (f) and the 2SLS identified an almost identical point estimate, suggesting that a migration dummy is exogenous to stochastic errors conditional on controlled observables. Indeed, this is the same finding as Beegle et al. (2008b). The instruments are strong enough and the exogeneity is satisfied by the Sargan test. While these numbers are slightly smaller than the point estimate of Beegle et

⁴⁰However, key implications were almost unaffected by including remittances.

⁴¹If we drop the initial household fixed effects, the level OLS and the DID OLS yielded 0.359 and 0.312 as a point estimate of migration effect, respectively.

al. (2008b), a standard T-test did not reject a hypothesis that they are equal to 0.36. As a result, our pooling estimation results recovered the almost identical migration effect with Beegle et al. (2008b). In columns (j), (k), and (l), migrants are split into inside and outside. The level OLS in (j) and the DID OLS in (k) show that the effect of outside migration is much larger than that of inside migration, suggesting that moving out of Kagera (probably, into urban areas) has a much higher value than just moving out of original communities. When these two migration dummies were instrumented in column (l), the migration effects were very poorly identified. This result is attributed to the fact that our instruments do not have strong explanatory power for outside migration. This issue is mirrored in Table 2.10, when estimating the reduced form of inside and outside migration equations.

By pooling migrants, the level OLS in column (a) shows that migration results in an increase of 41 percentage points in income growth. The migration effect is reduced to 34% by using a DID specification in column (b), suggesting an upward bias in a level OLS (positive selection into migration). The migration dummy was instrumented in column (c), where instruments are somewhat weaker than in a consumption regression but the exogeneity is again satisfied by Sargan test. As opposed to a consumption regression, the point estimate declined further from the DID OLS, although the effect became insignificantly different from zero, probably due to the increase in the standard errors. The finding that 2SLS reduces the migration effect in earnings but not in consumption is consistent with the view that migrants' earnings are measured with noise and that the current observables in DID are an adequate control for the other sources of bias not associated with the measurement errors. When migrants are split into inside and outside, columns (g) and (h) ensure the importance of moving out of Kagera as in consumption regressions. As seen in a consumption regression, since we do not have good instruments that explain outside migration, the 2SLS estimation result in column (i) very poorly identifies the migration effects.

2.10 Appendix: Income Measurement Issues

While measurement errors are always an inevitable concern for researchers who conduct empirical studies, they might not necessarily detract from the comparability of measured variables between individuals and over time when there is little heterogeneity in the behaviour of individuals. This

is likely to be the case in waves 1 to 4 in our data. However, taking into account the significant difference in the pattern of individual labour allocation between migrants and non-migrants in wave 5 and the occupational mobility towards non-farm activities in all groups; it seems that small measurement errors in earnings from a particular activity in wave 5 make us lose the comparability of total labour income between individuals and waves. As this forces us to be very cautious about correctly measuring earnings in wave 5, this section explains potential income measurement issues in our data.

2.10.1 Suggestive Evidence on Latent Income

The survey questionnaire allows us to construct two possible income measures for each individual: per worker weekly income and per capita weekly income (average of the total labour income of household), both of which are computed as the sum of wage earnings, farming self-employment income, and non-farm self-employment income. As the reference periods of recorded income were different between income generating activities in the questionnaire: 1 week for wage employment, 1 year (or 6 months in waves 2 to 4) for farming self-employment, and 2 weeks for non-farm self-employment, we simply took the time average to compute weekly income. As opposed to wage earnings that respondents aged 7 years or older individually reported, both farming and non-farm self-employment incomes were collected on a household basis. Thus, in per worker income measure, they were assigned to each household member in proportion to hours worked on those activities in the past 7 days, which were individually reported by respondents aged 7 years or older in a different section from sections collecting household-level self-employment income.⁴² As a result, per worker income measure is constructed only for panel individuals aged 7 years or older.

Given the way of constructing income measures, one potential concern is that we might systematically fail to capture household workers engaged in wage employment and non-farm self-employment due to the difference in the periodicity between income generating activities in the survey questionnaire. For example, if one does not happen to engage in those activities at the point of survey, we completely lose his contribution to household income, even though he usually makes a living from waged jobs or non-farm self-employed activities. If this is the case, we will underestimate household income relative to what it should be, as the household's main economic

⁴²Thus, adult and child workers are equally weighted in terms of their efficiency in the analysis.

activities concentrate more on wage employment and non-farm self-employment and as the access to such activities becomes more irregular. Taking into account the occupational mobility towards non-farm activities and generally episodic urban occupations, it is not unreasonable to think that we might fail to capture the full extent of workers available to (in particular, outside) migrants' households in wave 5.

In order to check the possibility, Appendix Table 2.20 highlights the extent to which individuals aged 23 to 40 years, who are potentially the active workers of a household, diversified their income generating activities in the past 7 days by exploiting the information on hours worked that they reported. It is remarkable that about 27% of outside migrants engaged in no activities in wave 5.⁴³ Many possibilities in the urban sector can explain the large proportion of non-working outside migrants. For example, this might suggest the generally higher rate of unemployment in cities, or urban workers' preferences for jobs with higher wage rate but fewer hours worked required. It is also consistent with the fact that urban workers have more channels to earn non-labour income. However, if this is an artificial result reflecting the periodicity problem, it turns out that we are failing to capture the true activities of each urban household.

In addition to the periodicity problem, a relatively large number of households did not report the fortnight's income from non-farm self-employed businesses in wave 5 despite some of their members reporting positive hours worked on those activities in the past 7 days. This is clear from comparing the number of non-farm self-employed observations in wave 5 between rows (3) and (4) in Appendix Table 2.18, which presents descriptive statistics at the household level. Hereinafter, including the table, households are defined as migrants in waves 1 to 4 if they supply at least one member who moved out of his original community between 1991 and 2004, and in wave 5 if they contain at least one panel migrant as a member. For example, it is critical that only 11 households of outside migrants reported income from non-farm self-employment, whereas 83 households actually engaged in it.⁴⁴

⁴³We also conducted the same exercise at the household level in Appendix Table 2.21, where about 38% of outside migrants' households engaged in no activities in wave 5. However, the number might be exaggerated because outside migrants' households are relatively young (see (12) in Appendix Table 2.18) and, thus, likely to contain more children who are not working normally. Thus, we believe that Appendix Table 2.20 is more informative because it removes the influence of the difference in the age structure of the household between groups. Moreover, a similar exercise was conducted for individuals aged 20 to 40 years and 15 to 40 years. As expected, stretching the grouping age as a whole increased the proportion of non-workers in all groups and waves. For example, the proportion of non-working outside migrants in wave 5 was about 32% and 40% for 20 to 40 and 15 to 40 age groups, respectively.

⁴⁴This problem is also reflected in the difference in the number of observations between migrants' per worker weekly income and per worker weekly hours worked in wave 5 in Appendix Table 2.17.

All the above suggest that we might be underestimating the earnings of (outside) migrants' *household* in wave 5. To see this more formally, we regressed household (annual) consumption on household (weekly) income by group with the control of remittances received and other non-labour income earned in the past 6 months in Appendix Table 2.22.⁴⁵ Suppose that a household's income is correctly measured, we expect that the elasticity of consumption with respect to labour income is almost identical in all groups (and, probably, over time). While the estimated coefficients are relatively stable between groups in waves 1 to 4, we see the large variation in columns (d) to (f) in wave 5. The elasticity is much smaller in the migrants' regressions in wave 5 and, in particular, the null of zero correlation is not rejected in outside migrants. The regression results imply that we are seriously failing to capture the potential income of migrants' households in wave 5. In other words, it is highly likely that household income (latent income) exists that is not contained in our income measure.

2.10.2 Empirical Test for Latent Income

The preceding discussions provided suggestive evidence that we might be underestimating the income of household in wave 5 and that the problem is more serious in migrants' earnings. To support this claim further, we estimated the per capita (annual) consumption of panel individuals (the average of the total annual consumption of household) in wave 5 with the per capita (weekly) income in Appendix Table 2.23, controlling for potential sources of bias in several ways. Suppose that we are failing to capture the income of migrants' households in wave 5, controlling for an attenuation bias associated with the problem, as well as biases arising from other sources, should substantially correct the correlation between income and consumption.

Testing Strategy

Our specification for per capita consumption in wave t is

$$\log c_{iht} = \phi_0 + \phi_1 \log y_{iht}^{c*} + \delta_t + \eta_h + \epsilon_{iht}, \quad (2.34)$$

⁴⁵One might want to say that the primary income source of migrants' households in the urban sector is non-labour income rather than labour income. Indeed, about 38% of 145 outside migrants' households, which received any remittances (gift or loan) in the past 6 months in wave 5, did not report any labour income in our data, in contrast to only about 2% of 1192 non-migrants' and about 8% of 789 inside migrants' households, which received remittances. This might support the substitutability between labour and non-labour income for outside migrants, although this can also be explained by the latent (labour) income problem discussed in this section.

where c_{iht} and y_{iht}^{c*} are the true per capita consumption and the income of an individual i in a household h , respectively. We assume that $y_{iht}^c = y_{iht}^{c*} \cdot \exp(-u_{iht})$, where y_{iht}^c is observed per capita income, and a variable u_{iht} represents the latent income or simply a measurement error in observed income. Since the measurement error is expected to be much more serious in wave 5, we assumed that $u_{iht} = 0$ in the earlier four waves. Any unobserved time-invariant household characteristics that determine household consumption (and probably household income simultaneously) are captured by η_h , and a variable δ_t captures a wave-specific common shock and ϵ_{iht} is a random error term with mean zero.⁴⁶ We assume that observable per capita income is strictly exogenous with respect to this random error term.

Two main empirical challenges in estimating per capita consumption in wave 5 become clearer to see in the following tractable model: an unobserved heterogeneity across households (η_h) and a measurement error (u_{ih5}), both of which are expected to be correlated with per capita income ($\log y_{ih5}^c$):

$$\log c_{ih5} = \phi_0 + \phi_1 \log y_{ih5}^c + \delta_5 + \phi_1 u_{ih5} + \eta_h + \epsilon_{ih5}. \quad (2.35)$$

Our strategies to control for them can mainly be classified into two groups: (a) a directly estimating equation (2.35), instrumenting the per capita income and (b) conducting 2SLS estimations after eliminating the unobserved heterogeneity.

For the former strategy (strategy (a)), we firstly instrumented per capita income in wave 5 with per worker income in the same wave and dummies for the (original) locations in wave 4,⁴⁷ although this restricts the observations exploited in the estimation only into panel workers rather than panel individuals. The correlation of instruments must be strong because both income measures are constructed from the same information source, and the pre-determined location seems to be orthogonal to both the measurement error in the current income and the unobserved heterogeneity. On the other hand, the exogeneity of per worker income is somewhat doubtful. If it is also measured with errors, it is likely that measurement errors in both income measures are correlated with each other. Moreover, it is not unreasonable to think that the current household heterogeneity

⁴⁶However, δ_t cannot be identified in the subsequent regressions in this sub-section because our estimations are based upon observations in a single wave.

⁴⁷This restricts migrants into those who moved out of their original communities after wave 4 because those who migrated before wave 4 were not surveyed in wave 4.

influences the current per worker income. However, if the correlation of errors between the two income measures is mitigated in the process of constructing each measure and the unobserved household heterogeneity is not a serious problem, it might work as a good instrument. We treat this set of instruments as a place to start.

In the next set of instruments for both inside and outside migrants, we replaced the per worker income by the first-lag of per capita income ($\log y_{ig4}^c$, where g denotes a household in the earlier four waves, and $h \neq g$ for migrants). Since it ($\log y_{ig4}^c$) is determined by the characteristics of migrants' original household, it seems to be uncorrelated with both the heterogeneity of their current household and the measurement error in their current income. Potential concern for using it as an instrument is that moving out of original communities might break the correlation between the lagged income in their previous household and the current income in their new household. A similar exercise is not reported for non-migrants because most of them stayed in the same household during sample periods ($h = g$),⁴⁸ and it is likely that the time-invariant household heterogeneity drives not only their current per capita income, but also their previous income.

The key of the second strategy (strategy (b)) is to eliminate the unobserved household heterogeneity before conducting 2SLS estimations, and a standard procedure to do it with a panel dimension is to take the first-difference of equation (2.35). In the first place, we conducted this exercise for non-migrants. While this can eliminate the heterogeneity, the measurement error remains in the first-differenced equation. Since it is potentially correlated with the first-differenced per capita income ($\Delta \log y_{ih5}^c$), the estimated income-consumption correlation would be biased if we do not control for it. To address the issue, we instrumented the first-differenced per capita income with per capita income ($\log y_{ih3}^c$) and household size in wave 3. Since it seems that per capita income in the earlier four waves has strong serial correlation because of short time-intervals between waves, the per capita income in wave 3 will be correlated with the first-differenced per capita income ($\Delta \log y_{ih5}^c$) through its correlation with per capita income in wave 4. The exogeneity of the set of instruments is ensured if we believe that both per capita income and household size more than 10 years ago have no correlation with the measurement error in wave 5.

While taking the first-difference is a standard way, this cannot directly be applied to migrants

⁴⁸Note that $h \neq g$ even for non-migrants if they moved out of their original household to form their new household within the same community after wave 4, although there are not many.

because the household that they belonged to in wave 5 is different from their original household, and simply taking the first-difference cannot eliminate the time-invariant household heterogeneity. Thus, after rearranging equation (2.35), we estimated the following transformed consumption by using 2SLS for migrants:

$$\log c_{ih5} - \Delta \log c_{ig4} = \text{constant} + \phi_1(\log y_{ih5}^c - \Delta \log y_{ig4}^c) + \phi_1 u_{ih5} + \eta_h + (\epsilon_{ih5} - \Delta \epsilon_{ig4}), \quad (2.36)$$

where h and g denote a household that a migrant i belonged to in wave 5 and in the earlier four waves, respectively.⁴⁹ This specification allows us to eliminate the unobserved heterogeneity of an original household (η_g), not that of a current household (η_h). However, eliminating the previous heterogeneity enables us to exploit $\Delta \log y_{ig4}^c$ as a primary instrument for a transformed income term, $(\log y_{ih5}^c - \Delta \log y_{ig4}^c)$, because they are correlated with each other by construction, and we no longer have the unobserved heterogeneity of previous household in equation (2.36), which must be correlated with the instrument $(\Delta \log y_{ig4}^c)$. Moreover, since $\Delta \log y_{ig4}^c$ is, together with another instrument of household size in wave 4, determined by the characteristics of the original household, those instruments should be exogenous to composite error terms, including the unobserved heterogeneity of the current household.⁵⁰

Results

Appendix Table 2.23 presents the results of all specifications discussed above. Since income and consumption are exploited on a per capita basis, these estimations are essentially equivalent to household-level regressions. Since all individuals in the same household are recorded with exactly the same numerical values, standard errors are adjusted so that they could be robust to clustered residuals within each (current) household group. Since the validity of instruments is always a concern, the test of overidentifying restriction is implemented in all 2SLS estimations, and the p-values of the Hansen's J-test are reported, as well as those of the F-test to check the joint significance of instruments in the first-stage regressions.

Let us begin with the simple OLS estimation results of equation (2.35) in columns (a), (e), and

⁴⁹This specification restricts migrants into those who moved out of their original communities after wave 4 because their welfare measures in wave 4 are not observed if they migrated before wave 4.

⁵⁰As a result, we can have $E(\Delta \log y_{ig4}^c(\phi_1 u_{ih5} + \eta_h + \epsilon_{ih5} - \Delta \epsilon_{ig4})) = 0$ under our maintained assumptions.

(i). These again report the much smaller income elasticity of consumption for migrants, although the problem is slightly mitigated here by restricting our sample into panel individuals in contrast to Appendix Table 2.22, where the unit of observations is a household. As already discussed, we expect that the elasticity be biased for two reasons: an unobserved heterogeneity across households and a measurement error in earnings. While we are unable to draw inferences about the direction of bias that such unobservables generate, it might be worth saying a few words. For example, if a household is of high quality in the sense that it can earn higher income and have better access to consumption, the estimated elasticity will be biased upwards. On the other hand, the measurement error generates a standard attenuation bias. In particular, the serious measurement error in migrants' earnings seems to make their elasticity considerably biased towards zero.

In column (c), the first-difference of equation (2.35) was estimated for non-migrants by OLS. While the first-differencing usually worsens an attenuation bias associated with measurement errors in explanatory variables than a level equation, it might be able to correct the bias under our maintained assumption that the measurement error in earnings exists only in wave 5.⁵¹ In accordance with the idea, the estimated income elasticity in column (c) increased from that in (a), although there is no remarkable difference between them. While the exogeneity of current per worker income is theoretically doubtful, it seems to work well in column (b). Instruments are strong and the exogeneity is easily accepted by the test of overidentifying restriction. The estimated elasticity is again almost identical with that in column (a). Finally, the first-differenced per capita income was instrumented with per capita income and household size in wave 3 in column (d), where the rise in the elasticity is outstanding. However, weak instruments might have generated it. Overall, it appears that the measurement errors are not so serious in non-migrants' earnings.

The estimation results for inside and outside migrants reveal a similar picture. Both lagged

⁵¹To see this, we can consider the following simple model: $\log c_{iht} = \phi \log y_{iht}^{c*} + \eta_h + \epsilon_{iht}$, where $\log y_{iht}^c = \log y_{iht}^{c*} - u_{iht}$ in wave 5 but $\log y_{iht}^c = \log y_{iht}^{c*}$ in the earlier four waves. We assume that $E(\epsilon_{iht}) = E(u_{iht}) = E(\log y_{iht}^{c*}) = E(\log y_{iht}^{c*} \epsilon_{iht}) = E(\log y_{iht}^{c*} u_{iht}) = E(u_{iht} \epsilon_{iht}) = 0 \forall t$ and s . Then, a first-differenced OLS with data in waves 4 and 5 yields

$$\text{plim}_{n \rightarrow \infty} \hat{\phi} = \phi - \phi \frac{\sigma_u^2}{2(1 - \rho_{y*}^{45})\sigma_{y*}^2 + \sigma_u^2},$$

where $\sigma_u^2 \equiv \text{Var}(u_{iht})$, $\sigma_{y*}^2 \equiv \text{Var}(\log y_{iht}^{c*})$, and $\rho_{y*}^{45} \equiv \text{corr}(\log y_{iht5}^{c*}, \log y_{iht4}^{c*})$. It can be shown from this equation that an attenuation bias in the differenced equation will be smaller than that in a level equation, as long as $\rho_{y*}^{45} < \frac{1}{2}$. A 10-year gap appears to make the correlation of per capita income between waves 4 and 5 weak. In fact, we regressed the log of per capita income in wave 5 on that in wave 4, as well as dummies for location in wave 5 for all groups. The estimated correlation was 0.158, 0.080, and 0.024 for non-migrants, inside migrants, and outside migrants, respectively. Of these, only the case of non-migrants was significantly different from zero at a 1% level, whereas the other cases were not significant at any conventional level.

per capita income and previous location are such weak instruments that columns (g) and (k) yield the insignificant and spurious income elasticity. The weakness of instruments indicates that migrants are breaking the link with their original household. On the other hand, 2SLS estimations in columns (f), (h), (j), and (l) give rise to larger elasticity than OLS does. All instruments are strong and Hansen's J-test easily accepts their exclusion restriction. It seems that they are successfully correcting an attenuation bias to some extent.⁵²

To sum up the preceding discussions, it seems that we are underestimating the household income in wave 5 for two main types of (migrants' and, in particular, outside migrants') households: (a) those that did not happen to engage in wage employment and non-farm self-employment during the reference periods of survey questionnaire, and (b) those that engaged in non-farm self-employed activities during the periods but did not report any income from them. After recognizing these problems, we decided to exploit per worker weekly income as a primary income measure in the analysis. Our preference for this results from it being more accurate than per capita income for two reasons: firstly, per worker income is conceptually unaffected by the latent income problem and accurate as long as reported income is correct. This is because the way of constructing the per worker measure ensures that collected household income is allocated only to those who actually earned it in accordance with their contribution. On the other hand, per capita income is very vulnerable to the problem. Since we are not correctly measuring earnings at the household level, neither are we at the per capita level. Secondly, we expect that the per capita income measure is generally very sensitive to the demographic structure of household. The small proportion of workers in outside migrants' households in wave 5 (row (14) in Appendix Table 2.18) implies lower per capita income. However, one can note from row (15) in Appendix Table 2.18 that the size of migrants' households is smaller than that of non-migrants in wave 5 and that it is likely to increase per capita income. As a result, it seems that how per capita income is influenced by the demographic structure of household is more complicated than we think and that it more easily

⁵²However, the income elasticity is still smaller than that of non-migrants. The difference among groups might be caused by the fact that our measures of household income and consumption contain the (same) imputed value of food consumption of home produced crops. Given the way of measuring income and consumption, it is likely that moving away from farming self-employment breaks the correlation between them. Since outside migrants hardly engaged in farming self-employment in wave 5, the weak correlation might reflect the mechanism. In fact, the mean proportion of own consumption of home produced crops relative to the total household consumption in wave 5 was about 6% for the households of outside migrants in contrast to about 36% and 27% for that of non-migrants and inside migrants, respectively, although it was about 43%, 46%, and 41% in waves 1 to 4 for the household of non-migrants, inside migrants, and outside migrants, respectively. Own consumption of home produced crops might also be a part of story in explaining migrants' small income elasticity of consumption.

detracts from the comparability between individuals and waves when there is large heterogeneity in the household structure and the pattern of individual labour allocation between activities. In fact, we found a large number of unreasonable observations in wave 5, many migrants (in particular, outside migrants) with lower per capita weekly income but higher per capita annual consumption. All these push us to exploit per worker income more reliably.

For per worker income to be used in the empirical analysis, it still calls attention to several aspects. Firstly, per worker income does not contain many outside migrants in wave 5 (approximately 3.8% of total panel workers earning income), and even when it does, outside migrants' per worker income is mainly calculated from earnings information on non-farm wage activities (approximately 67% of total outside migrant workers earning income), not on non-farm self-employed activities. In addition, the accuracy of measured per worker income is not guaranteed either, and it seems that it is still measured with some noise, for which interested readers should read Appendix 2.10.3.

2.10.3 Per Worker Earnings Functions of Each Activity

While per worker income is conceptually insensitive to the problem of latent income, it does not necessarily mean that it is entirely free from measurement errors. In order to test the accuracy of reported income, we estimated per worker earnings functions by activity (wage employment, farming self-employment, and non-farm self-employment), group, and wave (pooled waves 1 to 4 and wave 5). Suppose that we are not correctly measuring the income of a particular group from a particular activity in a particular wave, therefore, earnings function should be different between groups and over time. The OLS estimation results are reported in Appendix Tables 2.24, 2.25, and 2.26.⁵³ It seems that migrants' earnings functions in both wage employment and non-farm self-employment are poorly specified relative to those of non-migrants' earnings in wave 5. This is a remaining concern in our per worker income measure.

⁵³We used a quadratic specification for the level of schooling in the outside migrants' wage earnings function in wave 5 because wage jobs in urban labour markets will be more heterogeneous than those in rural areas in terms of a variety of sectors and both technologies exploited and skills required in each sector. In addition, inside and outside migrants are pooled into one group in farming and non-farm self-employment income functions because few outside migrants reported income from those activities in wave 5.

2.11 Appendix: More on Descriptive Statistics

2.11.1 Occupational Mobility

While Kagera is a rural area, it has seemingly been difficult for people to survive only on self-employed farming. Figure 2.2 plots each income generating activity's share of weekly hours worked in total by migration status. In all groups of panel individuals, farming self-employment gradually lost the leading position, whereas non-farm wage employment and non-farm self-employed businesses started to become more important.⁵⁴ In particular, outside migrants spend more than 90% of total hours worked on non-farm wage employment and non-farm self-employment in wave 5 and, as a result, they replace farming self-employment as a primary income source.

However, Figure 2.2 does not allow us to distinguish between two potential mechanisms driving the picture. For example, all panel workers in all age groups may uniformly be moving away from farming self-employment to non-farm activities. Alternatively, only new participants in labour markets (i.e., children in wave 1) might exclusively be entering into non-farm activities once they become working-age adults, whereas the previous elderly did not change the type of jobs that they engaged in. To disentangle these explanations, Appendix Table 2.27 documents the transition of the activity shares of panel workers interviewed in both waves 1 and 5 by migration status and by age in wave 1.

The activity shares of both non-migrants and inside migrants report a similar pattern of transition. Panel workers aged 30 or above in wave 1 show no drastic change in the pattern of labour allocation between activities over time, but those who were aged 7 to 30 years old in wave 1 substantially moved away from self-employed farming to the other activities. The activity selected as an alternative is slightly different between the age categories of 7 to 20 years and 21 to 30 years. For the latter age group, non-farm self-employment seems to be a primary option, whereas non-farm wage employment is an equally available option for the former age group. It can be noted from Table 2.7 that panel individuals in the former age group (in particular, 16 to 20 years in wave 1) are more educated than the latter group in wave 5. Thus, all these suggest that while panel workers aged 7 to 30 years in wave 1 have substantially moved away from farming self-employment

⁵⁴One caution about the trend in the early 1990s is that part of the occupational mobility may be generated by an individual's response to smooth income against the economic recession (Kochar, 1999). In fact, we found that the proportion of non-farm businesses aged 3 months below relative to the total number of them was larger in waves 2 to 4 than in the other two waves.

to non-farm activities, it is likely that non-farm wage jobs are mainly accessible to the currently more educated previous young (7 to 20 years in wave 1). Turning to the previous youngest (below 7 years in wave 1), their first jobs are primarily farming self-employment in both groups of non-migrants and inside migrants, although it seems that the latter has slightly wider options, such as non-farm wage jobs and non-farm self-employment.

The transition from farming self-employment to non-farm activities is much more outstanding for outside migrants. Although the number of observations is very small in some age categories, it is likely that the outside option of the relatively old in wave 1 (21 to 40 years) is limited to non-farm self-employed businesses in wave 5, whereas non-farm wage jobs are equally accessible to the relatively young in wave 1 (7 to 20 years). Since those who were aged 7 to 20 years in wave 1 are more educated than the other age groups in wave 5 (see Table 2.7), this might again reflect that only the educated are able to engage in non-farm wage jobs.

Finally, we checked the correlation between the intertemporal change in activity shares and the attainment of education in Appendix Table 2.28 to see more clearly how the occupational mobility is related to the educational attainment. In pooling all groups aged 7 years old or above in wave 1, the educational attainment is strongly correlated with the occupational shift from farming self-employment to non-farm wage jobs. It seems that the educational attainment of the young aged 7 to 20 years in wave 1 largely drives the correlation.

2.11.2 Hours Worked, Hourly Income, and Occupational Disaggregation

Per worker weekly hours worked and hourly income (time average) are reported in Appendix Table 2.17. The movement of hours worked is consonant with that of per worker weekly income. There are dips between waves 1 and 5, but it appears that workers work longer hours in wave 5 than they did in wave 1. In particular, the hours worked of outside migrants are substantially longer than those of the other groups in wave 5. The movement of hourly income seems complicated. One point that could be said is that the hourly income of inside migrants declined from wave 1 to 5 in both mean and median. We have already seen that their per worker weekly income increased from wave 1 to 5, so it appears that the income rise results from the increase in working opportunities (hours worked) rather than the increase in the income (or wage) rate.

Per worker weekly income, weekly hours worked, and hourly income are further occupationally

disaggregated in Appendix Table 2.19. Although this sort of disaggregation sometimes makes descriptive statistics uninformative by significantly reducing the number of observations, the table still contains noticeable information. Firstly, 189 outside migrants are identified as workers in wave 5 (those who reported positive hours worked) in Appendix Table 2.17. About 43% of them are engaged in non-farm wage employment and another 43% are in non-farm self-employment, which again shows the importance of those activities for outside migrants. In addition, data presented that migrants' hourly income (whether inside or outside) in non-farm wage employment and non-farm self-employment rose from wave 1 to 5 and that migrants' hourly income in those activities was larger than that of farming self-employment in wave 5 with a few exceptions. This might partly explain the mechanism driving the occupational mobility in Figure 2.2, which shows that migrants' main activities have been shifting from self-employed farming to non-farm businesses over time. Taking it together with information on hours worked, migrants are on average receiving larger earnings from non-farm activities than from farm activities in wave 5. Even for non-migrants, non-farm activities seem to be more profitable than self-employed farming in wave 5 in the hourly income sense, although we cannot see a clear rise in hourly income in non-migrants' non-farm self-employed activities from wave 1 to 5 as opposed to migrants.

2.12 Appendix: Descriptions of Welfare Variables

Per Capita Annual Consumption

Per capita annual consumption is the average of the total annual consumption of a household. The processed household-level consumption is publicly available and composed of food consumption (seasonal and non-seasonal) and non-food consumption (education and health expenditures, miscellaneous non-food expenditures, etc.). See *Kagera Health and Development Survey – Consumption Expenditure Data* for details at <http://www.edi-africa.com/research/khds/introduction.htm>.

Per Worker Weekly Income

Per worker weekly income is the sum of the following individual income: wage earnings, non-farm self-employment income, and farming self-employment income. The KHDS collected individual

wage earnings in the past 7 days. This includes salary after deduction of all taxes, per diem allowances, bonuses, incentives, gratuities, overtime income, payment in kind, and so forth. Non-farm self-employment income is the sum of net profits from all non-farm household businesses. This is calculated by adding the value of products consumed or used by households for their use to net cash revenues from those businesses. Since income reference periods for non-farm self-employed activities were the past 2 weeks, we simply took the time average to compute weekly income. To construct an individual income measure, the household income was finally allocated to each household member in proportion to the individual hours worked on those activities in the past 7 days, which were individually reported in a different section from sections collecting household-level self-employment income. Farming self-employment income is the household-level annual (or 6-month in waves 2 to 4) net profits from self-employed farming and livestock activities. This is calculated by adding the value of crops and crop- or livestock-related home products consumed or used by households or lost for some reasons (e.g., natural disasters, thefts) to net cash revenues from those activities. As in the case of non-farm self-employment income, the weekly income is the time average of annual (or 6-month) income and the household-level weekly income was assigned to each household member in proportion to the individual hours worked on those activities in the past 7 days.

Per Capita Weekly Income

Per capita weekly income is the average of the total weekly income of household. The household income consists of the income from the activities described before: wage employment, non-farm self-employment, and farming self-employment.

Price Index

All monetary values in this chapter are transformed into baseline prices (1991 prices) by using the Laspeyres and Fisher indices, both of which are in the public domain and constructed by household-wave group. See *Kagera Health and Development Survey – Price index* for details at <http://www.edi-africa.com/research/khds/introduction.htm>.

Table 2.1: Migrants by Origin

Original location	Non-migrants	Migrants (inside)	Migrants (outside)	No. of panel individuals in wave 1
Biharamulo	0.65	0.19	0.15	320
Bukoba Rural	0.61	0.29	0.09	1229
Bukoba Urban	0.58	0.30	0.11	658
Muleba	0.70	0.26	0.03	544
Ngara	0.74	0.21	0.03	401
Karagwe	0.74	0.22	0.03	506
Total	0.65	0.26	0.07	3658

Note: The number is the proportion relative to the total number of panel individuals in wave 1 in each location.

Table 2.2: Migrants by Age at Migration

Age at migration	Inside of Kagera		Outside of Kagera	
	Male	Female	Male	Female
Below 7 years old	0.11	0.06	0.04	0.06
7 to 20 years old	0.43	0.57	0.52	0.59
21 to 30 years old	0.31	0.25	0.30	0.25
31 to 40 years old	0.05	0.04	0.05	0.03
41 to 50 years old	0.03	0.01	0.02	-
51 to 60 years old	0.02	0.01	0.02	0.01
61 years old or above	0.01	0.02	0.01	0.03
No. of migrants	461	775	170	185

Notes: (1) Migrants are panel individuals in wave 5 who migrated at some point between 1991 and 2004. (2) The number is the proportion relative to the total number of migrants in each destination-gender category.

Table 2.3: Migrants by Year of Migration

Year of migration	Inside of Kagera		Outside of Kagera	
	Male	Female	Male	Female
1991	0.00	0.00	0.01	0.02
1992	0.04	0.04	0.02	0.07
1993	0.05	0.08	0.07	0.06
1994	0.13	0.13	0.10	0.11
1995	0.08	0.13	0.09	0.11
1996	0.06	0.08	0.05	0.05
1997	0.06	0.06	0.04	0.10
1998	0.07	0.06	0.10	0.03
1999	0.07	0.07	0.07	0.08
2000	0.09	0.07	0.11	0.05
2001	0.05	0.04	0.04	0.04
2002	0.08	0.08	0.11	0.12
2003	0.11	0.09	0.08	0.07
2004	0.03	0.01	0.04	0.01
	461	775	170	185

Notes: (1) Migrants are panel individuals in wave 5 who migrated at some point between 1991 and 2004. (2) The number is the proportion relative to the total number of migrants in each destination-gender category.

Table 2.4: Reason for Migration

	Pooling			Inside of Kagera			Outside of Kagera		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Economic reason									
No job / wanted better job	0.07	0.02	0.04	0.06	0.02	0.03	0.08	0.03	0.05
Business Opportunities	0.21	0.04	0.11	0.17	0.03	0.08	0.34	0.10	0.21
Land not available	0.11	0.03	0.06	0.13	0.03	0.07	0.04	0.03	0.03
Schooling	0.09	0.06	0.07	0.06	0.04	0.05	0.17	0.14	0.15
Family related reason									
Marriage	0.03	0.47	0.29	0.03	0.52	0.34	0.01	0.26	0.14
Divorce	0.00	0.03	0.02	0.00	0.03	0.02	-	0.02	0.01
Widowhood	0.05	0.02	0.03	0.06	0.02	0.03	0.04	0.01	0.02
Death of parents	0.00	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00
Illness of household members	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.01
Other family problems	0.05	0.02	0.03	0.07	0.02	0.04	-	0.01	0.00
Political reason									
Posted to new area	0.02	0.00	0.01	0.01	0.00	0.00	0.04	0.02	0.03
Political / economic problems	0.07	0.06	0.07	0.09	0.07	0.07	0.02	0.05	0.04
Natural disasters	0.24	0.17	0.20	0.25	0.15	0.19	0.21	0.28	0.25
No. of migrants	626	955	1581	459	770	1229	167	185	352

Notes: (1) Migrants are panel individuals in wave 5 who migrated at some point between 1991 and 2004. (2) The number is the proportion relative to the total number of migrants in each destination-gender category.

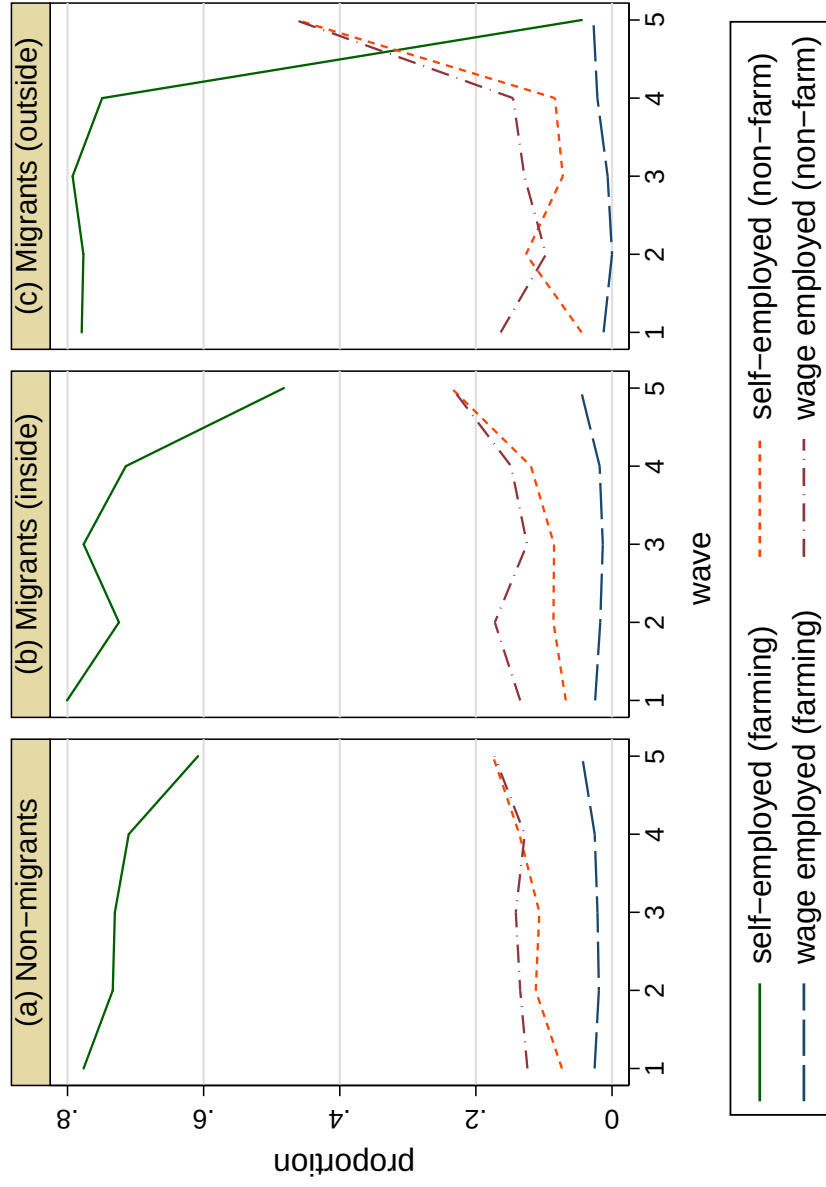


Figure 2.2: Transition of Activity Shares in Total Hours Worked

Note: Activity shares are the proportion of weekly hours worked of each activity relative to the total weekly hours worked.

Table 2.5: Summary of Key Variables - Individual Level

Non-migrants				Migrants (inside)				Migrants (outside)			
Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.
(1) Per capita annual consumption excluding remittances (US dollars, 1991 price)											
Wave 1	208.86	(136.23)	171.05	2426	217.77*	(135.17)	183.53	224.51*	(138.45)	189.48	291
Wave 2	185.95	(141.11)	159.67	2422	178.48	(111.74)	154.27	186.12	(126.14)	149.93	276
Wave 3	169.90	(105.47)	146.04	2353	164.91	(93.32)	146.72	177.50	(104.63)	161.70	247
Wave 4	166.75	(97.06)	147.03	2223	165.65	(90.22)	147.88	168.33	(105.11)	154.02	222
Wave 5	208.67	(154.09)	170.39	2825	273.81***	(233.29)	201.39	503.40***	(464.01)	374.30	352
(2) Per capita weekly income (US dollars, 1991 price)											
Wave 1	4.32	(8.46)	2.34	2401	4.35	(7.74)	2.62	4.51	(9.66)	2.53	280
Wave 2	2.69	(3.12)	1.90	2401	2.60	(2.94)	1.87	4.38	(17.10)	1.72	264
Wave 3	2.56	(2.75)	2.00	2318	2.39	(2.54)	1.95	2.54	(3.60)	1.91	245
Wave 4	2.67	(4.35)	1.95	2197	2.59	(4.05)	1.86	2.45	(3.48)	1.75	218
Wave 5	3.95	(7.81)	2.40	2780	3.99	(9.04)	1.84	4.67	(9.19)	1.63	218
(3) Per worker weekly income (US dollars, 1991 price)											
Wave 1	8.33	(20.22)	4.10	1338	7.75	(26.56)	3.19	7.50	(22.93)	2.57	156
Wave 2	4.71	(6.77)	2.91	1407	4.10	(10.98)	1.97	545	(4.89)	1.60	152
Wave 3	4.29	(6.36)	2.63	1448	3.45***	(6.52)	2.00	2.86***	(5.32)	1.55	133
Wave 4	4.53	(9.13)	2.83	1344	3.50***	(6.61)	1.96	2.83***	(5.10)	1.50	117
Wave 5	7.28	(16.38)	3.67	2114	8.59**	(19.76)	4.60	14.43***	(22.23)	10.63	120
(4) Per worker weekly hours worked											
Wave 1	24.75	(16.80)	23.00	1352	19.39***	(15.24)	15.00	18.71***	(15.37)	12.70	159
Wave 2	21.51	(16.79)	18.00	1429	17.40***	(16.42)	12.00	555	(13.91)	11.00	159
Wave 3	20.06	(15.63)	17.00	1470	16.48***	(14.57)	12.00	559	(14.28)	9.00	135
Wave 4	20.23	(16.63)	17.50	1357	15.57***	(15.87)	10.70	487	(13.65)	10.10	118
Wave 5	28.42	(19.52)	25.00	2157	29.99***	(20.55)	25.00	947	(26.08)	54.00	189
(5) Gender (one if male)											
Wave 1-4	0.52	(0.49)	-	9425	0.37***	(0.48)	-	3564	0.48**	-	1038
Wave 5	0.52	(0.49)	-	2832	0.38***	(0.48)	-	1289	0.47	-	355
(6) Education (years)											
Own											
Wave 1-4	2.56	(3.08)	0.00	9426	2.73***	(3.15)	1.00	3564	3.12***	(3.34)	1038
Wave 5	4.69	(3.22)	6.00	2811	5.49***	(3.35)	7.00	1280	6.70***	(3.49)	350
Average of household											
Wave 1-4	2.67	(1.46)	2.55	9426	2.74**	(1.49)	2.61	3564	2.99***	(1.45)	1038
Wave 5	3.48	(1.94)	3.44	2808	3.90***	(2.43)	3.50	1280	5.45***	(2.95)	347
Household head											
Wave 1-4	4.39	(3.22)	4.00	9426	4.32	(3.39)	4.00	3564	4.61*	(3.48)	1038
Wave 5	4.88	(3.46)	6.00	2811	6.12***	(3.80)	7.00	1280	8.06***	(4.32)	350
(7) Age (years)											
Own											
Wave 1-4	22.29	(19.28)	16.00	9425	15.33***	(12.13)	13.00	3564	14.85***	(11.81)	1038
Wave 5	33.00	(19.36)	27.00	2832	26.99***	(12.33)	25.00	1289	26.27***	(11.30)	355
Average of household											
Wave 1-4	21.66	(8.23)	19.90	9425	21.55	(7.12)	20.14	3564	21.82	(7.02)	1038
Wave 5	25.18	(11.11)	22.63	2832	21.79***	(9.23)	19.20	1289	21.43***	(7.15)	355
Household head											
Wave 1-4	48.16	(15.31)	48.00	9426	49.92***	(16.03)	50.00	3564	52.21***	(16.50)	1038
Wave 5	49.72	(17.52)	49.00	2811	39.48***	(15.22)	35.00	1280	35.62***	(13.56)	350

Notes: (1) Unit of observations is a person-wave. (2) In per capita measures, all members in the same household-wave group are recorded with the same numerical values. (3) The equality of means between non-migrants and migrants (inside) and that between non-migrants and migrants (outside) are examined by T-tests, assuming unequal variance. The degree of freedom is approximated by Satterthwaite. *** denotes the rejection of equality at 1%, ** at 5%, and * at 10%.

Table 2.6: Summary of Key Variables - Household Level

Non-migrants				Migrants (inside)				Migrants (outside)			
Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.
(1) Household annual consumption (US dollars, 1991 price)											
Wave 1-4	955.44	(811.86)	745.39	1282	1246.71***	1005.80	1819	1445.38***	(938.73)	1236.74	620
Wave 5	983.77	(760.58)	808.49	1354	1092.60***	886.73	1094	1669.54***	(1141.01)	1457.44	297
(2) Household weekly income (US dollars, 1991 price)											
Total											
Wave 1-4	15.33	(27.32)	9.48	1258	21.57***	13.21	1803	25.24***	(72.69)	14.68	607
Wave 5	17.70	(35.98)	11.52	1312	15.06*	7.98	986	11.63***	(19.30)	7.32	183
Farm wage employed											
Wave 1-4	2.64	(3.15)	1.71	106	3.28	1.14	147	2.16	(2.08)	2.13	41
Wave 5	4.70	(4.03)	3.47	160	4.80	3.39	98	7.79	(5.55)	6.15	7
Non-farm wage employed											
Wave 1-4	9.06	(11.08)	5.90	303	9.15	5.63	423	10.97*	(11.89)	6.94	174
Wave 5	14.34	(19.42)	9.75	289	18.64	10.92	215	17.82	(25.42)	13.36	85
Farm self-employed											
Wave 1-4	10.71	(20.28)	6.99	1230	16.40***	10.53	1798	18.61***	(70.77)	10.60	609
Wave 5	9.52	(31.33)	6.58	1279	7.36**	5.33	889	3.46***	(3.41)	2.50	106
Non-farm self-employed											
Wave 1-4	12.40	(31.62)	4.33	264	11.87	3.78	438	12.25	(27.02)	4.23	182
Wave 5	14.36	(21.37)	7.66	448	20.86**	10.51	189	17.49	(12.50)	15.38	11
(3) Household weekly hours worked											
Total											
Wave 1-4	60.09	(44.46)	51.00	1223	80.57***	69.65	1770	86.50***	(59.41)	76.00	607
Wave 5	70.55	(49.63)	61.00	1292	53.97***	42.00	903	53.76***	(32.48)	54.00	186
Farm wage employed											
Wave 1-4	18.06	(15.72)	14.00	109	17.63	12.00	152	15.72	(11.47)	12.00	43
Wave 5	22.24	(19.29)	16.00	161	20.50	13.00	99	39.00***	(10.96)	39.00	8
Non-farm wage employed											
Wave 1-4	43.14	(32.03)	40.00	306	43.25	40.00	440	53.84***	(37.83)	45.00	179
Wave 5	45.67	(31.55)	42.00	307	48.12	45.00	214	54.83**	(27.64)	55.00	83
Farm self-employed											
Wave 1-4	44.12	(36.69)	35.00	1124	61.81***	52.60	1727	60.82***	(47.32)	50.00	591
Wave 5	49.80	(41.51)	41.00	1171	38.25***	27.00	667	16.27***	(12.07)	16.00	33
Non-farm self-employed											
Wave 1-4	28.97	(29.19)	18.00	301	29.79	18.00	475	31.39	(31.73)	22.00	199
Wave 5	36.59	(33.93)	26.00	416	39.46	30.50	276	55.43***	(33.14)	57.00	83
(4) Average education of households (years)											
Wave 1-4	2.87	(2.12)	2.66	1290	2.73**	2.60	1820	3.06**	(1.53)	2.83	621
Wave 5	3.39	(2.14)	3.33	1355	3.89***	3.50	1096	5.64***	(2.90)	5.45	297
(5) Education of household heads (years)											
Wave 1-4	4.10	(3.47)	4.00	1290	4.20	4.00	1820	4.60***	(3.49)	4.00	621
Wave 5	4.96	(3.48)	7.00	1357	6.12***	7.00	1097	8.08***	(4.31)	7.00	300
(6) Average age of households (years)											
Wave 1-4	30.36	(19.67)	25.89	1290	22.63***	20.87	1820	22.57***	(7.89)	21.00	621
Wave 5	24.92	(12.92)	21.00	1357	21.73***	19.00	1097	21.57***	(7.30)	20.07	300
(7) Age of household heads (years)											
Wave 1-4	47.93	(17.64)	47.00	1290	49.79***	50.00	1820	52.43***	(16.46)	53.00	621
Wave 5	45.18	(18.41)	40.00	1350	38.70***	34.00	1090	35.40***	(12.57)	32.00	293
(8) Proportion of workers in households											
Wave 1-4	0.78	(0.38)	0.70	981	0.69***	0.66	1820	0.68***	(0.29)	0.66	621
Wave 5	0.08	(0.38)	0.66	1355	0.46***	0.33	1097	0.30***	(0.37)	0.20	300
(9) Household size											
Wave 1-4	4.13	(2.30)	4.00	1283	6.78***	6.00	1820	7.63***	(3.69)	7.00	621
Wave 5	4.79	(2.68)	4.00	1357	4.59*	4.00	1097	4.21***	(2.61)	4.00	300

Notes: (1) Unit of observations is a household-wave. (2) Households are defined as migrants in wave 1 to 4 if they supply at least one member who moved out of original community between 1991 and 2004 and in wave 5 if they contain at least one panel migrant in them. (3) The equality of means between non-migrants and migrants (inside) and that between non-migrants and migrants (outside) are examined by T-tests, assuming unequal variance. The degree of freedom is approximated by Satterthwaite. *** denotes the rejection of equality at 1%, ** at 5%, and * at 10%.

Table 2.7: Migration and Educational Attainment by Age in Wave 1

Age in wave 1		Wave 1		Wave 5		Change		No. of panel individuals in wave 1
		Mean	Std.	Mean	Std.	Mean	Std.	
Below 7 years old	Non-migrants	-	-	4.40	(2.60)	4.40	(2.60)	609
	Migrants (inside)	-	-	4.27	(2.87)	4.27	(2.87)	198
	Migrants (outside)	-	-	4.91	(3.17)	4.91	(3.17)	58
7 to 15 years old	Non-migrants	1.50	(1.99)	6.05	(3.09)	4.49	(3.20)	587
	Migrants (inside)	1.68	(2.07)	6.15	(3.08)	4.47	(3.03)	427
	Migrants (outside)	2.00	(2.05)	7.73	(3.02)	5.73	(3.10)	134
16 to 20 years old	Non-migrants	5.76	(2.43)	6.52	(2.34)	0.77	(1.70)	253
	Migrants (inside)	6.20	(2.23)	7.17	(2.37)	0.97	(1.68)	174
	Migrants (outside)	6.84	(2.23)	8.26	(2.56)	1.42	(1.99)	50
21 to 30 years old	Non-migrants	5.84	(2.77)	6.17	(2.65)	0.33	(1.25)	277
	Migrants (inside)	6.02	(2.71)	6.32	(2.74)	0.30	(1.16)	82
	Migrants (outside)	7.24	(3.41)	7.36	(3.47)	0.12	(0.43)	25
31 years old or above	Non-migrants	3.11	(3.05)	3.53	(3.26)	0.42	(1.23)	680
	Migrants (inside)	3.88	(4.17)	4.49	(4.33)	0.61	(1.48)	85
	Migrants (outside)	4.84	(4.84)	5.31	(4.89)	0.47	(1.65)	19

Note: Individuals are panel respondents who were interviewed in both wave 1 and 5.

Table 2.8: Reason for Migration by Educational Attainment

	Migrants (inside)						Migrants (outside)					
	> mean attainment			≤ mean attainment			> mean attainment			≤ mean attainment		
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Economic reason												
No job / wanted better job	0.10	0.01	0.05	0.06	0.03	0.04	0.05	0.04	0.05	0.10	0.03	0.06
Business Opportunities	0.18	0.02	0.08	0.19	0.03	0.09	0.20	0.09	0.20	0.39	0.16	0.26
Land not available	0.09	0.02	0.05	0.20	0.06	0.11	0.01	-	0.00	0.04	0.09	0.11
Schooling	0.09	0.07	0.08	-	0.00	0.00	0.25	0.24	0.24	0.08	0.01	0.04
Family related reason												
Marriage	0.04	0.57	0.36	0.03	0.53	0.37	0.01	0.25	0.12	0.02	0.29	0.17
Divorce	-	0.02	0.01	0.00	0.04	0.03	-	0.01	0.00	-	0.03	-
Widowhood	0.08	0.03	0.05	0.04	0.01	0.02	0.06	0.01	0.04	-	-	-
Death of parents	0.00	0.00	0.00	0.00	0.00	0.00	-	-	-	-	0.01	0.00
Illness of household members	0.00	0.00	0.00	0.00	0.00	0.00	-	0.01	0.00	0.02	0.01	0.01
Other family problems	0.05	0.02	0.03	0.11	0.03	0.05	-	-	-	-	0.01	0.00
Political reason												
Posted to new area	0.02	0.00	0.01	0.00	0.00	0.00	0.05	0.03	0.04		0.01	0.02
Political / economic problems	0.08	0.05	0.06	0.09	0.07	0.08	0.01	0.04	0.02	0.08	0.06	0.07
Natural disasters	0.21	0.12	0.15	0.20	0.11	0.14	0.21	0.24	0.22	0.10	0.24	0.18
No. of migrants	186	283	469	148	305	453	87	83	170	48	65	113

Notes: (1) Migrants are panel respondents who were interviewed in both wave 1 and 5. (2) The number is the proportion relative to the total number of migrants in each category. (3) The mean attainment means that the mean educational growth of non-migrants from wave 1 to 5 (2.22 years).

Table 2.9: Informative Evidence on Exclusion Restrictions

Dependent variables: log of	per worker income (weekly) in wave 5			per capita consumption (annual) in wave 5		
	OLS (a)	OLS (b)	OLS (c)	OLS (d)	OLS (e)	OLS (f)
Migrants	0.328*** (0.072)	-	0.321*** (0.072)	-	0.264*** (0.033)	-
Inside migrants	-	0.306*** (0.073)	-	0.300*** (0.073)	-	0.192*** (0.033)
Outside migrants	-	0.468*** (0.170)	-	0.454*** (0.170)	-	0.523*** (0.067)
Individual characteristics in wave 5						
Age (years)	0.095*** (0.007)	0.095*** (0.007)	0.094*** (0.007)	0.094*** (0.007)	0.005* (0.003)	0.005* (0.003)
Age squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)
Education (years)	0.036*** (0.010)	0.035*** (0.010)	0.011 (0.021)	0.012 (0.021)	0.093*** (0.011)	0.086*** (0.011)
Education squared	-	-	0.002 (0.002)	0.002 (0.002)	-	-
Male dummy	0.304*** (0.078)	0.304*** (0.078)	0.305*** (0.078)	0.305*** (0.078)	0.044 (0.033)	0.045 (0.033)
Instruments for selection into migration						
Rainfall deviation in origin from cluster average (1986 to 2000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.052 (0.047)	0.000 (0.000)
× Aged 5 to 15 in wave 1	-0.265** (0.108)	-0.266** (0.108)	-0.259** (0.108)	-0.260** (0.109)	-0.011 (0.047)	-0.063 (0.047)
HH head's spouse dummy in wave 1	-0.227** (0.108)	-0.229** (0.108)	-0.221** (0.108)	-0.223** (0.108)	-0.021 (0.050)	-0.012 (0.048)
Female	0.127 (0.097)	0.137 (0.098)	0.128 (0.097)	0.138 (0.098)	-0.013 (0.050)	0.001 (0.050)
× Aged 5 to 15 in wave 1	-	-	-	-	-	-
Instruments for selection into schooling						
Father's education	0.003 (0.014)	0.002 (0.014)	0.002 (0.014)	0.002 (0.014)	0.008 (0.006)	0.008 (0.006)
× Aged 5 to 15 in wave 1	0.027 (0.016)	0.026 (0.016)	0.026 (0.016)	0.026 (0.016)	-0.005 (0.007)	-0.007 (0.007)
Mother's education	5.107*** (0.202)	5.124*** (0.203)	5.248*** (0.224)	5.295*** (0.230)	10.642*** (0.113)	10.600*** (0.112)
Constant	Yes	Yes	Yes	Yes	Yes	Yes
IHHFE	Yes	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes	Yes
Joint significance of instruments						
F-statistic	5.26	5.24	5.08	5.08	0.75	0.73
p-values	0.000	0.000	0.000	0.000	0.609	0.571
R-squared	0.533	0.534	0.534	0.534	0.552	0.558
No. of observations	2562	2562	2562	2562	3645	3645

Notes: (1) Unit of observations is a person. (2) Sample is panel workers in income regressions, whereas panel individuals in consumption regressions. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity. (5) Remittance expenditures are excluded from the measure of consumption. (5) Initial location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (6) IHHFE stands for initial household fixed effects.

Table 2.10: Reduced Form Equations: Migration and Schooling

Dependent variables:	one if migrate and zero otherwise			Schooling level in wave 5
	Pooling	Inside migrants	Outside migrants	
	OLS (a)	OLS (b)	OLS (c)	OLS (d)
Individual characteristics in wave 1				
Age (years)	0.005** (0.002)	0.006*** (0.002)	-0.000 (0.001)	-0.060*** (0.011)
Age squared	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	0.000 (0.000)
Education (years)	0.014*** (0.003)	0.006* (0.003)	0.007*** (0.002)	0.576*** (0.020)
Male dummy	-0.100*** (0.022)	-0.075*** (0.021)	-0.025* (0.013)	0.621*** (0.116)
Instruments for selection into migration				
Rainfall deviation in origin from cluster average (1986 to 2000)	0.000* (0.000)	0.000 (0.000)	0.000** (0.000)	0.000** (0.000)
× Aged 5 to 15 in wave 1				
HH's head dummy in wave 1	-0.184*** (0.031)	-0.174*** (0.030)	-0.010 (0.016)	0.282 (0.174)
HH head's spouse dummy in wave 1	-0.243*** (0.033)	-0.220*** (0.032)	-0.023 (0.018)	-0.024 (0.179)
Female × Aged 5 to 15 in wave 1	0.158*** (0.031)	0.177*** (0.029)	-0.019 (0.019)	0.439** (0.172)
Instruments for selection into schooling				
Father's education × Aged 5 to 15 in wave 1	0.014*** (0.004)	0.008* (0.004)	0.006** (0.002)	0.102*** (0.023)
Mother's education × Aged 5 to 15 in wave 1	-0.011** (0.004)	-0.016*** (0.004)	0.004 (0.003)	0.182*** (0.027)
Constant	1.010*** (0.050)	1.067*** (0.047)	-0.056* (0.030)	3.809*** (0.280)
IHHFE	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes
Joint significance of instruments				
F-statistic	27.25	22.86	4.42	58.63
p-values	0.000	0.000	0.000	0.000
R-squared	0.456	0.438	0.388	0.666
No. of observations	3689	3689	3689	3658

Notes: (1) Unit of observations is a person. (2) Sample is panel individuals. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity. (5) Initial location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (6) IHHFE stands for initial household fixed effects.

Table 2.11: IV Approach: Structural Estimation Results

Dependent variables:		log of per worker income (weekly) in wave 5				log of per capita consumption (annual) in wave 5			
		Eqn-by-eqn		Eqn-by-eqn		Eqn-by-eqn		Eqn-by-eqn	
		OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Migrants		0.361*** (0.071)	0.339*** (0.072)	0.232 (0.277)	0.235 (0.392)	0.339*** (0.032)	0.269*** (0.031)	0.315*** (0.093)	0.277*** (0.062)
Education in wave 5 (years) (Individual / HH average)		-	0.042*** (0.009)	0.094*** (0.015)	0.247*** (0.052)	-	0.094*** (0.011)	0.115*** (0.030)	0.077*** (0.002)
Age in wave 5			0.098*** (0.007)	0.089*** (0.007)	0.043*** (0.012)		0.004* (0.002)	0.004 (0.002)	0.002*** (0.000)
Age squared in wave 5			-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)		-0.000 (0.000)	-0.000 (0.000)	-0.000*** (0.000)
Male dummy			0.277*** (0.055)	0.238*** (0.057)	0.176** (0.071)		0.042* (0.023)	0.041 (0.026)	0.025*** (0.007)
HH head in wave 1			-0.295*** (0.108)	-0.287*** (0.106)	-0.248** (0.125)		-	-	-
HH head's spouse in wave 1			-0.357*** (0.103)	-0.304*** (0.101)	-0.269** (0.126)		-	-	-
Constant			5.235*** (0.215)	5.700*** (0.175)	4.441*** (0.437)		10.716*** (0.092)	9.540*** (0.283)	9.939*** (0.523)
IHHFE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint significance of instruments in the first stage									
Migrants									
F-statistic	-	-	-	8.92	7.38	-	-	20.07	-
p-values	-	-	-	0.000	0.000	-	-	0.000	-
Education									
F-statistic	-	-	-	98.95	10.08	-	-	11.12	-
p-values	-	-	-	0.000	0.000	-	-	0.000	-
Sargan test (p-values)	-	-	-	0.001	0.490	-	-	0.226	-
Joint significance of migration and schooling effects									
F-test (p-values)	-	-	-	0.000	0.000	-	-	0.000	0.000
R-squared	0.523	0.528	0.528	0.519	0.396	0.520	0.551	0.549	0.550
No. of obs.	2562	2562	2562	2562	2562	3645	3645	3645	3645

Notes: (1) Unit of observations is a person. (2) Sample is panel workers in income regressions, whereas panel individuals in consumption regressions. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity in OLS. (5) Remittance expenditures are excluded from the measure of consumption. (5) Initial location is classified into 6 categories: Bihamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (6) IHHFE stands for initial household fixed effects.

Table 2.12: Control Function (CF) Approach: Earnings

Dependent variable:	log of per worker income (weekly) in wave 5						
	Linear education				Quadratic education		
	OLS (a)	OLS (b)	CF (c)	CF (d)	OLS (e)	CF (f)	CF (g)
(A) Pooling							
Migrants	0.361*** (0.071)	0.337*** (0.074)	0.002 (0.312)	0.328*** (0.072)	0.327*** (0.074)	-0.046 (0.370)	0.317*** (0.077)
Individual education (years) in wave 5	-	0.041*** (0.010)	0.108*** (0.021)	0.092*** (0.016)	-0.017 (0.024)	0.032 (0.034)	0.029 (0.031)
Education squared in wave 5	-	-	-	-	0.006** (0.002)	0.008*** (0.003)	0.007** (0.002)
Migrants × (Demeaned) education	-	0.003 (0.017)	0.002 (0.018)	0.002 (0.017)	0.075* (0.044)	0.102** (0.045)	0.089* (0.049)
Migrants × (Demeaned) education squared	-	-	-	-	-0.007* (0.003)	-0.010** (0.004)	-0.009** (0.004)
Age (years) in wave 5	0.098*** (0.007)	0.089*** (0.007)	0.086*** (0.009)	0.078*** (0.007)	0.088*** (0.007)	0.085*** (0.009)	0.077*** (0.008)
Age squared in wave 5	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Male dummy	0.277*** (0.055)	0.238** (0.057)	0.132 (0.092)	0.198*** (0.063)	0.246*** (0.057)	0.131 (0.096)	0.203*** (0.069)
HH head in wave 1	-0.295*** (0.108)	-0.287*** (0.106)	-0.334** (0.139)	-0.276** (0.119)	-0.276** (0.107)	-0.332** (0.159)	-0.263** (0.116)
HH head's spouse in wave 1	-0.357*** (0.103)	-0.305*** (0.102)	-0.317** (0.147)	-0.237** (0.100)	-0.295*** (0.102)	-0.316* (0.172)	-0.221** (0.106)
Migration control	-	-	0.206 (0.200)	-	-	0.227 (0.232)	-
Schooling control	-	-	-0.080*** (0.025)	-0.072*** (0.020)	-	-0.089*** (0.031)	-0.078*** (0.024)
Constant	5.235*** (0.215)	5.687*** (0.196)	4.724*** (0.438)	4.793*** (0.559)	5.845*** (0.271)	4.854*** (0.656)	4.522*** (0.476)
IHHFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.523	0.528	0.492	0.532	0.530	0.496	0.534
No. of obs.	2562	2562	1834	2562	2562	1834	2562
	OLS (h)	OLS (i)		CF (j)	OLS (k)		CF (l)
(B) Inside and outside migrants							
Inside migrants	0.332*** (0.072)	0.314*** (0.073)		0.309*** (0.075)	0.304*** (0.073)		0.297*** (0.073)
Outside migrants	0.547*** (0.169)	0.505** (0.252)		0.463 (0.294)	0.489* (0.271)		0.435 (0.420)
Individual education (years) in wave 5	-	0.042*** (0.010)		0.091*** (0.017)	-0.017 (0.024)		0.028 (0.029)
Education squared in wave 5	-	-		-	0.006** (0.002)		0.007*** (0.002)
Inside migrants × (Demeaned) education	-	-0.000 (0.018)		-0.001 (0.022)	0.075 (0.046)		0.090* (0.051)
Inside migrants × (Demeaned) education squared	-	-		-	-0.008* (0.004)		-0.009* (0.004)
Outside migrants × (Demeaned) education	-	-0.002 (0.055)		0.002 (0.062)	0.057 (0.163)		0.083 (0.264)
Outside migrants × (Demeaned) education squared	-	-		-	-0.006 (0.009)		-0.008 (0.013)
Age (years) in wave 5	0.098*** (0.007)	0.089*** (0.007)		0.078*** (0.008)	0.088*** (0.007)		0.077*** (0.008)
Age squared in wave 5	-0.000*** (0.000)	-0.000*** (0.000)		-0.000*** (0.000)	-0.000*** (0.000)		-0.000*** (0.000)
Male dummy	0.270*** (0.056)	0.233*** (0.057)		0.193*** (0.065)	0.241*** (0.057)		0.199*** (0.062)
HH head in wave 1	-0.294*** (0.108)	-0.286*** (0.106)		-0.276** (0.106)	-0.276** (0.107)		-0.263*** (0.115)
HH head's spouse in wave 1	-0.362*** (0.103)	-0.309*** (0.102)		-0.240** (0.108)	-0.299*** (0.102)		-0.225** (0.104)
Schooling control	-	-		-0.071*** (0.020)	-		-0.077*** (0.019)
Constant	5.237*** (0.215)	5.740*** (0.203)		4.807*** (0.531)	5.934*** (0.303)		4.278*** (0.346)
IHHFE	Yes	Yes		Yes	Yes		Yes
Location dum. (wave 1)	Yes	Yes		Yes	Yes		Yes
R-squared	0.523	0.529		0.532	0.531		0.535
No. of obs.	2562	2562		2562	2562		2562

Notes: (1) Unit of observations is a person. (2) Sample is panel workers. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity in OLS and estimated by bootstrapping with 100 replications in the control function (CF) approach. (5) Initial location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (6) IHHFE stands for initial household fixed effects.

Table 2.13: Control Function (CF) Approach: Consumption

Dependent variable:	log of per capita consumption (annual)				
	Level (wave 5)				Growth (wave 1 to wave 5)
	OLS (a)	OLS (b)	CF (c)	CF (d)	OLS (e)
(A) Pooling					
Migrants	0.339*** (0.032)	0.198*** (0.032)	0.175* (0.096)	0.198*** (0.033)	0.251*** (0.030)
HH average education (years) (level (wave 5) / growth (wave 1 to 5))	-	0.051*** (0.019)	0.101* (0.059)	0.064 (0.048)	0.046** (0.018)
Demeaned HH average education: level (wave 5) or growth (wave 1 to 5)					
× Migrants	-	0.068*** (0.019)	0.058*** (0.021)	0.068*** (0.023)	0.081*** (0.018)
Age (years) in wave 5	0.005** (0.002)	0.004* (0.002)	0.005 (0.003)	0.004* (0.002)	-
Age squared in wave 5	-0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-
Age in wave 1	-	-	-	-	0.001** (0.000)
Male dummy	0.086*** (0.024)	0.031 (0.023)	0.029 (0.037)	0.026 (0.028)	-
Migration control	-	-	0.021 (0.063)	-	-
Schooling control	-	-	-0.038 (0.053)	-0.013 (0.042)	-
Constant	10.003*** (0.082)	11.024*** (0.143)	9.745*** (0.210)	9.876*** (0.184)	0.928*** (0.123)
IHHFE	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes
R-squared	0.520	0.557	0.512	0.557	0.635
No. of obs.	3645	3645	2610	3645	3645
	OLS (f)	OLS (g)		CF (h)	OLS (i)
(B) Inside and outside migrants					
Inside migrants	0.238*** (0.032)	0.153*** (0.032)		0.153*** (0.035)	0.196*** (0.030)
Outside migrants	0.680*** (0.069)	0.426*** (0.077)		0.426*** (0.083)	0.440*** (0.078)
HH average education (years) level (wave 5) / growth (wave 1 to 5)	-	0.052*** (0.019)		0.054 (0.048)	0.045** (0.018)
Demeaned HH average education: level (wave 5) or growth (wave 1 to 5)					
× Inside migrants	-	0.055*** (0.019)		0.055** (0.022)	0.065*** (0.018)
× Outside migrants	-	0.057** (0.025)		0.057** (0.024)	0.088** (0.026)
Age (years) in wave 5	0.005** (0.002)	0.004* (0.002)		0.004* (0.002)	-
Age squared in wave 5	-0.000* (0.000)	-0.000 (0.000)		-0.000 (0.000)	-
Age in wave 1	-	-		-	0.001** (0.000)
Male dummy	0.075*** (0.024)	0.029 (0.023)		0.028 (0.028)	-
Schooling control	-	-		-0.001 (0.041)	-
Constant	10.113*** (0.082)	11.016*** (0.143)		10.646*** (0.288)	-1.488*** (0.077)
IHHFE	Yes	Yes		Yes	Yes
Location dum. (wave 1)	Yes	Yes		Yes	Yes
R-squared	0.532	0.562		0.562	0.639
No. of obs.	3645	3645		3645	3645

Notes: (1) Unit of observations is a person. (2) Sample is panel individuals. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity in OLS and estimated by bootstrapping with 100 replications in the control function (CF) approach. (5) Initial location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (6) IHHFE stands for initial household fixed effects.

Table 2.14: Appendix: Checking the Correlation between the Pre-migration Level of Schooling and Wealth

Dependent variables:	log of per capita non-food consumption (annual) in wave 1		log of per capita consumption (annual) in wave 1		log of household cultivated land (acre) in wave 1	
	OLS	OLS	OLS	OLS	OLS	OLS
	(a)	(b)	(c)	(d)	(e)	(f)
Individual education (years) in wave 1	0.016*** (0.005)	-	0.023*** (0.004)	-	0.013* (0.007)	-
HH average education (years) in wave 1	-	0.072*** (0.015)	-	0.095*** (0.014)	-	0.070*** (0.023)
Age (years) in wave 1	-0.001 (0.002)	0.001 (0.001)	-0.002 (0.002)	0.001 (0.001)	-0.005* (0.003)	-0.003* (0.002)
Age squared in wave 1	0.000* (0.000)	0.000 (0.000)	0.000** (0.000)	0.000 (0.000)	0.000** (0.000)	0.000** (0.000)
Male dummy	0.038** (0.018)	0.040** (0.017)	0.025 (0.016)	0.027* (0.015)	0.016 (0.026)	0.015 (0.025)
Constant	10.022*** (0.072)	9.843** (0.082)	10.434*** (0.064)	10.197*** (0.072)	1.941*** (0.087)	1.767*** (0.110)
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.091	0.112	0.120	0.165	0.182	0.192
No. of observations	3689	3689	3689	3689	3583	3583

Notes: (1) Unit of observations is a person. (2) Sample is all panel individuals. (3) Figures in () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity and clustered residuals within initial households. (5) Location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe.

Table 2.15: Appendix: 2SLS Estimations on Earnings

Dependent variable:	log of per worker income (weekly) in wave 5				
	OLS	2SLS	2SLS	2SLS	2SLS
	(a)	(b)	(c)	(d)	(e)
Migrants	0.339*** (0.072)	0.232 (0.277)	0.235 (0.392)	0.232 (0.409)	0.247 (0.287)
Education	0.042*** (0.009)	0.094*** (0.015)	0.247*** (0.052)	0.274*** (0.058)	0.275*** (0.057)
in wave 5 (years)	0.089*** (0.007)	0.078*** (0.007)	0.043*** (0.012)	0.037*** (0.013)	0.037*** (0.011)
Age in wave 5 (years)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Age squared in wave 5	0.238*** (0.057)	0.176** (0.071)	0.052 (0.109)	0.029 (0.115)	0.032 (0.101)
Male dummy	-0.287*** (0.106)	-0.296*** (0.103)	-0.248** (0.125)	-0.240* (0.130)	-0.235*** (0.086)
HH head in wave 1	-0.304*** (0.101)	-0.269** (0.126)	-0.046 (0.162)	-0.008 (0.169)	- (0.175)
HH head's spouse in wave 1	5.700*** (0.175)	4.245*** (0.282)	4.441*** (0.437)	4.702*** (0.676)	4.705*** (0.675)
Constant	Yes	Yes	Yes	Yes	Yes
IHHFE	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes
Joint significance of instruments in the first stage					
Migrants					
F-statistic	-	8.92	7.38	10.69	29.24
p-values	-	0.000	0.000	0.000	0.000
Education					
F-statistic	-	98.95	10.08	13.09	17.78
p-values	-	0.000	0.000	0.000	0.000
Sargan test (p-values)	-	0.001	0.490	0.568	0.769
Joint significance of migration and schooling effects					
F-test (p-values)	-	0.000	0.000	0.000	0.000
R-squared	0.528	0.519	0.396	0.358	0.358
No. of obs.	2562	2562	2562	2562	2562
Excluded instruments					
Education in wave 1	-	Yes	No	No	No
Age & age squared in wave 1	-	Yes	Yes	No	No
Rainfall shock	-	Yes	Yes	Yes	Yes
HH head in wave 1	-	No	No	No	No
HH head's spouse in wave 1	-	No	No	No	Yes
Young female in wave 1	-	Yes	Yes	Yes	Yes
Parents' education	-	Yes	Yes	Yes	Yes

Notes: (1) Unit of observations is a person. (2) Sample is panel workers. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity in OLS. (5) Initial location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (6) IHHFE stands for initial household fixed effects.

Table 2.16: Appendix: Impact Evaluation Results

Dependent variables: log of	per worker income (weekly)			per capita consumption (annual)		
	Level		Growth	Level		Growth
	(wave 5)		(wave 1 to 5)	(wave 5)		(wave 1 to 5)
	OLS	OLS	2SLS	OLS	OLS	2SLS
	(a)	(b)	(c)	(d)	(e)	(f)
(A) Pooling						
Migrants	0.405*** (0.067)	0.337*** (0.126)	0.248 (0.352)	0.338*** (0.034)	0.338*** (0.034)	0.334*** (0.112)
Individual characteristics at baseline (wave 1)						
Age (years)	0.049*** (0.005)	-0.079*** (0.013)	-0.082*** (0.013)	0.003 (0.002)	0.003 (0.002)	0.003 (0.002)
Age squared	-0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Education (years)	0.034*** (0.010)	-0.017 (0.017)	-0.015 (0.013)	0.002 (0.004)	0.002 (0.004)	0.002 (0.004)
Male dummy	0.275*** (0.050)	0.168* (0.093)	0.153* (0.082)	0.096*** (0.025)	0.096*** (0.025)	0.095*** (0.026)
Constant	7.115*** (0.166)	2.117*** (0.369)	0.654 (0.701)	10.888*** (0.035)	-0.353*** (0.035)	-0.103 (0.218)
IHHFE	Yes	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes	Yes
T-test (p-value)	0.495	0.855	0.752	0.527	0.527	0.823
Joint significance of instruments in the first stage						
Migrants						
F-statistic	-	-	11.14	-	-	38.60
p-values	-	-	0.000	-	-	0.000
Sargan test (p-values)	-	-	0.105	-	-	0.170
R-squared	0.519	0.675	-	0.506	0.586	-
No. of obs.	2579	1521	1521	3679	3679	3679
	OLS (g)	OLS (h)	2SLS (i)	OLS (j)	OLS (k)	2SLS (l)
(B) Inside and outside migrants						
Inside migrants	0.385*** (0.068)	0.302** (0.129)	0.182 (0.452)	0.245*** (0.032)	0.245*** (0.032)	0.117 (0.207)
Outside migrants	0.534*** (0.168)	0.550* (0.301)	-4.751 (4.361)	0.643*** (0.084)	0.643*** (0.084)	1.620 (1.012)
Individual characteristics at baseline (wave 1)						
Age (years)	0.049*** (0.005)	-0.080*** (0.013)	-0.086*** (0.017)	0.003 (0.002)	0.003 (0.002)	0.003 (0.002)
Age squared	-0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Education (years)	0.034*** (0.010)	-0.016 (0.017)	-0.019 (0.017)	0.002 (0.004)	0.002 (0.004)	0.005 (0.005)
Male dummy	0.271*** (0.050)	0.159* (0.093)	0.212* (0.117)	0.085*** (0.025)	0.085*** (0.025)	0.077** (0.031)
Constant	7.143*** (0.168)	2.158*** (0.369)	0.821 (0.904)	10.888*** (0.035)	-0.353*** (0.035)	-0.048 (0.236)
IHHFE	Yes	Yes	Yes	Yes	Yes	Yes
Location dum. (wave 1)	Yes	Yes	Yes	Yes	Yes	Yes
Joint significance of instruments in the first stage						
Inside migrants						
F-statistic	-	-	12.56	-	-	32.20
p-values	-	-	0.000	-	-	0.000
Outside migrants						
F-statistic	-	-	0.57	-	-	3.31
p-values	-	-	0.685	-	-	0.010
Sargan test (p-values)	-	-	0.291	-	-	0.249
R-squared	0.519	0.675	-	0.516	0.594	-
No. of obs.	2579	1521	1521	3679	3679	3679

Notes: (1) Unit of observations is a person. (2) Sample is panel workers in income regressions, whereas all panel individuals in consumption regressions. (3) Figures () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity in OLS. (5) T-test is the test for the null hypothesis that a coefficient on migration is equal to the suggested estimate in Beegle et al. (2008b), 0.36. (6) Remittance expenditures are excluded from the measure of consumption. (7) Initial location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. (8) IHHFE stands for initial household fixed effects.

Table 2.17: Appendix: Detailed Descriptive Statistics - Individual Level

Non-migrants					Migrants (inside)					Migrants (outside)				
Mean	Std.	Median	No. of obs.		Mean	Std.	Median	No. of obs.		Mean	Std.	Median	No. of obs.	
(1) Per capita annual consumption (US dollars, 1991 price)														
Wave 1	213.14	(139.26)	172.77	2426	221.23	(137.57)	186.43	972		229.20*	(142.52)	191.50	291	
Wave 2	192.49	(144.01)	164.43	2422	184.70	(114.85)	160.53	924		192.62	(128.94)	154.06	276	
Wave 3	176.16	(109.74)	149.44	2353	170.65	(97.44)	150.07	870		184.25	(107.77)	164.89	247	
Wave 4	172.47	(99.93)	150.15	2223	171.27	(93.29)	154.17	797		173.55	(107.91)	156.64	222	
Wave 5	214.91	(162.23)	173.96	2829	278.88***	(239.30)	204.85	1287		518.91***	(483.01)	374.30	352	
(2) Per capita annual consumption excluding remittances (US dollars, 1991 price)														
Wave 1	208.86	(136.23)	171.05	2426	217.77*	(135.17)	183.53	972		224.51*	(138.45)	189.48	291	
Wave 2	185.95	(141.11)	159.67	2422	178.48	(111.74)	154.27	924		186.12	(126.14)	149.93	276	
Wave 3	169.90	(105.47)	146.04	2353	164.91	(93.32)	146.72	870		177.50	(104.63)	161.70	247	
Wave 4	166.75	(97.06)	147.03	2223	165.65	(90.22)	147.88	797		168.33	(105.11)	154.02	222	
Wave 5	208.67	(154.09)	170.39	2825	273.81***	(233.29)	201.39	1287		503.40***	(464.01)	374.30	352	
(3) Per capita weekly income (US dollars, 1991 price)														
Wave 1	4.32	(8.46)	2.34	2401	4.35	(7.74)	2.62	967		4.51	(9.66)	2.53	280	
Wave 2	2.69	(3.12)	1.90	2401	2.60	(2.94)	1.87	908		4.38	(17.10)	1.72	264	
Wave 3	2.56	(2.75)	2.00	2318	2.39	(2.55)	1.95	863		2.54	(3.60)	1.91	245	
Wave 4	2.67	(4.35)	1.95	2197	2.59	(4.05)	1.86	790		2.45	(3.48)	1.75	218	
Wave 5	3.95	(7.81)	2.40	2780	3.99	(9.04)	1.84	1164		4.67	(9.19)	1.63	218	
(4) Per worker weekly income (US dollars, 1991 price)														
Wave 1	8.33	(20.22)	4.10	1338	7.75	(26.56)	3.19	507		7.50	(22.93)	2.57	156	
Wave 2	4.71	(6.77)	2.91	1407	4.10	(10.98)	1.97	545		2.86***	(4.89)	1.60	152	
Wave 3	4.29	(6.36)	2.63	1448	3.45***	(6.52)	2.00	553		2.95***	(5.32)	1.55	133	
Wave 4	4.53	(9.13)	2.83	1344	3.50***	(6.61)	1.96	478		2.83***	(5.10)	1.50	117	
Wave 5	7.28	(16.38)	3.67	2114	8.59*	(19.76)	4.60	862		14.43***	(22.23)	10.63	120	
(5) Per worker weekly hours worked														
Wave 1	24.75	(16.80)	23.00	1352	19.39***	(15.24)	15.00	509		18.71***	(15.37)	12.70	159	
Wave 2	21.51	(16.79)	18.00	1429	17.40***	(16.42)	12.00	555		14.45***	(13.91)	11.00	159	
Wave 3	20.06	(15.63)	17.00	1470	16.48***	(14.57)	12.00	559		14.89***	(14.28)	9.00	135	
Wave 4	20.23	(16.63)	17.50	1357	15.57***	(15.87)	10.70	487		13.78***	(13.65)	10.10	118	
Wave 5	28.42	(19.52)	25.00	2157	29.99**	(20.55)	25.00	947		49.76***	(26.08)	54.00	189	

Non-migrants				Migrants (inside)				Migrants (outside)				
Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	
(6) Per worker hourly income (US dollars, 1991 price)												
Wave 1	0.37	(0.92)	0.20	1337	0.44	(1.50)	0.21	507	0.38	(0.68)	0.19	156
Wave 2	0.28	(0.53)	0.17	1407	0.32	(0.80)	0.17	545	0.28	(0.42)	0.16	152
Wave 3	0.27	(0.33)	0.17	1448	0.26	(0.33)	0.17	553	0.32	(0.84)	0.21	133
Wave 4	0.28	(0.49)	0.18	1343	0.28	(0.38)	0.17	477	0.22**	(0.25)	0.15	117
Wave 5	0.30	(0.59)	0.16	2113	0.35**	(0.62)	0.19	862	0.41**	(0.51)	0.22	119
(7) Gender (one if male)												
Wave 1-4	0.52	(0.49)	-	9425	0.37***	(0.48)	-	3564	0.48**	(0.50)	-	1038
Wave 5	0.52	(0.49)	-	2832	0.38***	(0.48)	-	1289	0.47	(0.50)	-	355
(8) Education (years)												
Own												
Wave 1-4	2.56	(3.08)	0.00	9426	2.73***	(3.15)	1.00	3564	3.12***	(3.34)	2.00	1038
Wave 5	4.69	(3.22)	6.00	2811	5.49***	(3.35)	7.00	1280	6.70***	(3.49)	7.00	350
Average of household												
Wave 1-4	2.67	(1.46)	2.55	9426	2.74**	(1.49)	2.61	3564	2.99***	(1.45)	2.80	1038
Wave 5	3.48	(1.94)	3.44	2808	3.90***	(2.43)	3.50	1280	5.45***	(2.95)	5.25	347
Household head												
Wave 1-4	4.39	(3.22)	4.00	9426	4.32	(3.39)	4.00	3564	4.61*	(3.48)	4.00	1038
Wave 5	4.88	(3.46)	6.00	2811	6.12***	(3.80)	7.00	1280	8.06***	(4.32)	7.00	350
(9) Age (years)												
Own												
Wave 1-4	22.29	(19.28)	16.00	9425	15.33***	(12.13)	13.00	3564	14.85***	(11.81)	13.00	1038
Wave 5	33.00	(19.36)	27.00	2832	26.99***	(12.33)	25.00	1289	26.27***	(11.30)	25.00	355
Average of household												
Wave 1-4	21.66	(8.23)	19.90	9425	21.55	(7.12)	20.14	3564	21.82	(7.02)	20.79	1038
Wave 5	25.18	(11.11)	22.63	2832	21.79***	(9.23)	19.20	1289	21.43***	(7.15)	20.00	355
Household head												
Wave 1-4	48.16	(15.31)	48.00	9426	49.92***	(16.03)	50.00	3564	52.21***	(16.50)	53.00	1038
Wave 5	49.72	(17.52)	49.00	2811	39.48***	(15.22)	35.00	1280	35.62***	(13.56)	34.00	350

Notes: (1) Unit of observations is a person-wave. (2) In per capita measures, all members in the same household-wave group are recorded with the same numerical values. (3) The equality of means between non-migrants and migrants (inside) and that between non-migrants and migrants (outside) are examined by T-tests, assuming unequal variance. The degree of freedom is approximated by Satterthwaite. *** denotes the rejection of equality at 1%, ** at 5%, and * at 10%.

Table 2.18: Appendix: Detailed Descriptive Statistics - Household Level

Non-migrants				Migrants (inside)				Migrants (outside)			
Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.
(1) Household annual consumption (US dollars, 1991 price)											
Wave 1-4	955.44	(811.86)	745.39	1282	(903.24)	1005.80	1819	1445.38***	(938.73)	1236.74	620
Wave 5	983.77	(760.58)	808.49	1354	(822.24)	886.73	1094	1669.54***	(1141.01)	1457.44	297
(2) Per capita annual consumption (US dollars, 1991 price)											
Wave 1-4	264.65	(231.76)	201.96	1282	(124.16)	163.49	1819	202.47***	(135.23)	170.40	620
Wave 5	234.37	(187.37)	182.86	1354	(248.18)	210.02	1094	546.22***	(501.68)	399.76	297
(3) Household weekly income (US dollars, 1991 price)											
Total											
Wave 1-4	15.33	(27.32)	9.48	1258	(48.56)	13.21	1803	25.24***	(72.69)	14.68	607
Wave 5	17.70	(35.98)	11.52	1312	(29.68)	7.98	986	11.63***	(19.30)	7.32	183
Farm wage employed											
Wave 1-4	2.64	(3.15)	1.71	106	(6.01)	1.14	147	2.16	(2.08)	2.13	41
Wave 5	4.70	(4.03)	3.47	160	(5.65)	3.39	98	7.79	(5.55)	6.15	7
Non-farm wage employed											
Wave 1-4	9.06	(11.08)	5.90	303	(13.09)	5.63	423	10.97*	(11.89)	6.94	174
Wave 5	14.34	(19.42)	9.75	289	(34.41)	10.92	215	17.82	(25.42)	13.36	85
Farm self-employed											
Wave 1-4	10.71	(20.28)	6.99	1230	(45.55)	10.53	1798	18.61***	(70.77)	10.60	609
Wave 5	9.52	(31.33)	6.58	1279	(7.40)	5.33	889	3.46***	(3.41)	2.50	106
Non-farm self-employed											
Wave 1-4	12.40	(31.62)	4.33	264	(26.85)	3.78	438	12.25	(27.02)	4.23	182
Wave 5	14.36	(21.37)	7.66	448	(36.92)	10.51	189	17.49	(12.50)	15.38	11
(4) Household weekly hours worked											
Total											
Wave 1-4	60.09	(44.46)	51.00	1223	(56.57)	69.65	1770	86.50***	(59.41)	76.00	607
Wave 5	70.55	(49.63)	61.00	1292	(44.58)	42.00	903	53.76***	(32.48)	54.00	186
Farm wage employed											
Wave 1-4	18.06	(15.72)	14.00	109	(15.43)	12.00	152	15.72	(11.47)	12.00	43
Wave 5	22.24	(19.29)	16.00	161	(20.10)	13.00	99	39.00***	(10.96)	39.00	8
Non-farm wage employed											
Wave 1-4	43.14	(32.03)	40.00	306	(29.56)	40.00	440	53.84***	(37.83)	45.00	179
Wave 5	45.67	(31.55)	42.00	307	(30.46)	45.00	214	54.83***	(27.64)	55.00	83
Farm self-employed											
Wave 1-4	44.12	(36.69)	35.00	1124	(47.95)	52.60	1727	60.82***	(47.32)	50.00	591
Wave 5	49.80	(41.51)	41.00	1171	(36.74)	27.00	667	16.27***	(12.07)	16.00	33
Non-farm self-employed											
Wave 1-4	28.97	(29.19)	18.00	301	(34.69)	18.00	475	31.39	(31.73)	22.00	199
Wave 5	36.59	(33.93)	26.00	416	(31.46)	30.50	276	55.43***	(33.14)	57.00	83

Non-migrants				Migrants (inside)				Migrants (outside)				
Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	
(5) Per capita weekly income (US dollars, 1991 price)												
Total												
Wave 1-4	4.31	(7.79)	2.56	1258	3.25***	(6.94)	2.10	1803	3.58	(11.70)	2.07	607
Wave 5	4.18	(6.40)	2.60	1312	4.30	(13.03)	1.85	986	6.29*	(17.06)	1.77	183
Farm wage employed												
Wave 1-4	0.85	(1.23)	0.38	106	0.54**	(1.06)	0.19	147	0.37***	(0.46)	0.21	41
Wave 5	1.45	(2.17)	0.89	160	1.63	(1.98)	0.96	98	2.62	(2.53)	2.05	7
Non-farm wage employed												
Wave 1-4	2.80	(4.03)	1.59	303	1.36***	(2.05)	0.80	423	1.44***	(1.60)	0.91	174
Wave 5	4.20	(6.09)	2.39	289	7.12***	(12.41)	3.35	215	11.53***	(23.78)	5.63	85
Farm self-employed												
Wave 1-4	2.95	(5.85)	1.94	1230	2.51*	(6.54)	1.68	1798	2.68	(11.46)	1.52	609
Wave 5	2.04	(4.15)	1.48	1279	1.56***	(1.52)	1.16	889	0.76***	(0.74)	0.51	106
Non-farm self-employed												
Wave 1-4	3.41	(9.15)	1.00	264	1.66***	(3.77)	0.55	438	1.84**	(4.32)	0.53	182
Wave 5	3.39	(5.69)	1.61	448	6.32*	(22.76)	2.08	189	6.62	(7.62)	3.66	11
(6) Per capita weekly hours worked												
Total												
Wave 1-4	16.20	(11.70)	13.65	1218	12.37***	(7.72)	10.90	1770	12.04***	(8.07)	10.27	607
Wave 5	16.65	(11.78)	14.00	1292	15.17***	(14.21)	10.66	903	23.45***	(21.22)	16.33	186
Farm wage employed												
Wave 1-4	6.40	(9.34)	3.33	109	3.18***	(3.42)	2.20	152	2.58***	(2.18)	1.81	43
Wave 5	6.88	(10.39)	3.66	161	7.39	(8.60)	4.00	99	14.17	(12.64)	11.00	8
Non-farm wage employed												
Wave 1-4	14.36	(14.40)	9.00	305	6.63***	(6.50)	5.33	440	7.27***	(5.78)	5.83	179
Wave 5	13.49	(14.14)	9.00	307	20.05***	(18.62)	14.20	214	30.80***	(23.94)	25.00	83
Farm self-employed												
Wave 1-4	11.00	(7.61)	9.50	1120	9.51***	(6.32)	8.42	1727	8.30***	(5.49)	7.37	591
Wave 5	10.79	(7.63)	9.00	1171	8.22***	(6.75)	6.20	667	4.82***	(5.22)	3.00	33
Non-farm self-employed												
Wave 1-4	7.77	(9.64)	4.00	299	4.35***	(4.84)	2.50	475	4.97***	(7.01)	3.00	199
Wave 5	8.73	(10.33)	5.00	416	11.57***	(13.34)	7.00	276	18.47***	(16.92)	13.16	83

Non-migrants				Migrants (inside)				Migrants (outside)				
Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	
(7) Household remittances sent in the past 6 months (US dollars, 1991 price)												
Wave 1-4	28.20	(144.05)	7.27	1108	31.22	(108.36)	9.14	1605	37.75	(113.68)	13.85	565
Wave 5	21.80	(102.10)	5.29	1150	12.65***	(34.02)	3.97	793	30.65	(65.96)	9.00	174
(8) Household remittances received in the past 6 months (US dollars, 1991 price)												
Wave 1-4	44.79	(252.76)	10.07	1024	84.78*	(882.68)	14.66	1532	128.42	(1184.70)	24.57	553
Wave 5	30.20	(108.84)	8.13	1192	18.66***	(74.08)	4.55	789	17.98**	(48.39)	4.29	145
(9) Household non-labour income in the past 6 months (US dollars, 1991 price)												
Wave 1-4	83.17	(422.84)	13.89	783	126.53	(987.19)	15.03	1161	195.05	(1552.57)	23.46	433
Wave 5	78.15	(223.51)	16.20	185	159.26	(853.78)	17.86	118	133.79	(432.45)	23.91	26
(10) Average education of households (years)												
Wave 1-4	2.87	(2.12)	2.66	1290	2.73**	(1.56)	2.60	1820	3.06**	(1.53)	2.83	621
Wave 5	3.39	(2.14)	3.33	1355	3.89***	(2.46)	3.50	1096	5.64***	(2.90)	5.45	297
(11) Education of household heads (years)												
Wave 1-4	4.10	(3.47)	4.00	1290	4.20	(3.35)	4.00	1820	4.60***	(3.49)	4.00	621
Wave 5	4.96	(3.48)	7.00	1357	6.12***	(3.74)	7.00	1097	8.08***	(4.31)	7.00	300
(12) Average age of households (years)												
Wave 1-4	30.56	(16.67)	25.89	1290	22.63***	(8.23)	20.87	1820	22.57***	(7.89)	21.00	621
Wave 5	24.92	(12.92)	21.00	1357	21.73***	(9.52)	19.00	1097	21.57***	(7.50)	20.07	300
(13) Age of household heads (years)												
Wave 1-4	47.93	(17.64)	47.00	1290	49.72***	(16.72)	50.00	1820	52.43***	(16.46)	53.00	621
Wave 5	45.18	(18.41)	40.00	1350	38.70***	(15.08)	34.00	1090	35.40***	(12.57)	32.00	293
(14) Proportion of workers in households												
Wave 1-4	0.78	(0.38)	0.75	981	0.69***	(0.31)	0.66	1820	0.68***	(0.29)	0.66	621
Wave 5	0.68	(0.38)	0.66	1355	0.46***	(0.40)	0.33	1097	0.30***	(0.37)	0.20	300
(15) Household size												
Wave 1-4	4.13	(2.30)	4.00	1283	6.78***	(3.21)	6.00	1820	7.62***	(3.69)	7.00	621
Wave 5	4.79	(2.68)	4.00	1357	4.59*	(2.55)	4.00	1097	4.21***	(2.61)	4.00	300

Notes: (1) Unit of observations is a household-wave. (2) Households are defined as migrants in wave 1 to 4 if they supply at least one member who moved out of original community between 1991 and 2004 and in wave 5 if they contain at least one panel migrant in them. (3) Remittances take the form of either gift or loan. (4) The equality of means between non-migrants and migrants (inside) and that between non-migrants and migrants (outside) are examined by T-tests, assuming unequal variance. The degree of freedom is approximated by Satterthwaite. *** denotes the rejection of equality at 1%, ** at 5%, and * at 10%.

Table 2.19: Appendix: Per Worker Weekly Income, Weekly Hours Worked, and Hourly Income by Occupation

	Non-migrants				Migrants (inside)				Migrants (outside)			
	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.
(1) Per worker weekly income (US dollars, 1991 price)												
Farm wage employed												
Wave 1	5.38	(6.78)	3.08	55	3.16*	(2.73)	2.66	17	2.61**	(0.97)	2.61	2
Wave 2	1.92	(2.96)	0.59	37	0.94	(1.18)	0.40	13	-	-	-	-
Wave 3	1.91	(3.46)	0.42	54	2.00	(1.98)	1.64	15	1.62	(2.26)	1.62	2
Wave 4	1.64	(2.26)	0.75	43	0.46***	(0.74)	0.14	14	1.15	(1.58)	1.15	2
Wave 5	4.15	(3.73)	2.75	138	4.02	(3.19)	3.33	75	7.95	(6.06)	5.86	6
Non-farm wage employed												
Wave 1	10.23	(18.50)	5.56	113	9.27	(6.82)	7.63	28	8.99	(12.57)	6.81	13
Wave 2	6.44	(6.48)	5.08	114	6.72	(8.84)	4.44	37	8.20	(16.71)	2.11	8
Wave 3	5.68	(7.30)	4.25	122	5.71	(5.69)	4.39	31	3.56**	(2.00)	3.99	8
Wave 4	6.81	(6.25)	5.95	101	6.89	(8.25)	4.70	29	7.87	(12.44)	3.71	8
Wave 5	12.06	(14.74)	8.40	251	15.59	(24.03)	10.77	144	17.86*	(25.54)	12.75	84
Farm self-employed												
Wave 1	6.67	(17.27)	3.38	1269	6.99	(26.84)	2.70	486	7.07	(24.31)	2.41	148
Wave 2	3.59	(3.86)	2.57	1324	2.94***	(4.19)	1.69	520	2.26***	(2.32)	1.44	143
Wave 3	3.26	(4.00)	2.24	1377	2.61***	(2.59)	1.78	530	2.29***	(2.38)	1.32	130
Wave 4	3.36	(5.47)	2.35	1271	2.77**	(4.44)	1.74	456	2.25***	(3.58)	1.49	111
Wave 5	4.13	(13.18)	2.56	1852	4.71	(5.37)	3.44	687	3.90	(3.76)	2.83	30
Non-farm self-employed												
Wave 1	15.29	(36.27)	5.57	80	11.08	(17.07)	4.07	20	4.38**	(4.18)	2.74	4
Wave 2	9.67	(16.55)	3.44	113	16.85	(38.27)	3.09	30	5.09	(8.68)	1.45	9
Wave 3	6.96	(13.04)	2.48	135	6.69	(18.52)	1.45	48	7.49	(18.47)	1.08	9
Wave 4	7.16	(20.83)	2.38	149	5.55	(14.39)	2.11	37	2.71**	(2.47)	2.29	6
Wave 5	12.89	(19.25)	6.46	322	20.28	(49.50)	6.22	80	13.40	(7.43)	15.38	5

	Non-migrants				Migrants (inside)				Migrants (outside)			
	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.	Mean	Std.	Median	No. of obs.
(2) Per worker weekly hours worked												
Farm wage employed												
Wave 1	15.32	(10.77)	12.00	56	13.61	(9.82)	12.00	18	18.50	(19.09)	18.50	2
Wave 2	15.68	(12.65)	12.00	38	12.84	(7.69)	10.00	13	-	-	-	-
Wave 3	11.47	(8.81)	9.00	55	8.33*	(4.56)	8.00	15	6.50*	(2.12)	6.50	2
Wave 4	16.27	(20.05)	12.00	43	9.85*	(4.81)	11.00	14	17.50	(17.67)	17.50	2
Wave 5	19.10	(16.00)	12.00	141	17.40	(14.63)	12.00	76	43.00***	(9.38)	42.50	6
Non-farm wage employed												
Wave 1	36.82	(21.53)	40.00	113	37.57	(16.15)	40.00	28	37.46	(22.14)	45.00	13
Wave 2	34.00	(20.14)	35.00	122	42.65**	(23.29)	42.50	39	27.87	(15.69)	30.50	8
Wave 3	32.08	(19.44)	30.00	130	35.96	(24.37)	33.50	32	32.37	(15.35)	28.50	8
Wave 4	34.19	(20.15)	35.00	103	37.48	(26.97)	38.00	30	29.56	(14.64)	31.50	8
Wave 5	39.36	(24.23)	40.00	269	45.46**	(24.88)	45.00	147	53.36***	(24.50)	55.00	82
Farm self-employed												
Wave 1	20.34	(13.05)	18.50	1277	16.19***	(12.03)	12.00	488	15.45***	(11.77)	12.00	150
Wave 2	16.90	(11.97)	15.00	1334	13.27***	(11.14)	10.00	527	12.06***	(10.01)	9.00	148
Wave 3	15.58	(11.45)	12.35	1382	13.39***	(10.97)	10.00	534	12.16***	(11.19)	8.00	131
Wave 4	15.28	(10.36)	14.00	1275	11.72***	(9.23)	8.15	462	10.87***	(7.64)	9.25	112
Wave 5	19.87	(12.64)	18.00	1876	19.70	(11.57)	18.00	695	13.06***	(8.91)	12.00	32
Non-farm self-employed												
Wave 1	22.19	(19.43)	15.00	111	29.17	(21.46)	25.00	23	16.62	(10.68)	11.50	8
Wave 2	24.27	(21.62)	17.75	142	21.88	(21.10)	14.00	38	22.34	(27.55)	14.00	13
Wave 3	17.99	(17.65)	12.00	175	13.54*	(14.86)	10.00	58	13.20	(12.43)	7.00	11
Wave 4	19.16	(24.24)	10.00	195	18.82	(22.42)	9.00	48	19.50	(17.94)	14.00	7
Wave 5	29.16	(23.26)	21.50	368	33.70**	(24.64)	27.00	199	53.09***	(25.41)	57.00	82

Non-migrants					Migrants (inside)				
Mean	Std.	Median	No. of obs.		Mean	Std.	Median	No. of obs.	
					Migrants (outside)				
Mean	Std.	Median	No. of obs.		Mean	Std.	Median	No. of obs.	
(3) Hourly income (US dollars, 1991 price)									
Farm wage employed									
Wave 1	0.68	(1.41)	0.21	55	0.24**	(0.20)	0.21	17	0.24
Wave 2	0.22	(0.46)	0.05	37	0.08	(0.11)	0.03	13	-
Wave 3	0.26	(0.62)	0.04	54	0.33	(0.36)	0.10	15	0.32
Wave 4	0.11	(0.17)	0.05	43	0.04**	(0.05)	0.02	14	0.04
Wave 5	0.33	(0.40)	0.19	138	0.39	(0.44)	0.23	75	0.17**
Non-farm wage employed									
Wave 1	0.39	(0.63)	0.18	112	0.34	(0.50)	0.25	28	0.26
Wave 2	0.30	(0.86)	0.14	114	0.18	(0.26)	0.09	37	0.25
Wave 3	0.21	(0.26)	0.11	122	0.24	(0.45)	0.15	31	0.14
Wave 4	0.26	(0.40)	0.15	100	0.18	(0.16)	0.13	28	0.28
Wave 5	0.52	(0.86)	0.25	250	0.63	(1.55)	0.22	144	0.85
Farm self-employed									
Wave 1	0.36	(0.93)	0.19	1269	0.45	(1.53)	0.21	486	0.39
Wave 2	0.29	(0.55)	0.17	1324	0.32	(0.78)	0.17	520	0.29
Wave 3	0.28	(0.44)	0.18	1377	0.26	(0.32)	0.17	530	0.27
Wave 4	0.29	(0.58)	0.19	1271	0.30	(0.39)	0.19	456	0.23
Wave 5	0.27	(0.50)	0.14	1852	0.35**	(0.81)	0.19	687	0.43
Non-farm self-employed									
Wave 1	1.06	(0.45)	0.31	80	0.46**	(0.62)	0.21	20	0.29***
Wave 2	0.45	(0.75)	0.24	113	1.81	(7.67)	0.20	30	0.25**
Wave 3	0.39	(0.49)	0.18	135	0.36	(0.48)	0.22	48	1.23
Wave 4	0.40	(0.64)	0.19	149	0.31	(0.35)	0.16	37	0.14***
Wave 5	0.75	(1.51)	0.28	322	0.76	(1.40)	0.29	80	0.42**

Notes: (1) Unit of observations is a person-wave. (2) The equality of means between non-migrants and migrants (inside) and between non-migrants and migrants (outside) is examined by T-tests, assuming unequal variance. The degree of freedom is approximated by Satterthwaite. *** denotes the rejection of equality at 1%, ** at 5%, and * at 10%.

Table 2.20: Appendix: Diversification of Activities - Individual Level (Aged 23 to 40)

	Wave 1 to 4			Wave 5		
	Non-migrants	Migrants (inside)	Migrants (outside)	Non-migrants	Migrants (inside)	Migrants (outside)
Single activity	0.65	0.58	0.52	0.58	0.64	0.67
Two activities	0.20	0.22	0.21	0.24	0.17	0.03
Three activities	0.01	0.01	0.01	0.02	0.00	0.00
All activities	-	0.00	0.01	0.00	0.00	0.00
Nothing	0.12	0.16	0.23	0.14	0.16	0.27
Breakdown						
Farm wage employment only	0.00	0.00	-	0.01	0.01	0.02
Non-farm wage employment only	0.04	0.06	0.14	0.07	0.10	0.30
Farm self-employment only	0.56	0.48	0.34	0.41	0.40	0.08
Non-farm self-employment only	0.04	0.04	0.03	0.08	0.11	0.26
Farm wage, non-farm wage, farm self-emp.	0.00	-	-	0.00	-	-
Farm wage, non-farm wage, non-farm self-emp.	-	-	-	-	-	-
Farm wage, farm self-emp., non-farm self-emp.	0.00	0.00	-	0.00	0.00	-
Non-farm wage, farm self-emp., non-farm self-emp.	0.00	0.01	0.02	0.00	0.00	0.00
Farm wage, non-farm wage	-	-	-	0.00	0.00	-
Farm wage, farm self-emp.	0.03	0.02	0.01	0.04	0.05	-
Farm wage, non-farm self-emp.	-	0.00	-	-	0.00	-
Non-farm wage, farm self-emp.	0.06	0.08	0.09	0.06	0.02	0.00
Non-farm wage, non-farm self-emp.	0.03	0.00	-	0.01	0.00	0.01
Farm self-emp., non-farm self-emp.	0.10	0.10	0.10	0.12	0.8	0.02
All activities	-	0.00	0.01	0.00	0.00	0.00
Nothing	0.12	0.16	0.23	0.14	0.16	0.27
No. of observations	1640	350	82	947	685	194

Notes: (1) Unit of observations is a person-wave (aged 23 to 40 years old). (2) Reference periods are the past seven days. (3) The number is the proportion relative to the total number of observations in each category.

Table 2.21: Appendix: Diversification of Activities - Household Level

	Wave 1 to 4			Wave 5		
	Non-migrants	Migrants (inside)	Migrants (outside)	Non-migrants	Migrants (inside)	Migrants (outside)
Single activity	0.48	0.51	0.45	0.47	0.53	0.56
Two activities	0.30	0.35	0.40	0.39	0.24	0.04
Three activities	0.05	0.09	0.12	0.08	0.03	0.01
All activities	0.10	0.00	0.00	0.00	0.00	-
Nothing	0.04	0.02	0.02	0.04	0.17	0.38
Breakdown						
Farm wage employment only	0.00	0.00	-	0.00	0.01	0.02
Non-farm wage employment only	0.03	0.00	0.00	0.03	0.09	0.24
Farm self-employment only	0.41	0.49	0.42	0.39	0.34	0.07
Non-farm self-employment only	0.02	0.00	0.01	0.03	0.08	0.22
Farm wage, non-farm wage, farm self-emp.	0.01	0.01	0.00	0.01	0.00	-
Farm wage, non-farm wage, non-farm self-emp.	-	-	-	-	-	-
Farm wage, farm self-emp., non-farm self-emp.	0.00	0.01	0.02	0.02	0.01	0.00
Non-farm wage, farm self-emp., non-farm self-emp.	0.03	0.06	0.08	0.04	0.02	0.01
Farm wage, non-farm wage	-	-	-	0.00	0.00	-
Farm wage, farm self-emp.	0.04	0.05	0.03	0.07	0.05	0.00
Farm wage, non-farm self-emp.	-	-	-	-	0.00	-
Non-farm wage, farm self-emp.	0.11	0.14	0.18	0.11	0.06	0.00
Non-farm wage, non-farm self-emp.	0.00	0.00	-	0.00	0.01	0.01
Farm self-emp., non-farm self-emp.	0.13	0.15	0.18	0.19	0.11	0.02
All activities	0.10	0.00	0.00	0.00	0.00	-
Nothing	0.04	0.02	0.02	0.04	0.17	0.38
No. of observations	1443	1820	621	1357	1097	300

Notes: (1) Unit of observations is a household-wave. (2) Households are defined as migrants in wave 1 to 4 if they supply at least one member who moved out of original community between 1991 and 2004 and in wave 5 if they contain at least one panel migrant in them. (3) The number is the proportion relative to the total number of observations in each category.

Table 2.22: Appendix: Correlation between Income and Consumption (OLS)

Dependent variable:	Log of household consumption (annual)					
	Wave 1 to 4			Wave 5		
	Non-migrants (a)	Migrants (inside) (b)	Migrants (outside) (c)	Non-migrants (d)	Migrants (inside) (e)	Migrants (outside) (f)
Log of household income (weekly)	0.589*** (0.020)	0.543*** (0.025)	0.500*** (0.040)	0.449*** (0.032)	0.249*** (0.028)	-0.026 (0.030)
Remittances received in the past 6 months (dummy, one if exist)	0.058* (0.035)	0.067** (0.029)	0.102 (0.072)	0.105* (0.057)	-0.091** (0.042)	0.211 (0.137)
Non-labour income in the past 6 months (dummy, one if exist)	0.123*** (0.028)	0.072*** (0.022)	0.134*** (0.045)	0.119*** (0.037)	0.118* (0.065)	0.229 (0.164)
Constant	7.190*** (0.159)	8.070*** (0.225)	8.166*** (0.328)	8.251*** (0.268)	9.861*** (0.207)	12.538*** (0.231)
Location dummies	Yes	Yes	Yes	Yes	Yes	-
Time dummies	Yes	Yes	Yes	-	-	-
R-squared	0.631	0.637	0.545	0.499	0.285	0.022
No. of observations	1258	1803	607	1312	986	182

Notes: (1) Unit of observations is a household-wave. (2) Figures in () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity. (4) Households are defined as migrants in wave 1 to 4 if they supply at least one member who moved out of original community between 1991 and 2004 and in wave 5 if they contain at least one panel migrant in them. (5) Remittances take the form of either gift or loan. (6) Non-labour income contains income from pension or retirement funds, insurances, interest, lottery winnings, dowry, inheritance, sale of durable goods, sale of house or land, and so on. (7) Except for outside migrants in wave 5, location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe.

Table 2.23: Appendix: Empirical Test for Latent Income in Wave 5

Dependent variables:	Non-migrants			Migrants (inside)			Migrants (outside)		
	Difference in log of per capita consumption (annual) between wave 4 and 5			Log of per capita consumption (annual) in wave 5			Log of per capita consumption (annual) in wave 5		
	Eq (2.35)			Eq (2.35)			Eq (2.35)		
	OLS (a)	2SLS (b)	2SLS (d)	OLS (e)	2SLS (f)	2SLS (g)	OLS (i)	2SLS (j)	2SLS (k)
Log of per capita income (weekly) in wave 5	0.413*** (0.028)	0.415*** (0.038)	-	0.273*** (0.027)	0.412*** (0.058)	0.088 (0.163)	0.188*** (0.034)	0.201*** (0.091)	-0.254 (0.237)
Difference in log of per capita income (weekly)	-	-	0.548*** (0.166)	-	-	-	-	-	-
Transformed term of log of per capita income (weekly)	-	-	-	-	-	-	-	-	-
Constant	7.684*** (0.185)	7.994*** (0.252)	0.063*** (0.019)	8.781*** (0.189)	8.140*** (0.380)	10.260*** (1.039)	10.167*** (0.212)	10.119*** (0.643)	12.723*** (1.321)
Location dummies (wave 5)	Yes	Yes	-	Yes	Yes	Yes	-	-	-
Instrument sets	-	(A)	(B)	-	(A)	(C)	-	(A)	(D)
Joint significance of instruments in the first-stage	-	-	-	-	-	-	-	-	-
F-statistic	-	164.61	6.15	-	12.83	1.28	-	69.65	1.18
p-values	-	0.000	0.002	-	0.000	0.265	-	0.000	0.324
Hansen's J statistic (p-value)	-	0.382	0.337	-	0.196	0.501	-	0.114	0.388
R-squared	0.483	-	0.435	0.380	-	-	0.135	-	-
No. of observations	2780	1652	2024	1164	533	707	217	71	127
									114

Notes: (1) Unit of observations is a person. (2) Sample is all panel individuals except for (b), (f), and (j) where sample is panel workers. (3) Figures in () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity and clustered residuals within each (current) household. (5) Instrument set (A) contains log of per worker income in wave 5 and location dummies in wave 4. (6) Instrument set (B) contains log of per capita income and log of household size in wave 3. (7) Instrument set (C) contains log of per capita income and location dummies in wave 4. (8) Instrument set (D) contains log of household size in wave 4 and the difference in log of per capita income between wave 3 and 4. (9) Except for outside migrants in wave 5, location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe.

Table 2.24: Appendix: Per Worker Wage Earnings Functions (OLS)

Dependent variable:	Log of per worker wage earnings (weekly)					
	Wave 1 to 4			Wave 5		
	Non	Migrants	Migrants	Non	Migrants	Migrants
	-migrants	(inside)	(outside)	-migrants	(inside)	(outside)
	(a)	(b)	(c)	(d)	(e)	(f)
Age (years)	0.018 (0.026)	0.119** (0.058)	-0.225 (0.376)	0.059*** (0.017)	0.015 (0.026)	-0.002 (0.048)
Age squared	-0.000 (0.000)	-0.001** (0.000)	0.005 (0.006)	-0.000*** (0.000)	-0.000 (0.000)	0.000 (0.000)
Education (years)	0.077*** (0.028)	0.012 (0.037)	-0.035 (0.303)	0.049*** (0.017)	0.023 (0.019)	-0.120* (0.0695)
Education squared	-	-	-	-	-	0.011** (0.005)
Tenure (years)	0.086*** (0.023)	0.102 (0.074)	0.109 (0.152)	-	-	-
Tenure squared	-0.002*** (0.000)	-0.003 (0.003)	-0.009 (0.009)	-	-	-
Gender (dummy, one if male)	0.281 (0.172)	0.338 (0.292)	-1.104 (0.872)	0.215*** (0.144)	0.350** (0.144)	0.266 (0.209)
Constant	5.163*** (0.825)	4.566* (1.788)	1.043 (3.184)	6.197*** (0.456)	7.048*** (0.815)	7.189*** (1.095)
Employers' types (dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	-
Time dummies	Yes	Yes	Yes	-	-	-
R-squared	0.382	0.533	0.843	0.307	0.506	0.339
No. of observations	639	184	43	387	217	89

Notes: (1) Unit of observations is a person-wave. (2) Sample is panel workers. (3) Figures in () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity. (5) Information on tenure was not collected in wave 5. (6) Employers are classified into 6 categories: the government, state-owned companies, private employers, cooperative societies, religious institutions, and other. (7) Sectors are classified into farming/livestock and 12 non-farming categories. (8) In wave 5 excluding outside migrants and wave 1 to 4, location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe.

Table 2.25: Appendix: Per Worker Farming Self-employment Income Functions (OLS)

Dependent variable:	Log of per worker farm self-employment income (weekly)			
	Wave 1 to 4		Wave 5	
	Non -migrants (a)	Migrants (inside & outside) (b)	Non -migrants (c)	Migrants (inside & outside) (d)
Age (years)	0.014*** (0.001)	0.015*** (0.002)	0.011*** (0.002)	0.013* (0.006)
Education (years)	0.051*** (0.007)	0.015 (0.011)	0.055*** (0.011)	0.051** (0.022)
Gender (dummy, one if male)	-0.086** (0.041)	-0.002 (0.065)	-0.047 (0.072)	-0.396*** (0.146)
Log of household capital for farming	0.024 (0.016)	0.058** (0.025)	0.054* (0.029)	-0.008 (0.056)
Log of household cultivated land (acre)	0.101*** (0.031)	0.137*** (0.051)	0.060 (0.061)	0.127 (0.104)
Log of household livestock values (TSHS)	0.015 (0.014)	-0.022 (0.022)	0.064** (0.029)	0.144*** (0.038)
Constant	5.610*** (0.201)	5.380*** (0.254)	4.331*** (0.316)	5.435*** (0.709)
Location dummies	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	-	-
R-squared	0.179	0.103	0.153	0.259
No. of observations	2477	1094	656	199

Notes: (1) Unit of observations is a person-wave. (2) Sample is panel workers. (3) Figures in () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity. (5) Household capital is farm equipment (e.g., farm buildings, hand mills, tractors and so on) excluding hand tools. (6) In wave 1 to 4, location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. In wave 5, the category of outside Kagera is added to these categories.

Table 2.26: Appendix: Per Worker Non-farm Self-employment Income Functions (OLS)

Dependent variable:	Log of per worker non-farm self-employment income (weekly)			
	Wave 1 to 4		Wave 5	
	Non -migrants (a)	Migrants (inside & outside) (b)	Non -migrants (c)	Migrants (inside & outside) (d)
Age (years)	0.012** (0.005)	0.033** (0.015)	0.000 (0.005)	-0.035* (0.019)
Education (years)	0.052 (0.032)	0.044 (0.053)	-0.011 (0.027)	0.042 (0.070)
Gender (dummy, one if male)	0.557*** (0.176)	0.382 (0.256)	0.248* (0.148)	0.945*** (0.330)
Log of household capital for non-farm businesses	0.234*** (0.036)	0.288*** (0.068)	0.189*** (0.039)	0.086 (0.090)
Constant	3.771*** (0.502)	2.291*** (0.782)	5.423*** (0.456)	7.729*** (1.160)
Location dummies	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	-	-
R-squared	0.315	0.391	0.143	0.249
No. of observations	411	134	254	68

Notes: (1) Unit of observations is a person-wave. (2) Sample is panel workers. (3) Figures in () are standard errors. *** denotes significance at 1%, ** at 5%, and * at 10%. (4) Standard errors are robust to heteroskedasticity. (5) Household capital consists of buildings, vehicles like bicycles, tools and so on. (6) In wave 1 to 4, location is classified into 6 categories: Biharamulo, Bukoba Rural, Bukoba Urban, Muleba, Ngara, and Karagwe. In wave 5, the category of outside Kagera is added to these categories.

Table 2.27: Appendix: Transition of Activity Shares by Age in Wave 1

Age in wave 1	Wave 1			Wave 5			Change		
	Mean	Std.	No. of wor- kers	Mean	Std.	No. of wor- kers	Mean	Std.	No. of wor- kers
(1) Non-migrants									
Below 7 years old									
Farm wage	-	-	-	0.02	(0.12)	383	-	-	-
Non-farm wage	-	-	-	0.04	(0.19)	383	-	-	-
Farm self-emp.	-	-	-	0.89	(0.27)	383	-	-	-
Non-farm self-emp.	-	-	-	0.03	(0.17)	383	-	-	-
7 to 20 years old									
Farm wage	0.01	(0.94)	519	0.04	(0.17)	706	0.04	(0.18)	444
Non-farm wage	0.03	(0.17)	519	0.16	(0.34)	706	0.11	(0.35)	444
Farm self-emp.	0.93	(0.21)	519	0.62	(0.42)	706	-0.30	(0.44)	444
Non-farm self-emp.	0.01	(0.10)	519	0.16	(0.33)	706	0.15	(0.33)	444
21 to 30 years old									
Farm wage	0.01	(0.09)	239	0.05	(0.16)	242	0.02	(0.17)	208
Non-farm wage	0.10	(0.28)	239	0.12	(0.30)	242	-0.00	(0.30)	208
Farm self-emp.	0.77	(0.36)	239	0.67	(0.40)	242	-0.11	(0.41)	208
Non-farm self-emp.	0.10	(0.26)	239	0.16	(0.31)	242	0.09	(0.35)	208
31 to 40 years old									
Farm wage	0.03	(0.12)	191	0.05	(0.16)	186	0.02	(0.19)	169
Non-farm wage	0.09	(0.27)	191	0.07	(0.25)	186	-0.02	(0.19)	169
Farm self-emp.	0.79	(0.34)	191	0.77	(0.35)	186	-0.01	(0.36)	169
Non-farm self-emp.	0.07	(0.24)	191	0.10	(0.25)	186	0.01	(0.26)	169
41 years old or above									
Farm wage	0.01	(0.09)	403	0.01	(0.09)	361	0.00	(0.12)	308
Non-farm wage	0.06	(0.21)	403	0.04	(0.19)	361	-0.01	(0.17)	308
Farm self-emp.	0.87	(0.28)	403	0.88	(0.27)	361	0.00	(0.24)	308
Non-farm self-emp.	0.04	(0.17)	403	0.05	(0.19)	361	0.01	(0.19)	308
(2) Migrants (inside)									
Below 7 years old									
Farm wage	-	-	-	0.01	(0.10)	119	-	-	-
Non-farm wage	-	-	-	0.10	(0.31)	119	-	-	-
Farm self-emp.	-	-	-	0.78	(0.40)	119	-	-	-
Non-farm self-emp.	-	-	-	0.09	(0.27)	119	-	-	-
7 to 20 years old									
Farm wage	0.01	(0.08)	379	0.04	(0.17)	465	0.03	(0.17)	297
Non-farm wage	0.00	(0.08)	379	0.16	(0.35)	465	0.13	(0.33)	297
Farm self-emp.	0.96	(0.15)	379	0.60	(0.45)	465	-0.36	(0.46)	297
Non-farm self-emp.	0.00	(0.08)	379	0.19	(0.37)	465	0.20	(0.38)	297
21 to 30 years old									
Farm wage	0.02	(0.12)	63	0.08	(0.25)	69	0.05	(0.30)	56
Non-farm wage	0.08	(0.26)	63	0.15	(0.33)	69	0.06	(0.33)	56
Farm self-emp.	0.80	(0.34)	63	0.56	(0.45)	69	-0.21	(0.43)	56
Non-farm self-emp.	0.08	(0.24)	63	0.18	(0.37)	69	0.09	(0.29)	56
31 to 40 years old									
Farm wage	0.04	(0.17)	36	0.08	(0.23)	33	0.06	(0.21)	30
Non-farm wage	0.26	(0.41)	36	0.23	(0.41)	33	-0.02	(0.22)	30
Farm self-emp.	0.63	(0.41)	36	0.65	(0.43)	33	-0.01	(0.37)	30
Non-farm self-emp.	0.05	(0.16)	36	0.02	(0.08)	33	-0.02	(0.19)	30
41 years old or above									
Farm wage	0.03	(0.12)	31	0.11	(0.31)	30	0.11	(0.29)	19
Non-farm wage	0.17	(0.36)	31	0.02	(0.11)	30	-0.14	(0.34)	19
Farm self-emp.	0.69	(0.42)	31	0.65	(0.43)	30	-0.08	(0.57)	19
Non-farm self-emp.	0.09	(0.28)	31	0.20	(0.38)	30	0.11	(0.29)	19
(3) Migrants (outside)									
Below 7 years old									
Farm wage	-	-	-	0.00	(0.00)	10	-	-	-
Non-farm wage	-	-	-	0.30	(0.48)	10	-	-	-
Farm self-emp.	-	-	-	0.10	(0.31)	10	-	-	-
Non-farm self-emp.	-	-	-	0.60	(0.51)	10	-	-	-
7 to 20 years old									
Farm wage	0.01	(0.12)	120	0.01	(0.12)	119	0.00	(0.16)	79
Non-farm wage	0.00	(0.07)	120	0.44	(0.49)	119	0.39	(0.48)	79
Farm self-emp.	0.96	(0.16)	120	0.11	(0.30)	119	-0.84	(0.33)	79
Non-farm self-emp.	0.01	(0.07)	120	0.42	(0.48)	119	0.44	(0.49)	79
21 to 30 years old									
Farm wage	0.00	(0.00)	20	0.05	(0.22)	19	0.07	(0.27)	13
Non-farm wage	0.32	(0.45)	20	0.46	(0.50)	19	-0.01	(0.57)	13
Farm self-emp.	0.64	(0.44)	20	0.10	(0.31)	19	-0.35	(0.61)	13
Non-farm self-emp.	0.03	(0.10)	20	0.37	(0.48)	19	0.28	(0.50)	13
31 to 40 years old									
Farm wage	0.00	(0.00)	5	0.00	(0.00)	5	0.00	(0.00)	4
Non-farm wage	0.47	(0.50)	5	0.00	(0.00)	5	-0.34	(0.47)	4
Farm self-emp.	0.46	(0.50)	5	0.00	(0.00)	5	-0.58	(0.50)	4
Non-farm self-emp.	0.06	(0.13)	5	1.00	(0.00)	5	0.92	(0.15)	4
41 years old or above									
Farm wage	0.00	(0.00)	14	0.00	(0.00)	3	0.00	(0.00)	3
Non-farm wage	0.06	(0.24)	14	0.33	(0.57)	3	0.02	(0.04)	3
Farm self-emp.	0.87	(0.28)	14	0.66	(0.57)	3	0.04	(0.15)	3
Non-farm self-emp.	0.05	(0.17)	14	0.00	(0.00)	3	-0.07	(0.12)	3

Notes: (1) Individuals are panel respondents who were interviewed in both wave 1 and 5. (2) Activity shares are the proportion of weekly hours worked of each activity relative to the total weekly hours worked.

Table 2.28: Appendix: Correlation of the Transition of Activity Shares with Educational Attainment by Age in Wave 1

Age in wave 1	Activity shares	Non- migrants	Migrants (inside)	Migrants (outside)
Below 7 years old	Farm wage	-0.029	-0.163	-
	Non-farm wage	0.002	-0.097	-0.205
	Farm self	-0.006	0.108	-
	Non-farm self	0.104	0.015	0.205
	No. of workers	380	119	9
7 to 20 years old	Farm wage	-0.046	-0.051	0.000
	Non-farm wage	0.168	0.199	0.349
	Farm self	-0.108	-0.158	-0.184
	Non-farm self	-0.010	0.039	-0.216
	No. of workers	443	297	79
21 to 30 years old	Farm wage	-0.034	0.041	0.076
	Non-farm wage	0.081	-0.009	0.001
	Farm self	-0.019	-0.053	0.359
	Non-farm self	-0.031	0.045	-0.523
	No. of workers	205	55	12
31 to 40 years old	Farm wage	-0.037	0.080	-
	Non-farm wage	0.069	0.329	0.481
	Farm self	-0.013	-0.258	-0.555
	Non-farm self	-0.003	0.013	0.333
	No. of workers	166	29	4
41 years old or above	Farm wage	-0.051	-0.068	-
	Non-farm wage	-0.090	-0.100	0.944
	Farm self	0.103	-0.113	-0.903
	Non-farm self	-0.013	0.409	0.755
	No. of workers	306	19	3
7 years old or above	Farm wage	-0.006	-0.051	-0.053
	Non-farm wage	0.187	0.212	0.398
	Farm self	-0.181	-0.206	-0.311
	Non-farm self	0.065	0.098	-0.125
	No. of workers	1120	400	98

Notes: (1) Individuals are panel respondents who were interviewed in both wave 1 and 5. (2) Activity shares are the proportion of weekly hours worked of each activity relative to the total weekly hours worked.

Chapter 3

Heterogeneous Population and Labour Market Outcome in Urban Africa: Migrants' Earnings in Ghana and Tanzania¹

Abstract

This chapter investigates how rural-to-urban migrants' earnings compare with those of natives in urban labour markets in Africa. The human capital theory of migration suggests that on arrival, migrants' earning potential is lower than that of their native counterparts. However, after their adjustment period at the destination (assimilation), they tend to achieve earnings parity. Exploiting data drawn from urban household panel surveys in Ghana and Tanzania, this chapter attempts to identify the transition process by estimating an earnings function. This chapter makes a novel contribution to the present body of knowledge by providing a distinction between wage-dependent and self-employed migrants. This distinction is important, as migrating to start a business may require higher levels of financial and personal resources, compared to arriving to work for a wage. Thus, the two types of employment-motivated mobility are fundamentally different processes. The analysis uses the variation of migrants' urban experience (measured in years) and earnings across arrival cohorts to identify migrants' earnings growth. However, this identification strategy is problematic, as in the absence of good quality control of individuals, we tend to overestimate (underestimate) the rate of migrants' earnings growth when the quality of recent arrivals is lower (higher) than that of earlier cohorts. To address this problem, this chapter directly controls for the reasoning ability and cognitive skills of individuals by using their standardised test scores. After controlling for individuals' observed attributes and allowing for the differing returns to those attributes across urban residents, we find that the assimilation effect is observed *only* in a wage sector and, more robustly, that recently arrived wage employed migrants earn *less* than recently arrived self-employed migrants. Both findings suggest that wage employed migrants would achieve *higher* lifetime earnings if they entered a self-employed sector at their arrival, conditional on individuals' traits and the varying returns to those traits across urban residents. This is surprising, due to the common belief that wage jobs are preferable to self-employment. Then, why do not all migrants enter a potentially profitable self-employed sector? The possible answer points towards the importance of capital constraints in a decision to start a business.

¹ An earlier version of this chapter was presented at a seminar and the CSAE Conference: Economic Development in Africa (21st-23rd March, 2010) in Oxford. I am grateful to participants for helpful comments.

3.1 Introduction

One traditional view of migration is that it is an investment in human capital because it motivates workers to move from one place to another, where the return to their human capital is higher at the cost of movement (Schultz, 1961; Sjaastad, 1962; Bowles, 1970; Laber and Chase, 1971). The human capital theory suggests that, at the destination, migrants' earnings start at a lower level than that of comparable natives' earnings, but can rise to the natives' level after a certain transition period of adjustment. Based on this view, earlier empirical studies often investigated earnings differential between migrants and natives (Yezer and Thurston, 1976; Grant and Vanderkampt, 1980). The concept of adjustment was subsequently translated into the idea of migrants' assimilation (or localisation) and an empirical framework to analyse the process was made more sophisticated by Chiswick (1978), Borjas's series of works (1985, 1987, 1995, 1996), Borjas et al. (1992), Fiedberg (1992), and LaLonde and Topel (1991, 1997).

The assimilation literature has disproportionately focused on international migration, in particular, to the United States. The examination of earnings differential between immigrants and natives is important because the difference is directly related to income distribution in a host country. If there is a substantial earnings gap between immigrants and natives, immigrants' spatial mobility inevitably increases inequality in a host country, and the inequality, in turn, places greater burdens on social welfare systems (public health insurance, unemployment insurance, social security system, public schooling, housing services and other public resources) and can potentially lead to a political conflict resulting out of undermined social cohesion.

Using data drawn from urban household panel surveys in two developing countries, Ghana and Tanzania, this chapter attempts to identify the process of internal migrants' assimilation by applying the analytical framework to the earnings of rural-to-urban migrants. In contrast to international migration, this study is motivated by the worldwide historical experience, whereby urbanization is a key part of the process of economic development. As reported in Table 3.1, for example, the share of urban population in Asia was just 16.8% in 1950, but it has risen to 39.7% by 2005, in line with the consolidating economic improvement after the World War II. Since rural-to-urban migration plays a major role in the process of urbanisation, the pattern of economic development cannot be fully understood without knowing how the internal mobility is affecting

the structure of urban labour markets in less developed economies. The present study addresses this issue by asking the question of how rural-to-urban migrants' earnings compare with those of natives in urban labour markets in Africa. Focusing on African labour market is essential, because the region recorded the highest annual rate of growth of urban population in the world in the last half of the twentieth century, which makes the findings of this study relevant in a much broader context. As shown in Table 3.1, the rate of urban population growth in Ghana and Tanzania from 2000 to 2005 was more than double the world average as of 2005, although, according to the growth of economies, it seems to be slowing down.

Whilst there is no fundamental difference in the analytical framework between the current study and the previous works reported in existing assimilation literature, shifting the focus onto internal mobility taking place in less developed economies, away from international mobility heading towards more advanced economies (e.g., United States), allows this study to create strikingly novel and unique observations. In contrast to the literature, where wage employed migrants are exclusively surveyed, the present study makes a distinction between wage employed and self-employed migrants. In the countries of interest for this research, small-scale, self-employed activities constitute a large part of urban economy, relative to more developed economies. Starting and running a self-employed business may require a broad range of resources, including physical capital, such as equipment and vehicles, and social capital, such as a business network, as well as personal capabilities, such as self determination, persistence, business acumen etc.. In addition, qualification certificates may also be required to start a business. In comparison, working in a wage-based profession may require few those resources. Hence, it is highly likely that migrating to start a business and migrating to be a wage worker are fundamentally different processes. Thus, making a distinction between the two types of mobility is indispensable to the study of internal mobility taking place in less developed economies.

The chapter uses pooled data drawn from three-round household panel surveys, conducted in urban Ghana and Tanzania in 2003/2004, 2005 and 2006. The surveys collected information on whether labour force participants aged 15 to 60 were born in the cities they resided in at the time of the survey and, if not, when they relocated. The data contained a sufficient number of migrants (about 27% of total working observations in Ghana and about 36% in Tanzania), with their arrivals distributed over the past years with an adequate variation. Of those migrant

workers, approximately 38% and 39% were working in a wage sector in Ghana and Tanzania, respectively. These characteristics were sufficient to identify the process of internal migrants' assimilation, making a distinction between wage workers and self-employed workers.

To identify migrants' earnings growth at the destination, the analysis uses the variation of migrants' duration of stay in cities and their reported earnings. However, this identification strategy is problematic, as in the absence of a good quality control, the rate of migrants' earnings growth can be overestimated (underestimated) when the quality of recent arrivals is lower (higher) than that of earlier cohorts. To address the problem, the analysis directly controls for the reasoning ability and cognitive skills of individuals, in addition to a standard set of earnings determinants, such as age and education. The third-round survey conducted a range of tests on (Reven's) intelligence, language comprehension, English reading skills, and numeracy. The analysis uses scores from these tests to directly control for the unobserved quality of individuals that can significantly impact their earnings.

This study provides an assessment about the validity of the prevailing preconceptions that formal wage jobs are more preferable than informal self-employment and that a self-employment sector is a free-entry sector. By using data sets from the two countries of similar economic level² and by making a distinction between wage employed and self-employed migrants, the findings of this study challenge the unquestionably accepted views, pointing towards the importance of capital constraints in a decision to start a business.

The chapter starts with comprehensively reviewing the assimilation literature in Section 3.2, which provides theoretical explanations about the underlying mechanisms affecting migrants' earnings growth at the destination and carefully presents main econometric challenges in identifying the effect of assimilation. Section 3.3 outlines an framework for an empirical analysis, discussing the present identification strategy. The data overview is presented in Section 3.4, followed by the estimation results, given in Section 3.5, and finally the conclusion in Section 3.6.

²Both Ghana and Tanzania are classified as low-income economies by the World Bank as of 2010 - http://data.worldbank.org/about/country-classifications/country-and-lending-groups#Low_income.

3.2 Review on Assimilation Literature

Extant assimilation literature suggests that whilst migrants have an initial disadvantage in the earning potential, their income steadily rises, as they gradually adjust to working conditions and life at the destination. This raises crucial questions; what mechanisms constitute the assimilation process and what obstacles hinder the empirical identification of its effect? This section briefly reviews the literature related to these questions in order to motivate the development of the subsequent empirical strategy.

3.2.1 Theoretical Explanations

According to the assimilation literature, migrants' *individual* earnings increase as they spend time at the destination. There are two popular explanations for the observed phenomenon: the effect of migrants' learning on their job performance and the general change observed in the return to human capital imported from migrants' origins.

(i) Learning

Whether explicitly or implicitly, the literature has suggested two types of migrants' learning as an underlying mechanism driving their earnings growth at the destination: human capital accumulation and job mobility. Both suggest that migrants' earnings gradually rise as they learn at the destination.

Human Capital Accumulation: Skills Acquisition

In order to survive in destination markets, it is highly likely that migrants need to acquire destination-specific skills, such as language, local customs and ways of conducting business transactions, depending upon jobs that they are aiming for. The concept of this type of informal learning has often been emphasised in the literature of geographical economics. For example, Glaeser (1999) theoretically formalised the role of urban agglomeration in creating informal learning (imitation) opportunities for everyone living in cities and the consequent spillovers of human capital. Closely related is the study by Glaeser and Maré (2001), who called attention to the role of informal learning in explaining wage premium existing in cities. In addition to the geographical economics literature, the concept of learning externality among migrants living in urban areas was introduced

into a similar framework to AK model popular in the growth literature by Lucas Jr (2004), who provided an explanation for gradual transition from a traditional agricultural economy to a modern industrialised one. One distinctive feature of this type of learning is that migrants accumulate human capital (skills) at the destination, conditional on their ability. However, the same is not true in relation to job mobility described next.

Job Mobility: Information Collection

During the time in the destination, conditional on human capital imported from their origins, migrants acquire locale-specific knowledge, such as the types of jobs available, types of available formal and informal institutions, the condition of local job markets, and even whether they are capable and adaptable enough to earn living and thus survive. This learning process eventually decreases, or even eliminates, the disparity between their knowledge and that of comparable local workforce, allowing them to find and retain well-paid jobs. This process generates the growth of migrants' earnings even in the absence of human capital (skills) accumulation. It seems that the human capital literature of migration studies largely relies on this learning mechanism for explaining migrants' earnings profile at the destination.

Similar to the literature on job turnover (e.g., Jovanovic, 1979; Miller, 1984), a robust and comprehensive theoretical model of the process of migrants' job mobility should adequately address the imperfect information on the quality of match between skills and jobs attained, which stems from uncertainty about both ability (or type) of workers and jobs available (or skills demanded). Both parties of the firms (if wage jobs) and workers potentially contribute to such imperfection. Based on the literature review, Yezer and Thurston (1976) seem to be the first to introduce the aspect of job search to the traditional human capital model of migration to account for the empirically observed positive correlation between migrants' earnings and their length of stay in the destination. Their model is based upon imperfect information on jobs available that tends to affect migrant workers' job search. Whilst the theory of job search conventionally assumes that workers know the distribution of wages but not the actual wage for any given job, Yezer and Thurston (1976) are considering the case where the distribution is also imperfectly known by workers.

Another theoretical mechanism was presented by Yamauchi (2004) to provide an explanation for migrant workers' success in destination labour markets. He focused on workers' type, not known

by both workers themselves and firms *ex ante*. In his model, based on a standard Bayesian learning framework, the quality of match is revealed by forming and experiencing it.³ Assuming that the gap between each task (within a job) and workers' skills is instantaneously adjusted by both firms (e.g., the reallocation of workers, the introduction of new tasks) and workers (e.g., the turnover of jobs) without any friction, he showed that, as both migrants and firms learned about migrants' abilities (types), migrants' earnings gradually increased, resulting in improving job coordination in destination labour markets.

This concept of job mobility has recently gained empirical attention. For example, an empirical model of on-the-job search was developed by Weiss et al. (2003) to account for the experience of highly skilled immigrants from the former Soviet Union to Israeli labour market, incorporating the structural parameters in the economy as estimated through a technique of dynamic programming. Their simulation based on the estimated parameters successfully explained the occupational mobility of highly skilled immigrants and the growth of their earnings, with a conclusion that their lifetime earnings, nevertheless, fell short of those of comparable natives by 57%.

Finally, migrants might accumulate several types of destination-specific capital, such as urban social networks including friends and colleagues. If this is the case, it is certainly possible that the effect of both types of learning explained above is augmented by such capital accumulation.

(ii) Change in Return to Human Capital Imported from Origins

The quality of human capital (work experience, education, professional qualifications) gained in migrants' place of origin might be difficult to evaluate for employers in the destination labour markets and, as a result, migrants' skills may not be easily transferrable to the destination workplaces. In such case, migrants' skills might initially be valued unfavourably, compared to the equivalent gained at the destination. However, as the quality of the imported human capital can be evaluated through a longer term employment, it is crucial to allow sufficient time for its return to reach that to the equivalent obtained at the destination. This is another theoretical explanation for the growth of migrants' earnings at the destination, suggesting that the increase in migrants' earnings arises from the unveiled quality of their imported skills, not their human capital (skills)

³In this sense, jobs in his model are experience goods, based on the terminology of job turnover literature (Jovanovic, 1979).

accumulation.

This explanation is often supported by empirical evidence, focusing on the immigration of Jews from the former Soviet Union to Israel, which started toward the end of 1989. For example, Friedberg (2000) drew upon Israeli census data in 1983 to show that human capital (education and labour market experience) obtained abroad was less valued than the equivalent obtained domestically, and that the difference was sufficient to fully explain the earnings disadvantage of immigrants. Eckstein and Weiss (2004) conducted the structural estimation to show that about 40% of an average wage increase of Jewish male immigrants from 1991 to 2000 could be explained by an increase in the return to their imported skills, with greater importance than any other investigated sources including immigrants' occupational mobility. Closely related to it, the simulated results of Weiss et al. (2003) concluded that approximately 75% of the gap in lifetime earnings between highly skilled immigrants from the former Soviet Union and otherwise comparable Israeli natives resulted from a lower market valuation of immigrants' imported skills, with the remaining 25% arising from frictions associated with nonemployment and job distribution mismatch.⁴

The immigration experience of Jews from the former Soviet Union examined in these papers can be characterised by the following three aspects: (a) a huge influx of immigrants with high level of education acquired in their country of origin and pre-immigration work experience in high skilled jobs;⁵ (b) their earnings growth, much of which can be accounted for by the increase in the return to their imported skills; and (c) their occupational upgrade (e.g., from unemployment to employment, or from low-income jobs to high-income jobs). Taking these factors into consideration, a simple model is developed in this sub-section in order to facilitate the understanding of the relationship between the growth of immigrants' earnings and their length of stay at the destination. In the model, firms must assess the level of immigrant workers' skills through the observation of their performance during the period of employment, as the quality of immigrants' education obtained in their place of origin is difficult to evaluate and, hence, does not work as a good signal of their skills.

We consider a wage profile determined by one firm which is permanently hiring one immigrant

⁴This implies that immigrants could have earned substantially higher lifetime income had their imported skills been valued in the same way as those of otherwise comparable natives from the time of their arrival.

⁵However, this does not have much impact on the employment and wages of natives because of the inflow of capital and increased exports (Weiss et al., 2003).

once drawn from the pool of many immigrants from the initial (probation) period $t = 0$.⁶ There are two types of immigrants in the economy: good, with educational level s_H and skills level a_G , and bad, with s_L and a_B ($s_H > s_L$ and $a_G > a_B$), where the level of education - although acquired in the country of origin - can easily be observed by the firm, but the level of skills cannot. The proportion of good immigrants is $\gamma > 0$, which is known by the firm, with common knowledge about the level of a_G and a_B . Rationally, it is assumed that the best wage contract for the firm is to equate wage with the expected skills of immigrants. Thus, to the extent that the level of immigrants' education is perfectly informative about the level of their skills, the firm chooses $w = a_G$ for immigrants with s_H and $w = a_B$ for them with s_L during all the time of employment.

However, it is possible that the quality of education obtained abroad is different from equivalent in a country of destination. If this is the case, the firm is unable to evaluate the skills of immigrants just by their educational level. Thus, given that migrants' qualifications are deemed to be unreliable aptitude predictors, the assumption needs to be made that the firm can adequately assess the aptitude of migrant workers only by observing their performance which is subject to some random shocks. If an employed worker is good, he/she produces y_H with some known probability $p_G > 0$, and y_L with the remaining probability; whereas a bad worker produces y_H with known probability $p_B > 0$, and y_L with the remaining probability ($y_H > y_L$ and $y_H - a_G > y_L - a_B > 0$). The level of y_H and y_L is observable by the firm. Not surprisingly, it is assumed that p_B is substantially low and that $p_G > p_B$. In each period $t \geq 1$ ($t = 1, 2, 3, \dots$), the firm chooses the wage so that

$$w_t = E[a|h_t], \quad (3.1)$$

where h_t is the history of the past performance of the employed worker $h_t = \{y_0, y_1, \dots, y_{t-1}\}$ with $h_0 = \emptyset$. Assuming that the order of high or low performance periods in the past does not generate additional information to estimate the skills of the worker when the firm chooses his or her wage for the coming period, the history can be described with the number of high performances in the

⁶Discussions here are largely motivated by Professor Marcel Fafchamps' lecture materials for the second year master students in economics at the University of Oxford in a 2007-2008 academic year, which are treating the enforcement mechanism of the contract between suppliers and buyers in a commodity market.

past, N . If the actual number is n at time t , the wage is determined so that it could be

$$w_t = E[a|N_t = n] = a_G \cdot \text{Prob}(a_G|N_t = n) + a_B \cdot \text{Prob}(a_B|N_t = n). \quad (3.2)$$

Applying Bayes's rule to (3.2), we can obtain (see Appendix 3.7.1 for the derivation)

$$w_t = \frac{a_G}{1 + \hat{\gamma} \cdot p_0^n \cdot p_1^{t-n}} + \frac{a_B \cdot \hat{\gamma} \cdot p_0^n \cdot p_1^{t-n}}{1 + \hat{\gamma} \cdot p_0^n \cdot p_1^{t-n}}, \quad (3.3)$$

where $\hat{\gamma} \equiv \frac{1-\gamma}{\gamma}$, $p_0 \equiv \frac{p_B}{p_G} < 1$ and $p_1 \equiv \frac{1-p_B}{1-p_G} > 1$. Under the assumption that the immigrant who is hired by the firm is of a good type, it is highly likely that the proportion of his/her achievements of high performance in the past increases as time passes, i.e., as $t \rightarrow \infty$, $\frac{n}{t} \rightarrow 1$. It follows that

$$\lim_{t \rightarrow \infty} w_t = a_G. \quad (3.4)$$

Thus, starting from the initial wage, $w_0 = \gamma \cdot a_G + (1 - \gamma) \cdot a_B$, the earnings of the skilled immigrant will eventually rise to a_G , as the length of employment increases. Furthermore, the mass immigration of Jews with high level of education and skills acquired in the former Soviet Union into Israel can be captured by large γ . This generates the faster growth of skilled immigrants' earnings with the relatively high level of initial wage. Suppose that natives with the similar level of skills are paid a_G , it is likely that immigrants' earnings converge to that level at high rate of increase.

Another feature of the immigration experience of Jews is immigrants' occupational upgrading. In Appendix 3.7.2, we extend the model to show that a huge influx of skilled immigrants (large γ) generates their occupational upgrading from unemployment to employment by providing the firm with an incentive to replace an initially hired worker with a new one with better skills through the mechanism of search.

3.2.2 Empirical Model

Much work to identify the assimilation effect has been done by Chiswick (1978), Borjas (1985, 1987, 1995, 1996), Borjas et al. (1992), Fiedberg (1992), and LaLonde and Topel (1991, 1997).

Using cross-sectional data, the literature often attempts to estimate

$$\log w_{it}^m = \alpha_1^m + \left[\alpha_2 g_i + \alpha_3 c_{it} + \alpha_4 c_{it}^2 + \sum_d \alpha_{5d} D_{id} \right] + \alpha_6^m a_{it} + \alpha_7^m \mathbf{z}_i + \delta_t^m + \epsilon_{it}, \quad (3.5)$$

$$\log w_{it}^n = \alpha_1^n + \alpha_6^n a_{it} + \alpha_7^n \mathbf{z}_i + \delta_t^n + \epsilon_{it}, \quad (3.6)$$

where w_{it}^l , $l \in (m, n)$, is the earnings of a migrant (m) or a native (n) worker i at time t , a_{it} is an individual's age, a time-specific common shock (in the case of estimating pooled data of repeated cross-section) is captured by δ_t^l , and ϵ_{it} is a random error term with zero mean. A vector $\mathbf{z}_i$ contains other determinants of earnings, such as the level of education and dummy variables for locations; however, for ease of exposition it is assumed to be time invariant. In the absence of terms in the square brackets for migrants, (3.5) and (3.6) are a simplified version of Mincerian earnings function. Noticeably, explanatory variables are valued differently across migrants and natives. The terms in the square brackets capture migration effects, and variables g_i and c_{it} reflect age at migration and migrants' cumulative destination experience at time t since their arrival (years), respectively. A variable D_{id} is a dummy for arrival cohorts (i.e., equal to one if the migrant arrived in a period d), which controls for each cohort's average quality that might differ across cohorts. In the actual estimation, migrants and natives are pooled as

$$\log w_{it} = M_i \log w_{it}^m + (1 - M_i) \log w_{it}^n, \quad (3.7)$$

where M_i is a dummy, equal to one for a migrant and zero otherwise. Suppose that migrants initially obtain low-paid jobs and that their earnings gradually rise to the level of similar natives' earnings, it is expected that a coefficient on M_i in (3.7) (i.e., $\alpha_1^m - \alpha_1^n$) is negative, $\alpha_3 > 0$, and $\alpha_4 < 0$ (concave earnings profile).

However, not all the parameters of interest can be identified in (3.7). Firstly, years since migration (c_{it}) and cohort dummies (D_{id}) are highly correlated, and due to their collinearity, either one is redundant and must be removed.⁷ In the subsequent analysis, both specifications are delineated. Secondly, it is difficult to identify a work experience effect, often measured by

⁷Theoretically speaking, it is still possible to identify the effects of years since migration, age at migration, and arrival cohorts simultaneously, as long as we assume that period effects and work experience effect are the same across groups, i.e., $\delta_t^m = \delta_t^n$ and $\alpha_6^m = \alpha_6^n$. This is a strategy taken by Borjas (1996). However, it still seems doubtful that the effects of years since migration and arrival cohorts can separately be identified in the actual estimation because they are highly correlated.

age, separately between migrants and natives. In particular, it is theoretically impossible to do so in a specification that includes years since migration, because $(a_{it} - c_{it}) \cdot M_i = g_i \cdot M_i$ (perfect multicollinearity). Even in a specification that retains cohort effects, it seems doubtful that the model can allow the return to work experience to vary across groups, as $a_{it}M_i$, g_iM_i , and $D_{id}M_i$ are still highly correlated in (3.7). Following the recommendations from the previous studies (e.g., Borjas, 1996), the subsequent analysis, thus, imposes an assumption that $\alpha_6^m = \alpha_6^n$.

3.2.3 Econometric Concerns

When cross-sectional data are used to estimate (3.7), great empirical difficulty arises in identifying the assimilation effect. Several main econometric concerns are explained in this sub-section.

(i) Difference in Cohort Quality

Using cross-sectional data essentially implies that the assimilation effect is estimated through a comparison of earnings across arrival cohorts. Given the likelihood that the average quality of observed migrants varies with arrival cohorts, this generates a fundamental identification problem in the absence of a good quality control. Figure 3.1, which is a modified version of Figure 2 in LaLonde and Topel (1997, p. 831), provides an intuitive representation of the issue. Let us assume hypothetical cross-sectional data in 2010 comprising two types of migrants: those who arrived in 1970 and in 2005. The vertical axis of the figure reflects the average quality (hence, earnings) of each cohort and the horizontal axis is years since migration. For simplicity, we assume that migrants do not assimilate and, hence, that each migrant earns a constant level of earnings during their stay at the destination (w_{05} for 2005 arrival and w_{70} for 1970 arrivals). The figure shows the case where 2005 arrivals always earn less than 1970 arrivals because their average quality is lower than 1970 arrivals.

There is no assimilation effect in this hypothetical data. However, since the migrants are observed at points A and B in 2010 data, their assimilation rate is estimated as a slope of line A-B in the absence of a good control for the quality difference across cohorts. As explained, destination experience variable (years) and cohort dummies cannot simultaneously be controlled for, due to multicollinearity in the actual estimation. Thus, it is likely that the rate of assimilation will be overestimated (underestimated) when the quality of recent arrivals declines (rises) over time. The

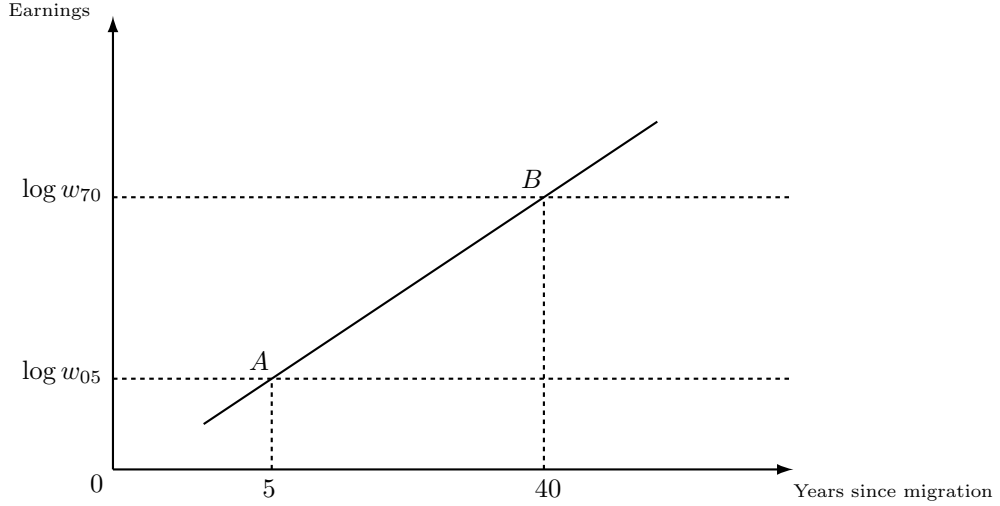


Figure 3.1: The Estimate of Assimilation Effect Based on Hypothetical Cross-Sectional Data in 2010

following sub-sections present several mechanisms that could produce this effect.

Sample Selection by Age at Migration

The first potential mechanism is noted by Friedberg (1992). In cross-sectional data, negative correlation between age at migration and years since migration is normally observed, as those who arrived later in their life and have been in the destination for many years are no longer in employment (e.g., 1970 arrivals who migrated at age 40 in 2010 data). Similarly, it is certainly true that child migrants who arrived a few years ago with their parents have not yet entered the labour market (e.g., 2005 arrivals who migrated at age 5 in 2010 data). Thus, the currently observed earlier cohorts are more likely to be dominated by those who arrived in their childhood, whereas the currently observed recent arrivals are more likely to be dominated by those who arrived in their adulthood. As migrants who arrive at the destination at young age would have the same social and cultural experience, such as education, as natives, they may be qualitatively similar to natives. Given the postulate that natives are more skilled than typical migrants, the selected sample of migrants in each cohort will result in overestimating the true assimilation effect by decreasing the relative quality of recent arrivals, as a large number of those who arrived previously would, in fact, be equivalent to natives. As already done in (3.7), a solution to this problem is to explicitly control for age at migration which could capture the quality difference across cohorts attributed to this sample selection problem. By using the US Census data, Friedberg (1992) showed that including age at migration into the model reduced both the cross-sectional difference in earnings

between early and recent arrivals (by about 25%) and the rate of migrants' assimilation.

Sample Selection by Return Migration

Within each new arrival cohort, it is certainly possible that a proportion of initial arrivals return home in a non-random way. Assuming a competitive destination labour market, it is likely that the least skilled group of initial arrivals would move back to their place of origin after unsuccessful attempts to assimilate into the new culture and workplace. Thus, within each cohort, migrants who remain in the destination will be a selected group, with higher abilities than the average for the cohort that initially migrated. As the selection mechanism associated with labour market competition is more evident for earlier arrival cohorts (because there will be no time for the mechanism to effectively operate for migrants who arrived recently), it is likely that the average quality of migrants observed in cross-sectional data is higher for earlier cohorts. This implies that migrants' initial disadvantage and earnings growth will be overestimated without controlling for the quality difference across cohorts.

Shift in the Origin Mix of Migrant Flows

The quality difference across cohorts in cross-sectional data could also arise from a dramatic change in the pattern of migrants' original regions. In the United States, for example, during the post-war periods there has been a substantial shift in the national origins of immigrant flows, away from the traditional European countries towards Asian and Latin American countries. Several studies found that the shift resulted in a sequential decline in the quality of new arrivals (and, implicitly, earnings) to the United States, generating an upwards biased cross-sectional estimate of assimilation effect (LaLonde and Topel, 1991; Borjas, 1995).

However, for the above issue to be relevant for the current study of internal mobility, the following two conditions must hold: firstly, the quality of workforce must systematically vary across regions within a country; and secondly, there must be a systematic shift in the origins of migrant flows away from high quality labourer producing areas to low quality ones. Since it is expected that the average quality of individuals is relatively uniform across villages within a rural area, this type of quality difference across cohorts is probably less important in the subsequent interpretation of the findings.

Of course, the sequential decline in the quality of recent arrivals is possible even in the absence of the shift in the geographical origins of migrant flows, if the average quality of people (including migrants) living in a particular area, from which migrants are exclusively supplied, has been continuously declining over time. However, given the recent economic development in the countries chosen for the present study, this phenomenon is unlikely.

(ii) Sample Selection into Migration and Employment

So far, a disproportionally large emphasis was placed on the quality difference among migrants. However, one fundamental selection problem in estimating (3.7) is that migrants might be different from natives in some unobservable way, influenced by the fact that they are already pre-selected from their origins (as only individuals with determination to succeed and accept a certain risk of unknown would even attempt a venture of moving from their natural surroundings). Hence, as not all characteristics of migrants are readily observable, their fair and accurate comparison with the native population might be difficult. When migrants are on average more (less) able than natives, for example, the initial disadvantage of migrants will be underestimated (overestimated) in the absence of a good control for the quality difference between groups.

Another empirical challenge is a standard sample selection problem in relation to employment. To the extent that those who earn positive income are selected from population in a non-random way, estimating (3.7) will make the estimated coefficients spurious without controlling for the quality of selected workers. In fact, this selection problem is much harder to address if a distinction between wage-dependent and self-employed workers is required, as it needs to take into account not just selection into employment, but also selection into wage or self-employment.

(iii) Recent Arrivals Artificially Created by Repeated Migration

Whilst the literature review has not identified any studies that have carefully addressed this issue, a systematic measurement error about migrants' true destination experience could generate the pseudo assimilation effect in cross sectional data. Due to the lack of reliable information, the previously conducted studies measured migrants' destination experience as the length of time since their most recent arrival year, which did not take into account whether or not they were repeated migrants. This procedure is, thus, likely to underestimate the true destination experience

of repeated migrants if they are treated as recent arrivals. In competitive destination labour markets, it is likely that the more able migrants would continuously stay in the destination once they arrive there. Based on this view, the repeated migrants are implicitly of relatively low quality. Thus, failing to measure their true destination experience artificially generates positive correlation between migrants' destination experience and their earnings in cross-sectional data, which at first glance appears to be the assimilation effect. This idea is theoretically formalized in Appendix 3.8.

As in the case of difference in cohort quality, explicitly controlling for the quality of migrants can partly mitigate this problem. Figure 3.2 illustrates this point effectively. This hypothetical cross-sectional data includes two types of migrants who have 20-year long destination experience: lifetime migrants who have continuously lived in the destination since their first arrival and repeated migrants whose recent arrival year was just five years ago. As before, for expositional simplicity, the assumption that there is no assimilation effect is made; thus, migrants' earnings depend only upon their quality. In the figure, lifetime migrants are more talented and consistently earn higher income w_L at the destination than repeated migrants, earning w_R .

To the extent that the destination experience of repeated migrants is correctly measured, the two types of migrants should be observed at points B and C and there is no positive correlation between migrants' destination experience and their earnings.⁸ Thus, earnings differential between the two groups arises solely from the difference in their intrinsic quality. However, using this cross-sectional data identifies the pseudo assimilation effect as a slope of line A-B, which arises when the destination experience of repeated migrants is not adequately measured. This problem can be overcome if the quality difference between migrants can perfectly be controlled. In such a case, the two types of migrants can be represented at the points that reveal the same level of earnings, but still differ in the measured destination experience, e.g., points B and D. Clearly the pseudo assimilation effect disappears in this data set.

The discussion suggests that controlling for the quality of each migrant can partly solve the problem by removing from an earnings function the correlation between the measurement error related to the destination experience of repeated migrants and the quality of each migrant. However, this does not imply that the quality control can manage all biases arising from the systematic

⁸Strictly speaking, since there is no variation of true destination experience in this hypothetical data set, the assimilation effect, albeit present, cannot be identified. However, this is a separate issue, and is out of scope of the current discussion.

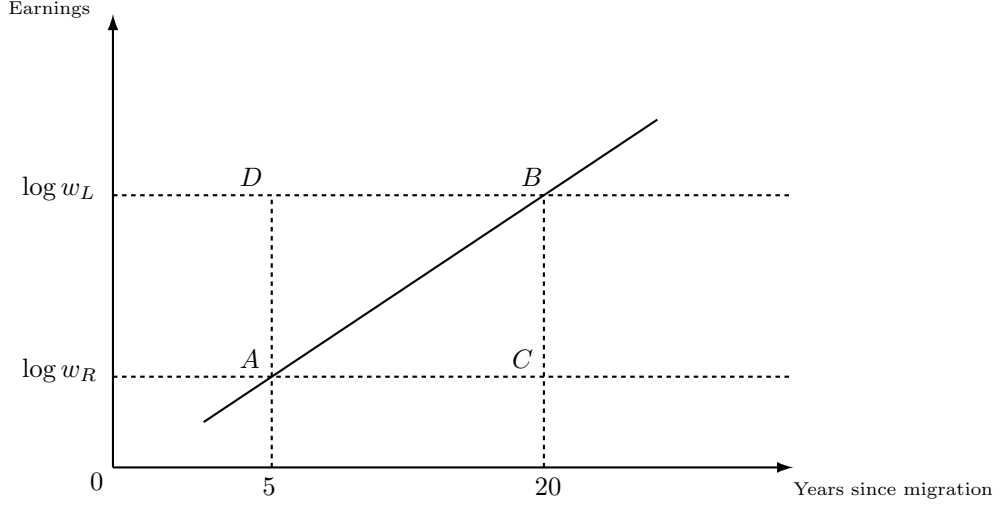


Figure 3.2: Artificially Created Recent Arrivals

measurement error. In other words, a standard measurement error bias remains, unless the true assimilation effect is zero.⁹ This is a remaining concern even in the presence of a good quality control.

3.3 Empirical Strategy

The empirical goal of this chapter is to identify the process of migrants' assimilation, making a distinction between wage employed migrants and self-employed migrants. Following discussions in Section 3.2, this section outlines the empirical strategy employed to achieve the goal.

3.3.1 Specification

The analysis uses panel data drawn from surveys conducted in three waves, as though they were repeated cross-sectional data. This is followed by modelling the monthly earnings of a worker i at time t (w_{it}) as

$$\log w_{it} = k_i \log w_{it}^k + h_i \log w_{it}^h, \quad (3.8)$$

⁹To see this, we can consider the following simple model based upon data containing migrants only: $\log w_i = \alpha c_i^* + \eta_i + \epsilon_i$, where c_i^* is true destination experience, $c_i = c_i^* + u_i$, and u_i is a systematic measurement error, such that $E(u_i) = E(c_i^* u_i) = E(\epsilon_i u_i) = 0$. The quality of individuals η_i is correlated with the measurement error, thus $E(\eta_i u_i) \neq 0$ (e.g., the true destination experience of repeated migrants of low quality is not adequately measured). It is also assumed that $E(\eta_i) = E(c_i^*) = E(c_i^* \epsilon_i) = E(c_i^* \eta_i) = 0$. Then, when η_i is unobserved, a standard algebra yields

$$\text{plim}_{n \rightarrow \infty} \hat{\alpha} = \alpha \left(\frac{\sigma_{c^*}^2}{\sigma_{c^*}^2 + \sigma_u^2} \right) + \left(\frac{\rho_{\eta u}}{\sigma_{c^*}^2 + \sigma_u^2} \right),$$

where $\sigma_u^2 \equiv \text{Var}(u_i)$, $\sigma_{c^*}^2 \equiv \text{Var}(c_i^*)$, and $\rho_{\eta u} \equiv \text{corr}(\eta_i, u_i)$. Controlling for the quality of individuals η_i can address the second term of bias, but the first term of standard measurement error bias remains unless $\alpha = 0$.

where k_i and h_i are dummies for wage-dependent and self-employed workers, respectively (i.e., $h_i = 1 - k_i$). Following (3.7), the earnings of a worker i in a sector $j \in (k, h)$ at time t are, in turn, modelled as

$$\log w_{it}^j = \beta_0^j + \left(\beta_1^j + \beta_2^j g_i + \beta_3^j c_{it} + \beta_4^j c_{it}^2 \right) \cdot M_i + \beta_5^{j'} \mathbf{x}_i + \beta_6^{j'} (\mathbf{x}_i - E(\mathbf{x}_i)) \cdot M_i + \delta_t^j + \delta_t^j \cdot M_i + \epsilon_{it}, \quad (3.9)$$

where most earnings determinants are captured by the vector $\mathbf{x}_i$. In practice, $E(\mathbf{x}_i)$, the population average of $\mathbf{x}_i$, is replaced with the sample mean.¹⁰ In this specification, negative β_1^j implies that a migrant worker in a sector j is at an initial earnings disadvantage relative to otherwise comparable natives in the same sector. Positive β_3^j and negative β_4^j yield the concave earnings profile (assimilation effect) of the migrant worker in a sector j . Section 3.2.3 (i) already delineated that including g_i was crucial if observed migrant workers were a sample selected by age on arrival. While our primary purpose of including g_i is to control for the selection-oriented quality difference across cohorts, the impact could have its own meaning. As found in Friedberg (1992), to the extent that migrants who arrived in their youth are more adept at adapting to new life at the destination than those who arrived in their adulthood, negative β_2^j is expected. On the other hand, if (labour market) experience in rural areas can easily be transferred to urban labour markets, it is not surprising to find positive β_2^j . A vector β_6^j captures the difference in the returns to $\mathbf{x}_i$ between migrants and natives in a sector j . As explained in Section 3.2.2, however, the specification does not allow the return to work experience (age) to vary across groups due to perfect multicollinearity. The period effect is allowed to affect migrants and natives differently within the sector.

Taking (3.8) and (3.9) together, we obtain

$$\begin{aligned} \log w_{it} = \beta_0^h + & \sum_{j \in (k, h)} \left[\left(\beta_1^j + \beta_2^j g_i + \beta_3^j c_{it} + \beta_4^j c_{it}^2 \right) \cdot M_i \cdot j_i + \beta_5^{j'} \mathbf{x}_i \cdot j_i + \beta_6^{j'} (\mathbf{x}_i - E(\mathbf{x}_i)) \cdot M_i \cdot j_i + \delta_t^j \cdot j_i + \delta_t^j \cdot M_i \cdot j_i \right] \\ & + (\beta_0^k - \beta_0^h) k_i + \epsilon_{it}. \end{aligned} \quad (3.10)$$

Equation (3.10) is the most general version of our specification and it allows the returns to $\mathbf{x}_i$ to vary across different types of workers (wage or self-employed, migrants or natives). Conditional on the heterogeneous returns, a coefficient on a wage sector dummy (i.e., $\beta_0^k - \beta_0^h$) captures the

¹⁰Subtracting the sample average, rather than population average, introduces a generated regressor problem. However, adjusting the standard errors typically has minor effects (Wooldridge, 2002, p. 613).

difference in the average return between sectors.

Imposing an additional assumption that there is no difference in the returns to $\mathbf{x}_i$ and the period effect between migrants and natives, but still allowing for the difference between sectors, equation (3.10) can be reduced to

$$\log w_{it} = \beta_0^h + \sum_{j \in (k, h)} \left[\left(\beta_1^j + \beta_2^j g_i + \beta_3^j c_{it} + \beta_4^j c_{it}^2 \right) \cdot M_i \cdot j_i + \beta_5^{j'} \mathbf{x}_i \cdot j_i + \delta_t^j \cdot j_i \right] + (\beta_0^k - \beta_0^h) k_i + \epsilon_{it}. \quad (3.11)$$

In the case of no difference in the returns to $\mathbf{x}_i$ and the period effect across all types of workers, the model becomes

$$\log w_{it} = \beta_0^h + \sum_{j \in (k, h)} \left[\left(\beta_1^j + \beta_2^j g_i + \beta_3^j c_{it} + \beta_4^j c_{it}^2 \right) \cdot M_i \cdot j_i \right] + \beta_5' \mathbf{x}_i + (\beta_0^k - \beta_0^h) k_i + \delta_t + \epsilon_{it}. \quad (3.12)$$

In Section 3.2.2, it was shown that the effects of destination experience and cohort dummies could not simultaneously be identified due to collinearity and, thus, that either one must be eliminated in estimating (3.7). In contrast to all the above specifications keeping years since migration, retaining a cohort dummy (in the simplest way) can model the earnings of a worker i in a sector j as

$$\log w_{it}^j = \gamma_0^j + \gamma_1^j M_{it}^R + \gamma_2^j g_i \cdot M_i + \gamma_3^{j'} \mathbf{x}_i + \gamma_4^{j'} (\mathbf{x}_i - E(\mathbf{x}_i)) \cdot M_{it}^R + \delta_t^j + \delta_t^j \cdot M_{it}^R + \epsilon_{it}, \quad (3.13)$$

where M_{it}^R is a dummy equal to one for a recent migrant at time t . In our short-period panel, this variable is treated as time-invariant ($M_{it}^R = M_i^R$), equal to one for the migrant who arrived between 1992 and 2006. Suppose that a migrant i in a sector j , who arrived within the recent 15 years, earns less than the others (earlier arrivals and natives) in that sector, negative γ_1^j is expected. As before, the difference in the returns to $\mathbf{x}_i$ between recent arrivals and the others in a sector j is captured by the vector γ_4^j .

Combining (3.8) and (3.13) yields

$$\log w_{it} = \gamma_0^h + \sum_{j \in (k, h)} \left[\gamma_1^j M_i^R \cdot j_i + \gamma_2^j g_i \cdot M_i \cdot j_i + \gamma_3^{j'} \mathbf{x}_i \cdot j_i + \gamma_4^{j'} (\mathbf{x}_i - E(\mathbf{x}_i)) \cdot M_i^R \cdot j_i + \delta_t^j \cdot j_i + \delta_t^j \cdot M_i^R \cdot j_i \right] + (\gamma_0^k - \gamma_0^h) k_i + \epsilon_{it}, \quad (3.14)$$

This specification, once again, allows the returns to $\mathbf{x}_i$ to vary across different types of workers (wage or self-employed, recent arrivals or the others). Based on the heterogeneous returns, the

difference in the average return between sectors is captured by a coefficient on a wage sector dummy (i.e., $\gamma_0^k - \gamma_0^h$).

Assuming that the returns to $\mathbf{x}_i$ and the period effect differ only across sectors, (3.14) can be simplified into

$$\log w_{it} = \gamma_0^h + \sum_{j \in (k, h)} \left[\gamma_1^j M_i^R \cdot j_i + \gamma_2^j g_i \cdot M_i \cdot j_i + \gamma_3^{j'} \mathbf{x}_i \cdot j_i + \delta_t^j \cdot j_i \right] + (\gamma_0^k - \gamma_0^h) k_i + \epsilon_{it}. \quad (3.15)$$

Imposing an assumption that all types of workers have the same returns to $\mathbf{x}_i$ and the period effect generates

$$\log w_{it} = \gamma_0^h + \sum_{j \in (k, h)} \left[\gamma_1^j M_i^R \cdot j_i + \gamma_2^j g_i \cdot M_i \cdot j_i \right] + \gamma_3' \mathbf{x}_i + (\gamma_0^k - \gamma_0^h) k_i + \delta_t + \epsilon_{it}. \quad (3.16)$$

In the subsequent analysis, a range of these specifications is estimated: (3.10), (3.11), (3.12), (3.14), (3.15) and (3.16).

In the above specifications, the vector $\mathbf{x}_i$ is of natural interest. In addition to a standard set of earnings determinants, such as age (including its squared term), education (including its squared term) and a gender dummy, ethnic group dummies are also included.¹¹ This has three important implications. Firstly, the ethnic group dummies can control for potential ethnic fractionalisation in a labour market (Barr and Oduro, 2002). Secondly, they help to identify the process of migrants' localisation, conditional on an ethnic network. The literature often places an emphasis on the role of social networks based upon original communities in helping migrants find jobs or adapt to new settings (Andersson, 2001; Munshi, 2003; McKenzie and Rapoport, 2007; Yamauchi and Tanabe, 2008). By putting in the ethnic group dummies, the analysis can ask if any systematic profile of migrants' earnings at the destination can be found and identified, conditional on the strength of each ethnic network. Thirdly, including ethnic group dummies allows for comparison of a representative migrant with a representative native by removing the influence of ethnicity. On the other hand, it would be undesirable to control for any characteristics of individual jobs by $\mathbf{x}_i$ (e.g., tenure, the size of businesses, etc.), as one explanation for migrants' assimilation process is

¹¹ Ethnic group is classified into nine categories in Ghana (Akan, Ewe, Ga-Dangme, Guan, Grusi, Gurma, Mande, Mole-Dagbani, Other) and twenty-three categories in Tanzania (Bondei, Chagga, Fipa, Gogo, Makonde, Masai, Ngoni, Nyakyusa, Nyamwezi, Nyasa, Ha, Haya, Jita, Kuria, Lyo, Pare, Rangi, Sambia, Sukuma, Wa-Arusha, Zaramo, Zigua, Other).

their search for well-paid jobs (Borjas et al., 1992). It is worth noting that, as usual, the estimation controls for location fixed effects.

3.3.2 Identification

Since the present study applies the same analytical framework as the assimilation literature, all identification issues discussed in Section 3.2.3 are of great concern, apart from the selection problem associated with age at migration, which can directly be controlled for. Similarly, the difference in cohort quality driven by the shift in the geographical origins of migrant flows is not a serious concern, for the reason explained in Section 3.2.3 (i). However, the remaining issues are great econometric challenges.

One possible way to address sample selection problems (non-random selection into migration, return migration, and wage/self employment) is to use a selection correction methodology by explicitly modelling each selection process. However, the strategy often yields evidence that is sensitive to a parametric assumption made in each process and needs good instruments that explain the selection but not earnings. Due to the complexity of selection mechanisms and the lack of good instruments in data, this is not the strategy taken in this study.

Apart from a standard measurement error bias associated with the destination experience of repeated migrants, the key to addressing all the other econometric issues is the extent to which the quality of observed individuals can be controlled for. Following Boissiere et al. (1985), the current study directly exploits variables measuring the reasoning ability and cognitive skills of individuals. Using wage labour force data in Kenya and Tanzania, Boissiere et al. (1985) attempted to distinguish the influence on earnings of native ability (Raven's test), cognitive skills (literacy and numeracy tests), and education, and found that the reasoning ability and education operated through the development of cognitive skills in determining earnings.¹²

The third-round survey included tests designed to measure the reasoning ability and cognitive skills of respondents. A multiple-choice format was used in most questions. As in Boissiere et al. (1985), the ability was measured by Raven's intelligence test, which is expected to reveal the ability

¹²The authors used the term "expanded human capital model" to describe the earnings function, including native ability, cognitive skills, and education. In the model, they found in both countries that controlling for cognitive skills substantially reduced the return to education of secondary school-leavers relative to primary school-leavers and that conditional on cognitive skills and education, no significant impact of reasoning ability on earnings existed.

of respondents that cognitive skills and education cannot provide. Cognitive skills are assessed in terms of language comprehension, English reading skills (word-reading, translation, and sentence evaluation) and numeracy. The language comprehension skills are evaluated by checking the degree of understanding a short paragraph. Since the information on test scores was available only for participants who took part in the third-round survey, the subsequent empirical analysis excluded all previously surveyed individuals who did not take the test.¹³ When these tests are too difficult or too easy, it is possible that the scores do not reveal any difference in native ability and skills among respondents, showing that all achieve either full or zero mark. Figures 3.3 and 3.4 provide the graphical representation of the distribution of test scores, where the full mark of each score is normalised to unity. Whilst there is a relatively large proportion of those who achieved zero mark in Raven’s test in both countries, and language comprehension in the case of Ghana, all scores still have sufficient variation and, hence, do not reject the utility.

Technically speaking, our identification strategy, relying on these test scores, boils down to the conditional independence assumption (CIA) (e.g., see Angrist and Pischke (2009, pp. 52-59) for a brief overview). As explained above, our study has to address a number of and, possibly, inextricably linked selection problems: non-random selection into migration (pre-selected migrants), return migration, and wage/self employment. Despite many types of urban residents, such as migrants and natives who can further break down into the wage employed, the self-employed, and the unemployed, the CIA means that conditional on the test scores (as well as other covariates, such as education and age), our observations exploited for an empirical analysis are a random sample of urban population, resulting in the elimination of any selection bias.

Whilst these test scores can be accepted to adequately capture individuals’ unobserved traits specific to ability and cognitive skills and they are, together with a vector $\mathbf{x}_i$, a powerful control for sample selection problems, a potential threat to the CIA should be recognised. If non-reasoning and non-cognitive attributes, such as diligence, motivation, and determination, are valued in labour markets and drive any non-random sample selection, this makes the interpretation of the estimation results difficult, as the test scores do not control for those attributes. For example, it is certainly possible that migrants are, on average, more self-determined than urban natives, as they are already pre-selected from their origins with full of strong ambition to succeed in cities. It is also likely that

¹³This resulted in excluding about 19% and 28% of worker-year observations in Ghana and Tanzania, respectively.

self-employed workers are more motivated than wage-dependent ones by virtue of their decision to run a business despite the risk of the investment. As our data set does not contain an appropriate proxy for these non-reasoning and non-cognitive traits, this issue will be discussed again in Section 3.5.2, when checking the robustness of the findings.

With respect to the standard measurement error bias related to repeated migration discussed in Section 3.2.3 (iii), unfortunately, the data set generated through the survey process does not contain rich enough information to address the problem. However, the subsequent analysis finds evidence to support the remarkably large earnings disadvantage of recently arrived wage employed migrants in both countries, conditional on observable human capital variables. This suggests that the evidence should not be solely attributed to the measurement error resulting from the failure to measure the true destination experience of repeated migrants.

3.4 Data

The data used in this chapter is drawn from the Urban Household Panel Surveys (UPS) in Ghana and Tanzania, which have been conducted in several urban areas in both countries since 2003 by the Centre for the Study of African Economies (CSAE) at University of Oxford, in collaboration with the Ghana Statistical Office (GSO) and the Tanzania National Bureau of Statistics (NBS).¹⁴ Whilst these surveys are still continuing as of 2008, this chapter uses the data of the first three rounds (2003/2004, 2005, and 2006) in the public domain of the CSAE.

The UPS samples are a stratified random sample of urban households, based upon geographic locations, drawn from the 2000 census in Ghana and the 2000 Household Budget Survey (HBS) in Tanzania. Whilst the initial sample was household-based, the UPS enabled us to construct a panel data on the basis of individuals interviewed, as the surveys collected various information about household members aged 15 to 60 in all rounds, such as earnings and education.¹⁵ Due to several reasons such as the absence of previous household members and the potential attrition of households,¹⁶ an unbalanced panel data from round 1 to 3 was eventually constructed on an

¹⁴These surveys have been made possible by funding from the Economic and Social Research Council (ESRC) and the Department for International Development (DFID) of the UK government.

¹⁵The data suggested that only a few respondents were aged out of this range, for some reason. But, since this was only a small proportion of sample observations (0.9% in Ghana and 6.5% in Tanzania), they were not removed from the data set in the analysis.

¹⁶Since information on households is not publicly available at the time of conducting this research, the rate of attrition is unfortunately unclear.

individual basis for the current study.

In addition to the members of the randomly selected households, the surveys conducted follow-up interviews for workers in manufacturing firms in Ghana, who had been contacted since 1995 in another survey of enterprises. However, this non-random sample in Ghana was excluded from the analysis.

Apart from earnings, age, and migrants' duration of stay in cities, the key variables of our interest, such as a migration dummy, education and gender, are time-invariant throughout the rounds, although a few of them have a little time-variation probably due to measurement errors. In addition, as described in Section 3.3.2, the test scores exploited as a quality control of individuals are available only for respondents who participated in the third-round survey and, as a result, the same test scores were used for the same individual throughout the rounds in an empirical analysis, making these scores time-invariant by construction. As this time-invariance property of key variables considerably prevents us from exploiting the panel advantage of data, the subsequent analysis pools the observations over the three rounds in both countries, resulting in 2,872 worker-round observations in Ghana and 1,793 in Tanzania. Pooling implies that the effect of assimilation is identified primarily using across-worker variation of data, rather than within-worker variation.

The key variables are summarised in Table 3.2 for Ghana and Table 3.3 for Tanzania, in which migrants are divided into five arrival cohorts. In this chapter, individuals are identified as migrants if their current living city differs from their place of birth. By this definition, even if a respondent who was (not) born in a city surveyed moves to another place and returns to the city between rounds, he/she is identified as an urban native (migrant) in that city. Due to a relatively short interval between rounds, however, this study regards this sort of between-round mobility as unlikely.

The characteristics of those variables are broadly described from two perspectives: the difference between sectors conditional on each group (the groups of migrants' arrival cohorts and natives), and the difference between groups conditional on each sector.

Monthly earnings, measured by usual pre-tax salary including allowances for wage employment and net cash takings (profits) from own activities for self-employment, are reported in row (1). As earnings distribution in these countries is highly skewed, we focus on the median as a better measure of central tendency. It can be seen that, with the exception of arrivals prior to 1965

in Ghana, wage employed workers earn more than self-employed workers in all groups in both countries. Given each sector, data demonstrate the tendency for the recent arrivals to earn less than earlier cohorts. This is consistent with either the theory of assimilation, or the lower average quality of recent arrivals.

Interestingly, the previously described measures of reasoning ability and cognitive skills show that wage-dependent workers achieved higher scores than self-employed workers in all tests and groups in both countries. To the extent that Raven's test measures individuals' ability entirely ascribed to heredity and that the cognitive skills measured by the test scores are obtained before joining labour force, this suggests an occupational sorting, in which more able and skilled workers exclusively enter the wage sector because their ability and skills may be more valuable in that sector. The systematic relationship between the sectoral choice and the test scores is a positive finding for the purpose of this study, which aims to rely on the scores as a quality control of workers. This is because the test scores are, indeed, measuring individuals' attributes associated with their performance in labour markets. Education in (3) (the highest completed level of education measured by years) reveals the similar pattern, suggesting that wage workers are more educated than the self-employed. This is not surprising because it is likely that the native ability and schooling determine the cognitive skills of individuals and that these variables are positively correlated.¹⁷ In contrast to the across-sector dimension, no clear picture emerges from the data analysis, which is common to both countries regarding the difference in test scores across groups conditional on each sector. This is, again, a positive finding for this study. One econometric concern discussed in Section 3.2.3 (i) was that, in the absence of a good quality control, the rate of assimilation in a cross-sectional estimation might be overestimated (underestimated) when the quality of recent arrivals is lower (higher) than that of earlier cohorts. Whilst the tables are just a descriptive summary, the chosen measures of reasoning ability and cognitive skills do not provide evidence to support the systematic quality difference across arrival cohorts.

It can be seen from (2) that workers' age is relatively similar between sectors in all groups in

¹⁷Indeed, data show the positive correlation between education and the test scores in both countries. The correlation coefficient is available upon request. However, the roles of education and reasoning ability in determining cognitive skills are not the only reason for the positive correlation among these variables. Regarding the relationship between education and the native ability, for example, more educated people are more used to taking tests and performing under controlled conditions, rather than that they are simply more intelligent. In contrast, it is likely that more intelligent individuals would be more educated, as their ability would make academic achievements easier to attain.

both countries. Whilst the age difference is not markedly different, migrants are on average older than natives in both sectors in both countries. Not surprisingly, all sectors in both countries show the tendency that migrants become older as they stay in cities longer time. This is because earlier cohorts inevitably consist of migrants who have been alive since their arrival. In all groups in both countries, there is a larger male proportion in a wage sector than in a self-employed sector, but there is no clear systematic gender difference between groups in both sectors in both countries. As anticipated from discussions in Section 3.2.3 (i), age at migration tends to monotonically decrease as migrants' urban experience increases in both sectors in both countries (Friedberg, 1992). Migrants who came to cities a long time ago and are still in labour force must have been quite young when they arrived. Similarly, those who arrived in recent years and are already in employment would have been mature when they migrated. With respect to the difference in age at migration between sectors, the data do not suggest any identifiable consistency across groups and countries.

3.5 Estimation Results

3.5.1 Main Results

Table 3.4 summarizes the main OLS estimates of interest in specifications (3.10), (3.11), (3.12), (3.14), (3.15) and (3.16) in both countries. The full regression results of (3.10), (3.11) and (3.12) are reported in Appendix Tables 3.6 (Ghana) and 3.7 (Tanzania) and those of (3.14), (3.15) and (3.16) are given in Appendix Tables 3.8 (Ghana) and 3.9 (Tanzania), respectively. The results tell a remarkably similar story across specifications and countries.

Across-group Comparison in Each Sector: Assimilation Effect

In the controls of individuals' observed characteristics, such as education, reasoning ability, cognitive skills and ethnicity, migrants in a wage sector suffer an earnings disadvantage at their arrival relative to otherwise comparable natives, and their earnings rise as their urban experience increases. For example, the simulation based on the results in column (c), which assumes that all $\mathbf{x}_i$ variables and the period effects are equivalently rewarded across groups, suggests that, in Ghana, wage employed migrants aged 20 on arrival initially earn 41% less than otherwise comparable wage employed natives and that it would take 11 years for their earnings to reach the level of comparable

native employees. Similarly, column (f) suggests that, in Tanzania, migrants who arrive at age 20 and enter the wage sector are at a 55% initial earnings disadvantage relative to comparable wage employed natives and their earnings converge to that of natives after 16 years of urban experience.¹⁸ Taking into account life expectancy at birth in these countries,¹⁹ the time during which wage employed migrants undergo lower earnings than comparable wage-employed natives is not inconsiderable. These simulated earnings profiles are given in Appendix Figure 3.5.

In contrast to the evidence in a wage sector suggesting migrants' gradual assimilation, the results do not reveal a similar story in a self-employed sector. Conditional on individuals' attributes, in this sector we do not have strong evidence to support the difference in earnings, attributed to migrants' destination experience, between migrants and natives.

Across-sector Comparison in Each Group

It is evident from the specifications (3.10), (3.11), (3.14) and (3.15) that if the returns to $\mathbf{x}_i$ and the period effects are allowed to vary across sectors, there is no significant difference in the average return between sectors ($\beta_0^k - \beta_0^h$ and $\gamma_0^k - \gamma_0^h$). This suggests that, on average, wage-employed natives and self-employed natives are equally well-off, conditional on individuals' observed attributes. In Tanzania, this is true even if the homogenous returns across all groups are imposed on a vector $\mathbf{x}_i$ and the period effects (columns (f) and (l)).

The above explanation for natives directly applies to migrants who live in an urban area for a long time. In contrast to those established (or lifetime) migrants, however, the earnings of recent arrivals substantially differ, depending upon the sector they work in. All estimation results suggest that recent arrivals working in a wage sector earn *less* than those in a self-employed sector, conditional on individuals' observed attributes. For example, column (g) in Ghana estimates that wage-employed migrants who arrived between 1992 and 2006 earn 35% less than comparable self-employed migrants who arrived during the same period,²⁰ with the numerical value of 37% in

¹⁸By using 1970, 1980 and 1990 census data, Borjas (1995) found that immigrants to the US who arrived in the late 1980s earned about 19.3% less than natives at the entry, conditional on educational attainment, and that the annual growth rate of earnings was about 0.5%. If immigrants to the US are compared to natives of the same ethnicity, the initial cost at the entry and the growth rate of earnings usually vary across ethnic groups. For example, LaLonde and Topel (1991) show based on the 1980 census data that Asians are initially earning 33% less than natives of the same ethnicity and their earnings rise at the rate of about 2.4% per year in the first 10 years after their arrival. Similarly, the initial gap with ethnic natives and the annual rate of earnings growth are estimated as 40% and about 3% for Middle Easterners, respectively. In comparison, Friedberg (2000) estimated the initial cost and the growth rate of earnings of Jewish immigrants from the former Soviet Union at about 25% and 0.8% per year, respectively.

¹⁹Life expectancy at birth between 2000 and 2005 is 57.1 for Ghana and 52.0 for Tanzania according to World Population Prospects: The 2008 Revision Population Database of the United Nations (<http://esa.un.org/unpp/>).

²⁰Since a coefficient on a wage employed dummy is insignificantly different from zero in column (g), there is

column (l) for Tanzania.

Across-group-Across-sector Comparison

The above across-group comparison implies that, in each sector, migrants who live in cities for a long time cannot be differentiated from natives, conditional on individuals' observed traits. Given this result, the urban residents can be re-grouped into (A) recently arrived wage employed migrants, (B) the other wage employed (i.e., wage employed natives and established migrants in a wage sector), (C) recently arrived self-employed migrants, and (D) the other self-employed (i.e., self-employed natives and established migrants in a self-employed sector). Since no strong evidence exists to support migrants' assimilation in the self-employed sector, and all but columns (c) and (i) show no significant difference in natives' average earnings between the wage and self-employed sector, it seems that the earnings of recently arrived self-employed migrants are not significantly different from those of both (B) the other wage employed and (D) the other self-employed, in the control of individuals' characteristics. Since recently arrived wage employed migrants earn less than recently arrived self-employed migrants, all these suggest that only recently arrived wage employed migrants earn less than the other groups in urban labour markets, conditional on individuals' attributes.

3.5.2 Robustness Checks

Two key points stand out from the estimation results in Table 3.4. Firstly, the concept of assimilation is observed *only* in a wage sector, and, secondly, recently arrived wage employed migrants earn *less* than recently arrived self-employed migrants, conditional on individuals' observed attributes. These findings draw a very surprising picture; after controlling for individuals' characteristics and allowing the returns to those characteristics to vary across urban residents, migrants achieve *lower* lifetime earnings by entering a wage sector than by entering a self-employed sector. Given the common belief that wage jobs are better than self-employed jobs in less developed countries, this result for migrants is truly remarkable.

no difference in the average earnings between the wage-dependent and self-employed native workers. In addition, column (g) rejects the hypothesis of assimilation in a self-employed sector. Consequently, all these findings suggest that recently arrived self-employed migrants are very similar to both the wage and self-employed natives in the control of individuals' attributes. Thus, it follows that the difference in the earnings of recent arrivals between sectors is reduced solely to the difference in earnings between recently arrived wage employed migrants and the other wage-employed workers, yielding an estimate, $1 - \exp(-0.432) \approx 0.35$.

Before interpreting the findings from this study, however, it is important to check the robustness of our findings. For that purpose, several exercises were conducted. Firstly, language comprehension, English reading skills, and numeracy might drive earnings not independently but in combination. This implies single, unique measure of cognitive skills, which is calculated as a sum of marks from each separate test. When this total score was used in the estimations, all the key results were almost unaffected. Secondly, in discrete specifications of (3.14), (3.15) and (3.16), the recent migrants were redefined by expanding (1991 to 2006 arrivals) or reducing (1993 to 2006 arrivals) the grouping period by one year. The implications remained unchanged. These estimation results are available from the author if required.

Despite the best effort to include explanatory variables that should be sufficient to control for the quality difference across individuals, it is still uncertain if some individuals' unobserved quality, not measured by our covariates, could contribute to the present findings. Unfortunately, this study cannot completely rule out the possibility that the lower average quality of recent arrivals compared to that of earlier cohorts is the reason for the positive correlation between migrants' urban experience and earnings in a wage sector.

With respect to earnings differentials between recently arrived wage employed migrants and recently arrived self-employed migrants, however, more evidence is available. By restricting the sample to recently arrived migrant workers only (i.e., removing the influence of potential quality difference across arrival cohorts), the following earnings function was estimated,

$$\log w_{it} = \pi_0 + \pi_1 k_i + \pi_2 g_i + \pi_3' \mathbf{x}_i + \pi_4' (\mathbf{x}_i - E(\mathbf{x}_i)) \cdot k_i + \delta_t + \delta_t \cdot k_i + \epsilon_{it}. \quad (3.17)$$

In the actual estimation, several asset dummies, such as dummies for a fridge and mobile phone, are also used as a dependent variable in (3.17) because the asset position is expected to be a form of labour market outcomes. It is not difficult to expect a negative π_1 , as Table 3.4 showed that recently arrived wage employed migrants earned less than recently arrived self-employed migrants.

Two main selection issues appear in estimating (3.17). One is sample selection into recent arrivals, as we are exploiting only the part of whole observations. The other is, as before, selection into wage/self employment, as the unemployed are excluded from the estimation. Therefore, the CIA again provides the justification for the causal interpretation of the negative π_1 . That is,

the recently arrived migrant workers exploited in the estimation are a random sample of urban population after controlling for $\mathbf{x}_i$ (test scores, education, age etc.).

Table 3.5 reports the estimated π_1 of earnings and asset equations. The full regression results are given in Appendix Tables 3.10 and 3.11 for Ghana and Appendix Tables 3.12 and 3.13 for Tanzania. It seems that restricting the sample size is generating imprecise and inefficient point estimates, in particular, in Tanzanian sample.²¹ However, almost all estimates are negative and some are significantly different from zero. For example, being wage employed significantly reduces the probability (relative to being self-employed) that recently arrived migrants would own a fridge and mobile phone.²²

The violation of the CIA arises from several sources. The first possibility is that the test scores are not powerful enough to control for the reasoning ability and cognitive skills of individuals. Tables 3.2 and 3.3, however, presented that education and the test scores are positively correlated with being in a wage sector. If native ability and cognitive skills, not captured by the test scores, also have positive correlation with being in a wage sector and such unobservables positively drive labour market outcomes, the estimated π_1 must have been biased upwards and, hence, the true π_1 should be more negative. This provides further support for the view that recent arrivals earn less by choosing to be wage workers than by entering the self-employed sector.

In contrast, to the extent that self-employed workers are more motivated and self-determined, and that such unobserved traits, unrelated to skills, are valued in labour markets, it is possible that the estimated π_1 is biased downwards. For those attributes to fully explain the identified large disadvantage of recently arrived wage employed migrants, however, those traits would have to be highly (and even more than ability and skills not captured by the test scores) valuable in urban labour markets. This study simply regards the view as unlikely.

²¹The number of wage employed workers is 92 out of 215 recently arrived migrant workers in Ghana, whereas 34 out of 94 in Tanzania.

²²Since no information on when migrants purchased these assets is available, it is possible that they purchased these assets before starting to work. Then, the causality is opposite, i.e., wealthier migrants are less likely to enter a wage sector. Whichever the causality is, the negative correlation between the asset possession and being a wage worker can produce the same implication for the question of why do not all migrants choose to be self-employed workers despite the strong indication that the sector seems more profitable for migrants, conditional on individuals' attributes and the varying returns to those attributes. One answer suggested below is the role of capital constraints to start a business.

3.5.3 Interpreting the Results

Why do not all recently arrived migrants choose to enter a self-employed sector, which seems more profitable than a wage sector, conditional on individuals' traits, such as ability and skills, and the differing returns to those traits?

One common explanation for the question of why not all migrants enter a self-employed sector, often assumed to be a free-entry sector in the literature, is that migrants aiming at working in a wage sector but having difficulty in finding good jobs prefer to be unemployed, because engaging in a self-employed activity reduces their time to search wage jobs (Fields, 1975). In this case, the observed wage employed migrants are selected from all migrants seeking for wage jobs, which might imply that the observed wage employed migrants have higher skills than migrants' population does on average, provided that a wage sector is skill-demanding. Since the CIA, relying on the test scores, is made partly to exclude this sort of selection-oriented difference among migrants (more generally, across urban residents), however, this will not provide convincing account of our findings.

Another possibility points towards the importance of capital constraints. Starting a business is more costly than being a wage worker because it needs not only physical capital, such as machinery, vehicles and equipment, but also social capital, such as a business network. Then, the present findings might be providing indirect evidence that migrants who are not in possession of such assets cannot enter a potentially profitable self-employed sector. Whilst the literature often presumes that everybody can enter the self-employed sector in urban labour markets, our findings challenge the view.

3.6 Conclusion

The present study investigated how rural-to-urban migrants' earnings compare with those of natives in urban labour markets, shedding light on the link between the recent concentration of population and the outcomes of labour markets in urban Africa. Exploiting data drawn from urban household panel surveys in Ghana and Tanzania, the study applied the analytical framework of assimilation, previously reported in literature, to internal migration. In contrast to international migration, making a distinction between wage employed and self-employed migrants is important in the study of internal mobility taking place in less developed countries. This is because small-scale self-

employed activities create many jobs in the economies of our interest, relative to more advanced economies, and those jobs are expected to be different from wage jobs in terms of skills and capitals required. It is highly likely that migrating to start a business and migrating to be wage employed are fundamentally different mobility processes.

The chapter started with a comprehensive review of the assimilation literature to provide theoretical explanations about underlying mechanisms behind migrants' earnings growth at the destination and to show the main econometric challenges in identifying the effect of assimilation. Based on the findings arising from the literature review, an empirical model was developed. The main econometric challenge in identifying the assimilation effect is that, in the absence of good control of individuals' characteristics, the quality difference across arrival cohorts can affect the rate of assimilation, making the differentiation of migrants' earnings growth from the cohort effects challenging. To address the problem, the analysis directly controlled for the reasoning ability, measured by Raven's intelligence test, and cognitive skills, measured by language comprehension, English reading, and numeracy tests, of individuals.

It was expected that migrants were at an initial earnings disadvantage on their arrival relative to comparable native workers and that their earnings would gradually converge to the same level. After controlling for individuals' attributes and allowing for the varying returns to those attributes across urban residents, however, the analysis supported this hypothesis *only* in the wage sector. With somewhat stronger confidence, we also found that recently arrived migrants earned *less* by choosing to be wage workers than by entering a self-employed sector. In contrast to the popular belief that wage jobs are preferable to self-employed jobs, these findings reveal a surprising picture that the lifetime earnings of wage employed migrants would be *higher* if they entered a self-employed sector on their arrival, conditional on individuals' traits and the heterogeneous returns to those traits. This prompts the question of why do not all migrants enter the self-employed sector, given that it is more profitable. In urban labour markets, capital constraints may constitute an important obstacle for migrants to start a business. It seems untrue that the self-employed sector is a free-entry sector for migrants.

3.7 Appendix: Model in Section 3.2.1 (ii)

3.7.1 Derivation of equation (3.3)

The key part of the derivation is to obtain the posterior belief of the firm. Applying Bayes's rule,

$$\text{Prob}(a_G|N_t = n) = \frac{\text{Prob}(a_G) \cdot \text{Prob}(N_t = n|a_G)}{\text{Prob}(a_G) \cdot \text{Prob}(N_t = n|a_G) + \text{Prob}(a_B) \cdot \text{Prob}(N_t = n|a_B)} \quad (3.18)$$

$$\begin{aligned} & \gamma \binom{t}{n} p_G^n (1 - p_G)^{t-n} \\ &= \frac{\gamma \binom{t}{n} p_G^n (1 - p_G)^{t-n}}{\gamma \binom{t}{n} p_G^n (1 - p_G)^{t-n} + (1 - \gamma) \binom{t}{n} p_B^n (1 - p_B)^{t-n}} \quad (3.19) \end{aligned}$$

$$= \frac{1}{1 + \left(\frac{1-\gamma}{\gamma}\right) \left(\frac{p_B}{p_G}\right)^n \left(\frac{1-p_B}{1-p_G}\right)^{t-n}} \quad (3.20)$$

$$= \frac{1}{1 + \hat{\gamma} \cdot p_0^n \cdot p_1^{t-n}}. \quad (3.21)$$

Similarly,

$$\text{Prob}(a_B|N_t = n) = \frac{\text{Prob}(a_B) \cdot \text{Prob}(N_t = n|a_B)}{\text{Prob}(a_G) \cdot \text{Prob}(N_t = n|a_G) + \text{Prob}(a_B) \cdot \text{Prob}(N_t = n|a_B)} \quad (3.22)$$

$$= \frac{\hat{\gamma} \cdot p_0^n \cdot p_1^{t-n}}{1 + \hat{\gamma} \cdot p_0^n \cdot p_1^{t-n}}. \quad (3.23)$$

Substituting (3.21) and (3.23) into (3.2), we get (3.3).

3.7.2 Extending the Model: Influx of Skilled Immigrants and Occupational Upgrading

The model presented in Section 3.2.1 (ii) is extended to show that a huge influx of skilled immigrants (large γ) generates their occupational upgrading from unemployment to employment by providing the firm with an incentive to replace an initially hired worker with a new one with better skills through the mechanism of search. To simplify this discussion, it is assumed that $p_G = 1$ and $p_B = 0$, suggesting that the firm can perfectly assess the skills of the employed worker after observing his performance in the initial period ($t = 0$). In other words, the wage convergence of good quality employed immigrants towards a_G described in equation (3.4) will be achieved on observing his performance in a single period.

After observing y_H in the initial period ($t = 0$), the best strategy for the firm is obviously to permanently employ the highly performing employee, which generates the present discounted value of the firm's payoff in the period 1 as

$$V_G = \frac{\hat{y}_H}{1 - \delta}, \quad (3.24)$$

where $\hat{y}_H \equiv y_H - a_G$ and δ is a discount factor. If the initially hired worker produced y_L in the period 0, the firm must choose between keeping the low performing employee and firing him to search for a new worker, with the associated cost c . As a simplification, the model assumes that searching, replacing the old worker with a new one, and producing output with the new worker simultaneously take place in one period. Then, after observing y_L in period 0, the present discounted value of the firm's payoff at $t = 1$ when searching becomes

$$V_S = -c + \gamma[y_H - \bar{a} + \delta V_G] + (1 - \gamma)[y_L - \bar{a} + \delta V_S], \quad (3.25)$$

where $\bar{a} \equiv \gamma a_G + (1 - \gamma)a_B$. With (3.24) and (3.25), we get

$$V_S = \frac{\gamma \hat{y}_H + (1 - \gamma)(1 - \delta)\hat{y}_L - c(1 - \delta)}{(1 - \delta)(1 - \delta + \delta\gamma)}, \quad (3.26)$$

where $\hat{y}_L \equiv y_L - a_B$. Thus, the firm initially matched with a low performing immigrant is willing to search in period 1 if and only if the present discounted value of searching exceeds that of keeping him:

$$V_B = \frac{\hat{y}_L}{1 - \delta} \leq V_S. \quad (3.27)$$

Since it can be shown that $\frac{\partial V_S}{\partial \gamma} > 0$,²³ the huge influx of skilled immigrants is likely to increase the probability of high quality unemployed immigrants being employed (occupational upgrading) by providing the firm that initially hired an unsuitable worker with an incentive to search for a new skilled worker.

²³ $\frac{\partial V_S}{\partial \gamma} = \frac{\hat{y}_H - \hat{y}_L + \delta c}{(1 - \delta + \delta\gamma)^2} > 0$.

3.8 Appendix: Recent Arrivals Artificially Created by Repeated Migration

3.8.1 Model

Even though the literature has not focused on this problem, a systematic measurement error about migrants' true destination experience could generate the pseudo assimilation effect in cross-sectional data. Due to the lack of accurate information, the previously conducted studies typically measured migrants' destination experience as the length of time since their most recent arrival, ignoring the possibility that some might have been repeated migrants. This procedure is, thus, likely to underestimate the true destination experience of repeated migrants, as it regards them as recent arrivals. In competitive destination labour markets, the more capable migrants are more likely to continuously stay in the destination once they arrive. This would imply that the repeated migrants are of relatively low quality; thus, failing to measure their true destination experience artificially generates positive correlation between migrants' destination experience and their earnings in cross-sectional data, which, at first glance, appears to be the assimilation effect.

To formalise the idea in an example of rural-to-urban migration, we consider a hypothetical rural economy (origin) where a continuum of risk-neutral agents with identical preferences and infinite lifetime horizons are living with the size of population normalized to 1 in period 0. In this model, the ability of agents is uniformly distributed as $a \sim U[0, 1]$. In each period $t \geq 1$ ($t = 1, 2, 3, \dots$), agents decide whether to stay in their origin or move to an urban city, depending on earnings that they are able to obtain in each location. Furthermore, for simplification of the model, we assume that agents are able to change locations without any cost. An agent i 's earnings from urban jobs in the period t , w_{it} , are defined as $a_i - \theta \cdot M(t - 1)$ with $0 < \theta < 1$, where the population of agents who stay in the city in the period t is expressed as $M(t)$ with $M(0) = 0$. With strictly positive θ , migrants' earnings in the period t decline with the number of agents in the city in the previous period.²⁴ In this sense, θ can be seen as the degree of competitiveness in urban labour markets with greater intensity reflected by its higher value. When an agent i stays in his origin, he can earn income w_{it}^R , assumed to be $w_{it}^R = 0 \forall i$ and t for the sake of simplicity.

²⁴One alternative model is $w_{it} = 2a_i - \theta \cdot (M(t - 1) + M(t))$. This enables migrants' current population to affect their current earnings. But, in both cases, the dependence of the earnings on the previous population needs to allow for an initial influx of migrants, either in the absence of, or in the weak presence of, competition in period 1. As key implications remain unchanged, in the present study the current population was excluded when defining migrants' earnings. The solution in the alternative model can be obtained from the author if required.

The population of those who stay in their origin in the period t is expressed as $R(t) = 1 - M(t)$ with $R(0) = 1$.

With these assumptions, an agent i decides to stay in the city in the period t if and only if the present discounted value of staying there exceeds that of staying in their origin:

$$\sum_{s=t-1}^{\infty} \delta^{s-t+1} [a_i - \theta \cdot M(s)] \geq 0, \quad (3.28)$$

where $0 < \delta < 1$ is the discount factor. To characterize the equilibrium path of the economy, we need the following lemma:

Lemma 2 *For all a and t , if an agent of type a' prefers to stay in the urban economy, all agents $a > a'$ are strictly better-off in the urban economy.*

Proof The proof is derived by contradiction. See Appendix 3.8.2.

Lemma 1 suggests that there exists an ability threshold $\bar{a}_t$ in each period below which all agents prefer to stay in their origin:

$$\sum_{s=t-1}^{\infty} \delta^{s-t+1} [\bar{a}_t - \theta \cdot M(s)] = 0. \quad (3.29)$$

Replacing $\bar{a}_t$ with $R(t)$ ²⁵ and $M(s)$ with $1 - R(s)$ in (3.29), we can get (see Appendix 3.8.3 for the derivation)

$$R(t) = \left(\frac{1}{1 + \theta} \right) \left[\theta + \left(\frac{\sqrt{1 + 4\theta\delta(1 - \delta)} - 1}{2\delta} \right)^t (-1)^t \right], \quad (3.30)$$

where $0 < \frac{\sqrt{1 + 4\theta\delta(1 - \delta)} - 1}{2\delta} < 1$. Now the equilibrium of the economy can be explained. Firstly, the fraction $\theta/(1 + \theta)$ ($\equiv R^*$) of agents decides to stay in their origin in the long-run steady state ($t \rightarrow \infty$). The more competitive urban labour markets are, the more agents prefer to stay in their origin ($\frac{dR^*}{d\theta} > 0$). Appendix Figure 3.6 shows the path to the steady state. In period 1, the most talented group of agents $a_i \in [R(1), 1]$ migrate to the city and each migrant i fully enjoys high income opportunities by earning a_i , due to the lack of competition in period 1. However, the influx of migrants into the city in period 1 makes it hard for them to survive in period 2 and, as a result,

²⁵Remember that the ability of agents is uniformly distributed in the interval $[0, 1]$.

the least talented group of first-period migrants $a_i \in [R(1), R(2)]$ move back to their origin in period 2. Since the return of some of the initial migrants in period 2 makes urban labour markets less competitive in period 3, moving to the city again becomes attractive for the most talented group of returners $a_i \in [R(3), R(2)]$ in that period. This is how the economy eventually reaches the steady state, oscillating between overpopulated and underpopulated states in the city relative to the steady state. As often cited in the literature, migration is a repeated pattern rather than a single, isolated process. Whilst economists frequently suggest that imperfect information about economic opportunities at the destination as a key mechanism driving such repeated migration (e.g. Vanzo, 1983), the presented model implies that it is possible even in the absence of any information problem.

When attempting to relate the present model to empirical studies, it is essential to address the assumption often made in extant literature that migrants' destination experience can uniformly be measured as the length of time since their most recent arrival year, thus disregarding the effect of repeated migration. According to the rule, migrants' urban experience in each period from period 1 to 12 was summarised by their ability group in Appendix Table 3.14. For example, in period 5, the ability group $a_i \in [R(2), 1]$ has an urban experience of five years because they have continuously lived in the city from period 1. Similarly, the group $a_i \in [R(4), R(2)]$ has lived in the city for three years since period 3. Interestingly, the destination experience of the ability group $a_i \in [R(5), R(4)]$ is counted as one year, although they are repeated migrants who have been in the city in periods 1 and 3. Appendix Table 3.14 shows that there is clearly a positive correlation between migrants' ability (hence, earnings) and their destination experience in the absence of each migrant's assimilation, if a particular period is selected and used for empirical analysis as cross-sectional data. The relationship between migrants' ability and their destination experience is formally summarised in Appendix 3.8.4.

3.8.2 Proof

Proof of Lemma 1: Suppose the contrary. If an agent of type a' prefers to stay in the urban economy and all agents $a > a'$ are weakly better-off in the rural economy at time t , it must be that $V_t(a') \geq 0$ and $V_t(a) \leq 0$ where $V_t(\cdot)$ is an agent's present discounted value of staying in the

city in the period t . This implies that $V_t(a') \geq V_t(a)$. Now we can show for all $t \geq 1$ that

$$V_t(a') - V_t(a) = \sum_{s=t-1}^{\infty} \delta^{s-t+1} (a' - a). \quad (3.31)$$

Since $a' < a$, this results in $V_t(a') < V_t(a)$ for all $t \geq 1$, which is a contradiction.

3.8.3 Derivation of equation (3.30)

The derivation comprises two parts: (a) deriving the second order difference equation of $R(t)$ and (b) solving it. Firstly, substituting $\bar{a}_t = R(t)$ and $M(s) = 1 - R(s)$ into (3.29) yields

$$(R(t) - \theta) \sum_{s=t-1}^{\infty} \delta^{s-t+1} + \theta \sum_{s=t-1}^{\infty} \delta^{s-t+1} R(s) = 0. \quad (3.32)$$

Now let us define the first term as $A(t)$ and the second term as $B(t)$ in (3.32). Then, we can show that

$$\delta A(t+1) - A(t) = \left(\frac{\delta}{1-\delta} \right) R(t+1) - \left(\frac{1}{1-\delta} \right) R(t) + \theta, \quad (3.33)$$

and

$$\delta B(t+1) - B(t) = -\theta R(t-1). \quad (3.34)$$

Then,

$$A(t) + B(t) = 0, \quad (3.35)$$

$$\Rightarrow \delta A(t+1) - A(t) + \delta B(t+1) - B(t) = 0, \quad (3.36)$$

$$\Rightarrow \left(\frac{\delta}{1-\delta} \right) R(t+1) - \left(\frac{1}{1-\delta} \right) R(t) + \theta - \theta R(t-1) = 0. \quad (3.37)$$

Arranging this, we obtain the second order difference equation of $R(t)$ as

$$R(t+1) - \frac{1}{\delta} R(t) - \frac{\theta(1-\delta)}{\delta} R(t-1) = -\frac{\theta(1-\delta)}{\delta}. \quad (3.38)$$

Solving (3.38) is straightforward. Using the lag operator, this can be written as

$$(1 - \mu_1 L)(1 - \mu_2 L)R(t) = -\frac{\theta(1 - \delta)}{\delta}, \quad (3.39)$$

where the characteristic roots μ_1 and μ_2 are

$$\mu_1 = \frac{1 - \sqrt{1 + 4\theta\delta(1 - \delta)}}{2\delta}, \quad (3.40)$$

$$\mu_2 = \frac{1 + \sqrt{1 + 4\theta\delta(1 - \delta)}}{2\delta}, \quad (3.41)$$

where $\mu_1 + \mu_2 = \frac{1}{\delta}$ and $\mu_1\mu_2 = -\frac{\theta(1-\delta)}{\delta}$. Since $-1 < \mu_1 < 0$ and $1 < \mu_2$, solving backward for μ_1 and forward for μ_2 yields

$$R(t) = \mu_1 R(t-1) + \frac{\theta(1 - \delta)}{\delta(\mu_2 - 1)} \quad (3.42)$$

$$= \mu_1^t R(0) + \frac{\theta(1 - \delta)}{\delta(\mu_2 - 1)} \sum_{j=0}^{t-1} \mu_1^j \quad (3.43)$$

$$= \mu_1^t + \frac{\theta(1 - \delta)}{\delta(\mu_2 - 1)} \sum_{j=0}^{t-1} \mu_1^j \quad (3.44)$$

$$= \mu_1^t + \frac{\theta(1 - \delta)}{\delta(\mu_2 - 1)} \left(\sum_{j=0}^{\infty} \mu_1^j - \mu_1^t \sum_{j=0}^{\infty} \mu_1^j \right), \quad (3.45)$$

where we used the initial condition $R(0) = 1$ from the second line to the third. Arranging $\sum_{j=0}^{\infty} \mu_1^j$ is not straightforward, as μ_1 is negative. But it can be done as

$$\sum_{j=0}^{\infty} \mu_1^j = 1 - |\mu_1| + |\mu_1|^2 - |\mu_1|^3 + |\mu_1|^4 - |\mu_1|^5 + \dots \quad (3.46)$$

$$= (1 + |\mu_1|^2 + |\mu_1|^4 + \dots) - (|\mu_1| + |\mu_1|^3 + |\mu_1|^5 + \dots) \quad (3.47)$$

$$= \frac{1}{1 + |\mu_1|}. \quad (3.48)$$

Substituting (3.48) into (3.45), we get

$$R(t) = \mu_1^t + \frac{\theta(1 - \delta)}{\delta(\mu_2 - 1)(1 + |\mu_1|)} (1 - \mu_1^t). \quad (3.49)$$

Since $\mu_1 + \mu_2 = \frac{1}{\delta}$ and $\mu_1\mu_2 = -\frac{\theta(1-\delta)}{\delta}$,

$$(\mu_2 - 1)(1 + |\mu_1|) = -(1 - \mu_2)(1 - \mu_1) \quad (3.50)$$

$$= \frac{(1 + \theta)(1 - \delta)}{\delta}. \quad (3.51)$$

With (3.49) and (3.51), we obtain

$$R(t) = \mu_1^t + \left(\frac{\theta}{1 + \theta}\right)(1 - \mu_1^t) \quad (3.52)$$

$$= \left(\frac{\theta}{1 + \theta}\right) + \left(\frac{1}{1 + \theta}\right)\mu_1^t \quad (3.53)$$

$$= \left(\frac{1}{1 + \theta}\right)(\theta + \mu_1^t). \quad (3.54)$$

Arranging this yields (3.30).

Finally, we want to ensure that $R(t)$ cannot take the value outside the unit interval for all $t \geq 1$ leading to the steady state. For the purpose, it is enough to show that $R(1) > 0$ and $R(2) < 1$ in the present model.

$$R(1) = \left(\frac{1}{1 + \theta}\right) \left(\theta - \frac{\sqrt{1 + 4\theta\delta(1 - \delta)} - 1}{2\delta}\right) \quad (3.55)$$

$$= \left[\frac{1}{2\delta(1 + \theta)}\right] \left[(2\delta\theta + 1) - \sqrt{1 + 4\theta\delta(1 - \delta)}\right] \quad (3.56)$$

$$= \left[\frac{1}{2\delta(1 + \theta)}\right] \left[\sqrt{(2\delta\theta + 1)^2} - \sqrt{(2\delta\theta + 1)^2 - 4\delta^2\theta(1 + \theta)}\right] > 0. \quad (3.57)$$

Next,

$$R(2) = \left(\frac{1}{1 + \theta}\right) \left[\theta + \left(\frac{\sqrt{1 + 4\theta\delta(1 - \delta)} - 1}{2\delta}\right)^2\right] \quad (3.58)$$

$$< \left(\frac{1}{1 + \theta}\right) \left[\theta + \left(\frac{\sqrt{1 + 4\delta(1 - \delta)} - 1}{2\delta}\right)^2\right]. \quad (3.59)$$

Since we can show with $0 < \delta < 1$ that

$$1 + 4\delta - 4\delta^2 < 1 + 4\delta + 4\delta^2 \quad (3.60)$$

$$\Leftrightarrow 1 + 4\delta(1 - \delta) < (2\delta + 1)^2 \quad (3.61)$$

$$\Leftrightarrow \sqrt{1 + 4\delta(1 - \delta)} - 1 < 2\delta, \quad (3.62)$$

$$R(2) < \left(\frac{1}{1 + \theta} \right) \left[\theta + \left(\frac{\sqrt{1 + 4\delta(1 - \delta)} - 1}{2\delta} \right)^2 \right] \quad (3.63)$$

$$< \left(\frac{1}{1 + \theta} \right) \left[\theta + \left(\frac{2\delta}{2\delta} \right)^2 \right] = 1. \quad (3.64)$$

3.8.4 Ability Group and Destination Experience in Each Period

This section formally summarises the relationship between migrants' ability and their destination experience in each period shown in Appendix Table 3.14. Firstly, when $t = 2n$ (even period) or $t = 2n - 1$ (odd period), we can partition migrants into n groups according to their ability for all n ($= 1, 2, 3, \dots, \infty$) as $(A_1, A_2, \dots, A_j, \dots, A_{n-1}, A_n)$ where A_1 is the least talented and A_n is the most talented group.

In an even period $t = 2n$, the group A_j contains migrants such that $a_i \in [R(2(n-j+1)), R(2(n-j))]$ for all j ($= 1, 2, \dots, n$) and all migrants in the group A_j have $2j$ years of continuous destination experience. For example, when $n = 3$ (period 6), migrants are partitioned into three groups (A_1, A_2, A_3) where A_1 contains migrants such that $a_i \in [R(6), R(4)]$ and two years of destination experience, A_2 with $a_i \in [R(4), R(2)]$ and four years, and A_3 with $a_i \in [R(2), 1]$ and six years. This is summarized in Appendix Table 3.14.

In an odd period $t = 2n - 1$, the group A_j contains migrants such that $a_i \in [R(2(n-j+1)), R(2(n-j))]$ for all $j \geq 2$ ($= 2, 3, \dots, n$), whereas the group A_1 ($j = 1$) with $a_i \in [R(2(n-j+1) - 1), R(2(n-j))]$. The group A_j has $2j - 1$ years of continuous destination experience. For example, when $n = 3$ (period 5), migrants are partitioned into three groups (A_1, A_2, A_3) where A_1 contains migrants such that $a_i \in [R(5), R(4)]$ and one year of destination experience, A_2 with $a_i \in [R(4), R(2)]$ and three years, and A_3 with $a_i \in [R(2), 1]$ and five years. This is, again, summarised in Appendix Table 3.14.

Table 3.1: Urban Population Data

Region	Percentage of urban population (%)			Average annual growth rate of urban population (%)		
	1950	1975	2005	1950 to 1975	1975 to 2000	2000 to 2005
Africa	14.5	25.7	37.9	4.76	4.06	3.38
Ghana	15.4	30.1	47.8	5.37	4.20	4.19
Tanzania	3.5	11.1	24.2	7.59	5.78	4.19
Asia	16.8	24.0	39.7	3.54	3.49	2.62
Europe	51.2	65.7	71.9	1.84	0.63	0.21
Latin America & the Caribbean	41.4	61.1	77.5	4.20	2.75	1.86
Northern America	63.9	73.8	80.7	1.98	1.32	1.42
World	29.1	37.3	48.6	2.89	2.52	2.07

Source: Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, World Urbanization Prospects: The 2007 Revision Population Database (<http://esa.un.org/unup/>).

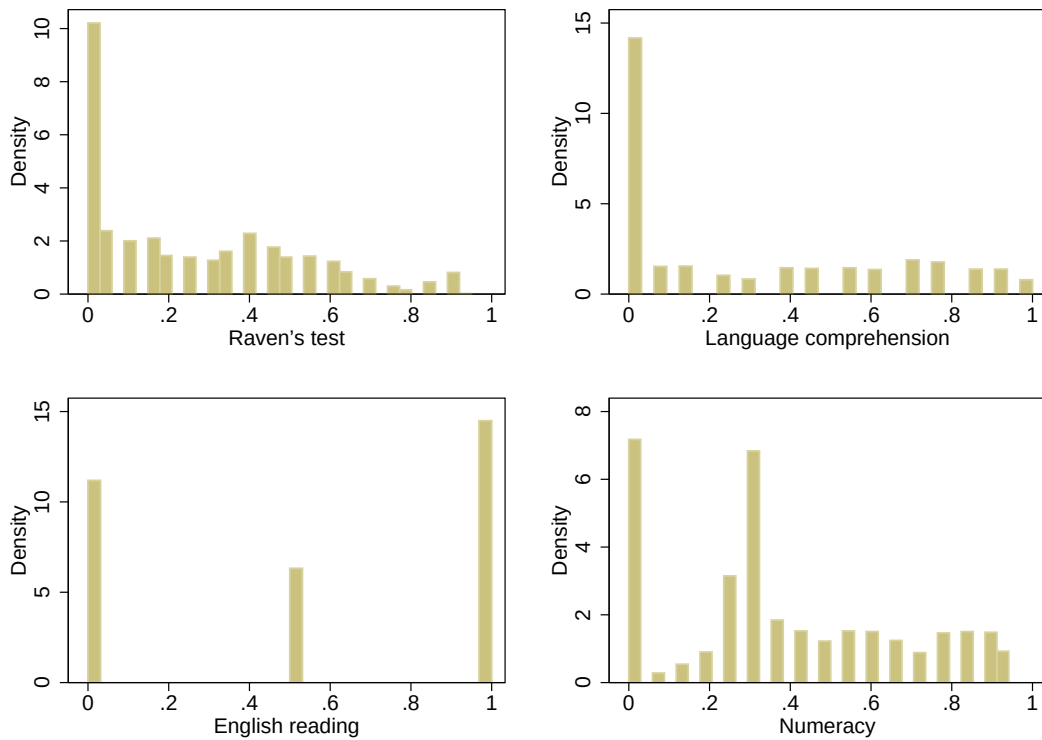


Figure 3.3: Histogram of Reasoning Ability and Cognitive Skills in Ghana

Notes: (1) The full mark of each test is normalized to be one. (2) The full mark of each test is 20, 13, 2, and 17 for Raven, language comprehension, English reading, and nemeracy, respectively.

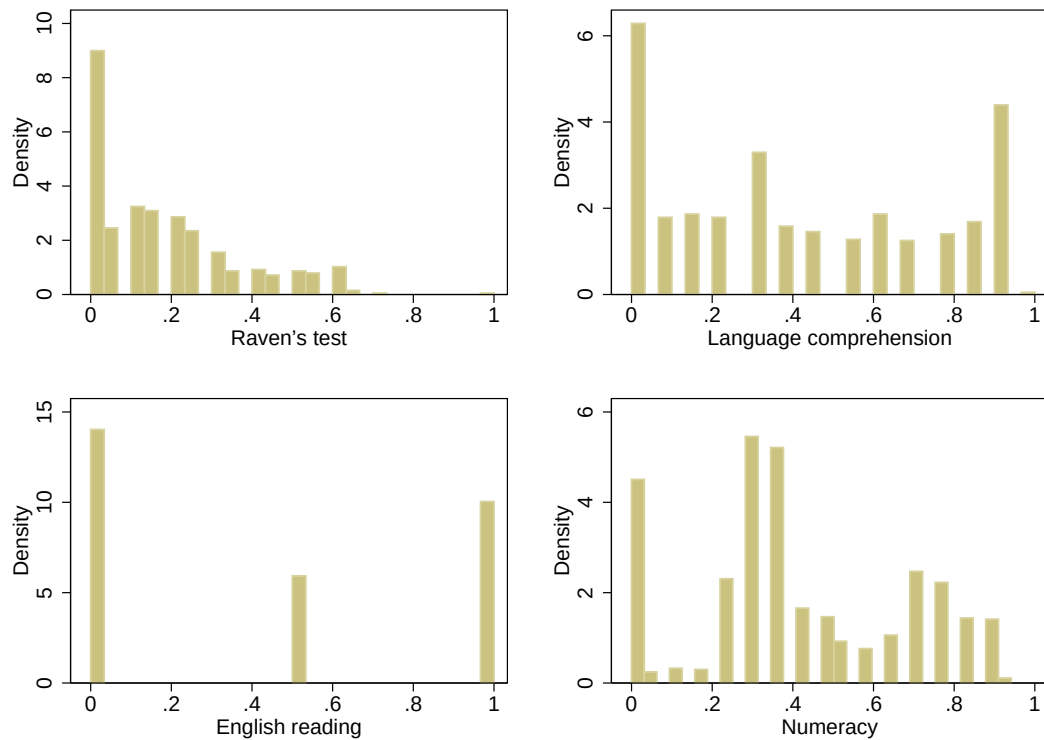


Figure 3.4: Histogram of Reasoning Ability and Cognitive Skills in Tanzania

Notes: (1) The full mark of each test is normalized to be one. (2) The full mark of each test is 20, 13, 2, and 17 for Raven, language comprehension, English reading, and nemeracy, respectively.

Table 3.2: Descriptive Statistics - Ghana

		Migrants						Natives
	Arrival years	1996 to to 2006	1986 to 1995	1976 to 1985	1966 to 1975	Before 1965	Total	
(1) Monthly income (US dollar)								
Wage	Mean	39.73	72.59	90.80	86.59	74.48	66.95	64.41
	Std.	(29.57)	(96.42)	(132.04)	(51.11)	(86.81)	(87.00)	(93.15)
	Median	32.93	48.52	65.86	86.03	43.91	46.18	42.88
	No. of obs.	85	88	50	26	4	257	664
Self	Mean	53.32	57.81	89.95	121.29	99.96	73.22	65.83
	Std.	(108.41)	(65.04)	(128.55)	(254.75)	(82.02)	(126.88)	(122.19)
	Median	28.27	36.34	50.90	52.69	78.41	41.42	36.59
	No. of obs.	120	138	129	44	13	448	1143
(2) Age (years)								
Wage	Mean	24.86	33.70	44.10	49.78	47.85	34.03	30.53
	Std.	(5.86)	(7.51)	45	(6.36)	(13.27)	(12.02)	(10.20)
	No. of obs.	110	91	54	24	7	290	826
Self	Mean	30.43	35.49	41.03	46.50	51.90	37.41	35.74
	Std.	(7.25)	(6.94)	(7.79)	(6.28)	(9.38)	(9.30)	(10.28)
	No. of obs.	126	144	138	49	15	477	1235
(3) Education (years)								
Wage	Mean	8.14	8.91	9.22	7.88	8.71	8.56	8.80
	Std.	(4.09)	(3.34)	(3.46)	(4.32)	(4.07)	(3.79)	(3.70)
	No. of obs.	110	91	54	26	7	292	825
Self	Mean	5.96	6.97	6.91	5.61	4.33	6.47	6.44
	Std.	(4.29)	(3.84)	(3.98)	(4.48)	(5.74)	(4.17)	(4.40)
	No. of obs.	126	143	135	49	15	473	1223
(4) Gender (one if male)								
Wage	Mean	0.50	0.82	0.68	0.68	0.71	0.66	0.59
	Std.	(0.50)	(0.38)	(0.46)	(0.47)	(0.48)	(0.47)	(0.49)
	No. of obs.	109	91	54	25	7	290	822
Self	Mean	0.36	0.31	0.26	0.35	0.40	0.32	0.31
	Std.	(0.48)	(0.46)	(0.44)	(0.48)	(0.50)	(0.46)	(0.46)
	No. of obs.	126	145	139	48	15	478	1235
(5) Age at migration								
Wage	Mean	19.93	20.04	19.10	16.37	6.28	19.18	-
	Std.	(5.33)	(7.43)	(11.84)	(5.05)	(7.51)	(7.93)	-
	No. of obs.	110	91	54	24	7	286	-
Self	Mean	25.19	21.39	17.41	13.52	9.19	20.04	-
	Std.	(7.19)	(6.67)	(7.22)	(6.55)	(7.62)	(8.15)	-
	No. of obs.	126	144	138	49	15	472	-
(6) Urban experience (years)								
Wage	Mean	5.92	14.65	26.03	34.41	46.17	15.73	-
	Std.	(2.74)	(2.89)	(3.10)	(2.71)	(2.71)	(10.72)	-
	No. of obs.	110	91	53	24	6	284	-
Self	Mean	6.23	15.09	24.71	33.97	44.45	18.17	-
	Std.	(2.48)	(2.70)	(2.90)	(2.69)	(2.46)	(10.20)	-
	No. of obs.	126	144	135	49	11	465	-
(7) Test scores								
(a) Raven's test (out of 20)								
Wage	Mean	6.36	6.28	7.48	7.12	7.80	6.58	5.97
	Std.	(5.06)	(5.31)	(5.03)	(5.20)	(6.09)	(5.15)	(5.52)
	No. of obs.	85	82	39	16	5	231	700
Self	Mean	3.72	4.29	4.06	3.54	2.27	3.92	3.49
	Std.	(4.65)	(4.81)	(4.34)	(5.26)	(4.05)	(4.64)	(4.30)
	No. of obs.	102	105	104	37	11	363	1004
(b) Language comprehension (out of 13)								
Wage	Mean	3.90	4.50	5.05	5.68	5.20	4.46	4.60
	Std.	(4.70)	(4.26)	(4.33)	(4.42)	(4.60)	(4.45)	(4.56)
	No. of obs.	85	82	39	16	5	231	700
Self	Mean	2.43	2.30	2.28	3.05	2.54	2.41	2.12
	Std.	(3.34)	(2.98)	(3.65)	(3.80)	(3.88)	(3.38)	(3.50)
	No. of obs.	102	105	104	37	11	363	1004
(c) English reading (out of 2)								
Wage	Mean	1.14	1.30	1.46	1.50	1.20	1.28	7.07
	Std.	(0.81)	(0.81)	(0.82)	(0.89)	(1.09)	(0.83)	(5.01)
	No. of obs.	85	82	39	16	5	231	6.00
Self	Mean	0.92	0.80	0.79	0.78	0.63	0.83	0.78
	Std.	(0.90)	(0.86)	(0.89)	(0.91)	(0.92)	(0.89)	(0.89)
	No. of obs.	102	105	104	37	11	363	1004
(d) Numeracy (out of 17)								
Wage	Mean	7.64	7.56	8.48	7.43	7.20	7.68	7.07
	Std.	(4.65)	(4.66)	(4.57)	(4.83)	(6.57)	(4.66)	(5.01)
	No. of obs.	85	82	39	16	5	231	700
Self	Mean	4.59	5.18	4.57	4.27	3.27	4.69	4.26
	Std.	(3.40)	(3.52)	(3.90)	(4.98)	(5.19)	(3.82)	(4.05)
	No. of obs.	105	105	104	37	11	363	1004

Note: Monthly income is expressed in US dollars as of 2000.

Table 3.3: Descriptive Statistics - Tanzania

		Migrants						Natives
		Arrival years	1996 to to 2006	1986 to 1995	1976 to 1985	1966 to 1975	Before 1965	Total
(1) Monthly income (US dollar)								
Wage	Mean	74.91	153.15	143.03	200.04	126.09	145.30	121.00
	Std.	(65.76)	(291.85)	(135.97)	(258.20)	(134.42)	(214.27)	(164.38)
	Median	61.38	102.78	100.13	104.64	78.37	94.04	83.81
	No. of obs.	34	71	65	45	25	240	352
Self	Mean	74.06	89.69	101.90	156.01	81.80	100.23	98.51
	Std.	(91.68)	(155.30)	(155.43)	(267.07)	(126.07)	(169.41)	(454.60)
	Median	48.01	47.02	52.57	63.53	50.06	50.06	60.01
	No. of obs.	53	112	96	58	45	364	727
(2) Age (years)								
Wage	Mean	28.21	37.92	44.77	52.48	52.95	42.51	36.67
	Std.	(6.64)	(7.46)	(8.56)	(7.20)	(12.37)	(11.52)	(10.71)
	No. of obs.	36	73	67	46	23	245	364
Self	Mean	33.59	36.18	41.96	49.34	53.71	41.56	37.40
	Std.	(12.28)	(7.93)	(8.12)	(7.78)	(11.29)	(11.30)	(11.93)
	No. of obs.	55	113	107	59	45	379	745
(3) Education (years)								
Wage	Mean	8.02	10.57	9.76	9.57	8.81	9.61	8.61
	Std.	(5.07)	(3.64)	(4.26)	(4.73)	(5.07)	(4.45)	(4.17)
	No. of obs.	36	73	67	47	27	250	369
Self	Mean	7.00	6.62	6.89	7.95	6.22	6.90	6.41
	Std.	(4.38)	(3.74)	(3.50)	(4.36)	(4.14)	(3.94)	(3.69)
	No. of obs.	55	116	109	60	50	390	769
(4) Gender (one if male)								
Wage	Mean	0.50	0.38	0.61	0.78	0.66	0.56	0.57
	Std.	(0.50)	(0.61)	(0.49)	(0.41)	(0.48)	(0.49)	(0.49)
	No. of obs.	36	73	67	47	27	250	371
Self	Mean	0.40	0.39	0.43	0.51	0.56	0.44	0.49
	Std.	(0.49)	(0.49)	(0.49)	(0.50)	(0.50)	(0.49)	(0.50)
	No. of obs.	55	116	109	60	50	390	775
(5) Age at migration								
Wage	Mean	22.41	23.03	20.62	17.96	9.73	20.08	-
	Std.	(6.11)	(7.32)	(8.46)	(6.89)	(8.93)	(8.42)	-
	No. of obs.	36	73	67	46	23	245	-
Self	Mean	27.37	21.64	17.61	15.90	8.26	18.85	-
	Std.	(11.70)	(8.19)	(7.71)	(7.50)	(7.85)	(10.00)	-
	No. of obs.	55	113	107	59	45	379	-
(6) Urban experience (years)								
Wage	Mean	6.80	15.89	25.14	35.52	45.61	23.10	-
	Std.	(2.40)	(2.69)	(3.22)	(2.97)	(4.80)	(11.58)	-
	No. of obs.	36	73	67	46	18	240	-
Self	Mean	7.21	15.53	25.34	34.44	47.19	22.90	-
	Std.	(2.38)	(3.06)	(2.80)	(2.45)	(5.55)	(11.70)	-
	No. of obs.	55	113	107	59	31	365	-
(7) Test scores								
(a) Raven's test (out of 20)								
Wage	Mean	6.20	6.07	4.27	3.96	4.00	5.01	4.68
	Std.	(4.60)	(4.77)	(4.17)	(3.67)	(3.94)	(4.40)	(3.85)
	No. of obs.	30	57	55	32	21	195	291
Self	Mean	4.35	4.61	4.05	3.08	2.75	3.95	3.49
	Std.	(3.41)	(3.74)	(3.36)	(3.57)	(2.86)	(3.49)	(3.33)
	No. of obs.	40	78	79	34	37	268	543
(b) Language comprehension (out of 13)								
Wage	Mean	5.90	7.10	6.41	6.81	4.52	6.42	6.18
	Std.	(3.60)	(3.78)	(4.44)	(4.62)	(4.00)	(4.15)	(3.99)
	No. of obs.	30	57	55	32	21	195	291
Self	Mean	5.77	5.75	4.75	4.73	4.27	5.13	6.09
	Std.	(3.33)	(3.63)	(3.93)	(3.41)	(4.38)	(3.78)	(4.22)
	No. of obs.	40	78	79	34	37	268	543
(c) English reading (out of 2)								
Wage	Mean	1.13	1.24	1.12	1.09	1.14	1.15	1.09
	Std.	(0.93)	(0.93)	(0.94)	(0.92)	(1.01)	(0.93)	(0.90)
	No. of obs.	30	57	55	32	21	195	291
Self	Mean	0.77	0.75	0.59	0.76	0.81	0.72	0.81
	Std.	(0.94)	(0.85)	(0.86)	(0.78)	(0.87)	(0.86)	(0.83)
	No. of obs.	40	78	79	34	37	268	543
(d) Numeracy (out of 17)								
Wage	Mean	8.56	9.64	8.05	7.96	6.57	8.42	8.49
	Std.	(4.86)	(4.00)	(4.53)	(4.83)	(2.87)	(4.39)	(4.02)
	No. of obs.	30	57	55	32	21	195	291
Self	Mean	8.00	7.65	6.92	7.00	5.86	7.16	7.59
	Std.	(4.08)	(3.71)	(3.93)	(3.35)	(4.81)	(3.98)	(4.05)
	No. of obs.	40	78	79	34	37	268	543

Note: Monthly income is expressed in US dollars as of 2000.

Table 3.4: Main Estimation Results (OLS)

	Ghana						Tanzania									
	(a)		(b)		(c)		(d)		(e)		(f)					
	Sp. (3.10)	Self	Wage	Sp. (3.11)	Self	Wage	Sp. (3.12)	Self	Wage	Sp. (3.10)	Self	Wage	Sp. (3.11)	Self	Wage	Sp. (3.12)
β_1^j	-0.674 (0.443)	0.124 (0.330)	-0.767 (0.499)	0.128 (0.334)	0.199 (0.312)	-0.832** (0.499)	0.199 (0.312)	-0.336 (0.453)	-1.456** (0.583)	-1.283** (0.525)	-0.499 (0.390)	-1.307*** (0.429)	-0.530 (0.372)			
β_2^j	0.008 (0.009)	-0.001 (0.010)	0.015* (0.009)	-0.003 (0.010)	-0.005 (0.010)	0.014 (0.009)	-0.005 (0.010)	0.001 (0.012)	0.023* (0.014)	0.025* (0.014)	0.001 (0.011)	0.025** (0.010)	0.004 (0.010)			
β_3^j	0.057* (0.032)	-0.026 (0.022)	0.058* (0.034)	-0.013 (0.022)	-0.011 (0.021)	0.062* (0.036)	-0.011 (0.021)	0.030 (0.021)	0.070** (0.030)	0.061** (0.027)	0.038** (0.019)	0.069*** (0.024)	0.031* (0.018)			
β_4^j	-0.112* (0.065)	0.092 (0.056)	-0.112 (0.068)	0.039 (0.054)	0.029 (0.052)	-0.115 (0.075)	0.029 (0.052)	-0.050 (0.035)	-0.103** (0.058)	-0.091* (0.052)	-0.063* (0.032)	-0.112** (0.047)	-0.048 (0.032)			
$\beta_0^k - \beta_0^h$	-0.423 (0.698)	-	-0.409 (0.636)	-	-	-0.193** (0.076)	-	-	-0.907 (1.184)	-1.082 (1.230)	-	0.088 (0.107)	-			
R-squared	0.177		0.160		0.147		0.147		0.327	0.292		0.244				
No. of obs.	1794		1794		1794		1794		977	977		977				
	(g)		(h)		(i)		(j)		(k)		(l)					
	Sp. (3.14)	Self	Wage	Sp. (3.15)	Self	Wage	Sp. (3.16)	Self	Wage	Sp. (3.14)	Self	Wage	Sp. (3.15)	Self	Wage	Sp. (3.16)
	Wage	Sp. (3.14)	Self	Wage	Sp. (3.15)	Self	Wage	Sp. (3.16)	Self	Wage	Sp. (3.14)	Self	Wage	Sp. (3.15)	Self	Wage
γ_1^j	-0.432** (0.206)	0.030 (0.184)	-0.281 (0.181)	-0.001 (0.149)	0.020 (0.142)	-0.365** (0.182)	0.020 (0.142)	-0.394* (0.223)	-0.303* (0.168)	-0.440* (0.225)	-0.204 (0.156)	-0.466** (0.184)	-0.141 (0.156)			
γ_2^j	0.011** (0.005)	-0.001 (0.005)	0.012** (0.004)	-0.002 (0.005)	-0.001 (0.005)	0.013*** (0.005)	-0.001 (0.005)	0.008 (0.005)	0.003 (0.005)	0.010* (0.005)	0.002 (0.005)	0.012*** (0.004)	0.000 (0.005)			
$\gamma_0^k - \gamma_0^h$	-0.529 (0.649)	-	-0.609 (0.635)	-	-	-0.207*** (0.075)	-	-0.907 (1.191)	-	-1.064 (1.152)	-	0.076 (0.106)	-			
R-squared	0.171		0.157		0.144		0.144		0.305	0.291		0.241				
No. of obs.	1801		1801		1801		1801		990	990		990				

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) The full regression results are in Table 3.6 for (a), (b) and (c), Table 3.7 for (d), (e) and (f), Table 3.8 for (g), (h), (i), and Table 3.9 for (j), (k), and (l).

Table 3.5: The Estimated π_1 in the Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS)

	Ghana (a)	Tanzania (b)
Dependent variables:		
Log of earnings	-0.402 (0.243)	-0.032 (0.184)
Asset dummies		
Own saving	-0.119 (0.074)	-0.028 (0.128)
Bicycle	-0.062 (0.040)	-0.071 (0.060)
Fridge	-0.243** (0.104)	-0.259* (0.131)
Jewellery	0.000 (0.046)	0.058 (0.090)
Mobile phone	-0.184* (0.096)	-0.313*** (0.115)
Motor vehicle	-0.050 (0.041)	-0.073 (0.066)
Music system	-0.134 (0.083)	-0.031 (0.045)
Radio	0.081 (0.118)	-0.005 (0.160)
Telephone	-0.027 (0.031)	-0.035 (0.047)
Television	-0.210** (0.103)	-0.045 (0.136)
Video recorder	-0.218*** (0.066)	-0.018 (0.040)
Watch	-0.073 (0.108)	0.071 (0.154)
No. of observations	215	94

Notes: (1) Unit of observations is a person-year of recently arrived migrants (1992 to 2006 arrivals). (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) The full regression results are in Tables 3.10, 3.11, 3.12 and 3.13

Table 3.6: Appendix: Main Estimation Results (OLS) - Continuous Measure: Ghana

Dependent variable:	log of earnings					
	Ghana					
	(a)		(b)		(c)	
	Sp. (3.10)		Sp. (3.11)		Sp. (3.12)	
	Wage ×	Self ×	Wage ×	Self ×	Wage ×	Self ×
Migrants	-0.674 (0.443)	0.124 (0.330)	-0.767 (0.499)	0.128 (0.334)	-0.832** (0.499)	0.199 (0.312)
Age at migration (zero if native)	0.008 (0.009)	-0.001 (0.010)	0.015* (0.009)	-0.003 (0.010)	0.014 (0.009)	-0.005 (0.010)
Urban experience (years) (zero if native)	0.057* (0.032)	-0.026 (0.022)	0.058* (0.034)	-0.013 (0.022)	0.062* (0.036)	-0.011 (0.021)
Urban experience squared $\times 10^{-2}$ (zero if native)	-0.112* (0.065)	0.092 (0.056)	-0.112 (0.068)	0.039 (0.054)	-0.115 (0.075)	0.029 (0.052)
Age (years)	0.057** (0.022)	0.063** (0.024)	0.059*** (0.020)	0.057** (0.024)	-	-
Age squared $\times 10^{-2}$	-0.042 (0.026)	-0.055* (0.033)	-0.045* (0.024)	-0.047 (0.033)	-	-
Education (years)	-0.042 (0.067)	0.016 (0.038)	-0.046 (0.051)	0.006 (0.033)	-	-
Education squared $\times 10^{-2}$	0.601 (0.397)	-0.096 (0.376)	0.638** (0.318)	-0.052 (0.333)	-	-
Male dummy	0.236* (0.140)	0.427*** (0.095)	0.279** (0.123)	0.453*** (0.083)	-	-
Test scores						
Raven's test	-0.002 (0.021)	0.007 (0.011)	-0.005 (0.016)	0.011 (0.010)	-	-
Language comprehension	0.025 (0.023)	0.035** (0.018)	0.040* (0.020)	0.023 (0.015)	-	-
English reading test	0.058 (0.080)	-0.132** (0.018)	0.039 (0.063)	-0.071 (0.058)	-	-
Numeracy test	-0.009 (0.298)	-0.004 (0.017)	-0.015 (0.023)	0.004 (0.015)	-	-
Migrants × Demeaned variables						
× Education (years)	-0.026 (0.105)	-0.046 (0.077)	-	-	-	-
× Education squared $\times 10^{-2}$	0.323 (0.639)	0.322 (0.766)	-	-	-	-
× Male dummy	0.208 (0.300)	0.033 (0.178)	-	-	-	-
× Test scores						
Raven's test	-0.016 (0.025)	0.020 (0.021)	-	-	-	-
Language comprehension	0.049 (0.047)	-0.073* (0.037)	-	-	-	-
English reading test	-0.093 (0.142)	0.291** (0.143)	-	-	-	-
Numeracy test	-0.012 (0.042)	0.039 (0.036)	-	-	-	-
Wage employed dummy	-0.423 (0.698)	-	-0.409 (0.636)	-	-0.193** (0.076)	-
Constant		11.055*** (0.521)		11.057*** (0.515)		10.906*** (0.327)
Ethnic group dum. (EG)		No		No		Yes
EG × Wage/Self		Yes		Yes		No
EG × Wage/Self × Migrants		Yes		No		No
Year dum. (YD)		No		No		Yes
YD × Wage/Self		Yes		Yes		No
YD × Wage/Self × Migrants		Yes		No		No
Location dummies		Yes		Yes		Yes
R-squared		0.177		0.160		0.147
No. of observations		1794		1794		1794

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 9 categories. See footnote 11. (5) Location is classified into 5 categories (Accra, Ashanti, Central, Eastern, and Western).

Table 3.7: Appendix: Main Estimation Results (OLS) - Continuous Measure: Tanzania

Dependent variable:	log of earnings						
	Tanzania						
	(a)		(b)		(c)		
	Sp. (3.10)		Sp. (3.11)		Sp. (3.12)		
	Wage	Self	Wage	Self	Wage	Self	
	×	×	×	×	×	×	
Migrants	-1.456** (0.583)	-0.336 (0.453)	-1.283** (0.525)	-0.499 (0.390)	-1.307*** (0.429)	-0.530 (0.372)	-
Age at migration (zero if native)	0.023* (0.014)	0.001 (0.012)	0.025* (0.014)	0.001 (0.011)	0.025** (0.010)	0.004 (0.010)	-
Urban experience (years) (zero if native)	0.070** (0.030)	0.030 (0.021)	0.061** (0.027)	0.038** (0.019)	0.069*** (0.024)	0.031* (0.018)	-
Urban experience squared × 10 ⁻² (zero if native)	-0.103** (0.058)	-0.050 (0.035)	-0.091* (0.052)	-0.063* (0.032)	-0.112** (0.047)	-0.048 (0.032)	-
Age (years)	0.017 (0.028)	0.041* (0.024)	0.015 (0.029)	0.038* (0.023)	-	-	0.034* (0.018)
Age squared × 10 ⁻²	-0.017 (0.037)	-0.047* (0.028)	-0.014 (0.038)	-0.043 (0.027)	-	-	-0.036 (0.022)
Education (years)	0.131 (0.239)	-0.079 (0.053)	0.202 (0.243)	-0.080* (0.042)	-	-	-0.039 (0.057)
Education squared × 10 ⁻²	-0.342 (1.278)	0.620* (0.315)	-0.559 (1.253)	0.653*** (0.243)	-	-	0.573* (0.297)
Male dummy	0.272** (0.113)	0.285*** (0.102)	0.280*** (0.090)	0.335*** (0.083)	-	-	0.260*** (0.074)
Test scores							
Raven's test	-0.005 (0.022)	-0.033* (0.018)	0.002 (0.014)	-0.032** (0.015)	-	-	-0.009 (0.010)
Language comprehension	0.048** (0.022)	-0.017 (0.019)	0.018 (0.018)	-0.025 (0.016)	-	-	-0.008 (0.013)
English reading test	0.083 (0.075)	0.181*** (0.063)	0.113* (0.064)	0.157*** (0.053)	-	-	0.143*** (0.042)
Numeracy test	0.013 (0.031)	0.019 (0.018)	0.009 (0.027)	0.031* (0.017)	-	-	0.025 (0.016)
Migrants × Demeaned variables							
× Education (years)	0.235 (0.215)	-0.017 (0.100)	-	-	-	-	-
× Education squared × 10 ⁻²	-0.949 (1.066)	0.159 (0.563)	-	-	-	-	-
× Male dummy	0.033 (0.181)	0.121 (0.176)	-	-	-	-	-
× Test scores							
Raven's test	0.017 (0.028)	-0.006 (0.033)	-	-	-	-	-
Language comprehension	-0.077** (0.038)	-0.034 (0.035)	-	-	-	-	-
English reading test	0.175 (0.128)	-0.079 (0.133)	-	-	-	-	-
Numeracy test	-0.021 (0.040)	0.046 (0.039)	-	-	-	-	-
Wage employed dummy	-0.907 (1.184)	-	-1.082 (1.230)	-	0.088 (0.107)	-	-
Constant		10.022*** (0.581)		10.196*** (0.523)			9.581*** (0.430)
Ethnic group dum. (EG)		No		No			Yes
EG × Wage/Self		Yes		Yes			No
EG × Wage/Self × Migrants		Yes		No			No
Year dum. (YD)		No		No			Yes
YD × Wage/Self		Yes		Yes			No
YD × Wage/Self × Migrants		Yes		No			No
Location dummies		Yes		Yes			Yes
R-squared		0.327		0.292			0.244
No. of observations		977		977			977

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 23 categories. See footnote 11. (5) Location is classified into 7 categories (Arusha, Dar es Salaam, Iringa, Morongoro, Tanga, Mwanza, and other).

Table 3.8: Appendix: Main Estimation Results (OLS) - Discrete Measure: Ghana

Dependent variable:	log of earnings						
	Ghana						
	(a)		(b)		(c)		
	Sp. (3.14)		Sp. (3.15)		Sp. (3.16)		
	Wage	Self	Wage	Self	Wage	Self	
	×	×	×	×	×	×	
Recent arrivals (1992 to 2006 arrivals)	-0.432** (0.206)	0.030 (0.184)	-0.281 (0.181)	-0.001 (0.149)	-0.365** (0.182)	0.020 (0.142)	-
Age at migration (zero if native)	0.011** (0.005)	-0.001 (0.005)	0.012** (0.004)	-0.002 (0.005)	0.013*** (0.005)	-0.001 (0.005)	-
Age (years)	0.057*** (0.021)	0.048** (0.024)	0.060*** (0.022)	0.040** (0.024)	-	-	0.054*** (0.015)
Age squared $\times 10^{-2}$	-0.042 (0.025)	-0.036 (0.033)	-0.045* (0.026)	-0.037 (0.032)	-	-	-0.041** (0.019)
Education (years)	-0.049 (0.061)	0.009 (0.034)	-0.044 (0.051)	0.008 (0.003)	-	-	-0.043* (0.024)
Education squared $\times 10^{-2}$	0.637* (0.356)	-0.064 (0.342)	0.624* (0.317)	-0.056 (0.331)	-	-	0.512** (0.219)
Male dummy	0.237* (0.126)	0.458*** (0.088)	0.289** (0.124)	0.447*** (0.083)	-	-	0.374*** (0.068)
Test scores							
Raven's test	-0.004 (0.018)	0.006 (0.011)	-0.004 (0.016)	0.010 (0.010)	-	-	0.003 (0.009)
Language comprehension	0.028 (0.020)	0.027 (0.016)	0.039* (0.020)	0.025 (0.015)	-	-	0.029** (0.012)
English reading test	0.046 (0.070)	-0.063 (0.061)	0.044 (0.062)	-0.073 (0.058)	-	-	-0.028 (0.043)
Numeracy test	-0.011 (0.025)	-0.001 (0.016)	-0.015 (0.024)	0.003 (0.015)	-	-	-0.004 (0.013)
Recent arrivals $\times$ Demeaned variables							
$\times$ Education (years)	-0.037 (0.157)	-0.027 (0.125)	-	-	-	-	-
$\times$ Education squared $\times 10^{-2}$	0.453 (1.134)	0.368 (1.331)	-	-	-	-	-
$\times$ Male dummy	0.524 (0.481)	-0.200 (0.194)	-	-	-	-	-
$\times$ Test scores							
Raven's test	-0.005 (0.028)	0.049* (0.023)	-	-	-	-	-
Language comprehension	0.073 (0.067)	-0.052 (0.042)	-	-	-	-	-
English reading test	0.015 (0.170)	-0.128 (0.163)	-	-	-	-	-
Numeracy test	-0.005 (0.060)	0.063 (0.051)	-	-	-	-	-
Wage employed dummy	-0.529 (0.649)	-	-0.609 (0.635)	-	-0.207*** (0.075)	-	-
Constant		11.210*** (0.512)		11.201*** (0.507)			10.948*** (0.328)
Ethnic group dum. (EG)		No		No			No
EG $\times$ Wage/Self		Yes		Yes			Yes
EG $\times$ Wage/Self $\times$ Recent arrivals		Yes		No			No
Year dum. (YD)		No		No			Yes
YD $\times$ Wage/Self		Yes		Yes			No
YD $\times$ Wage/Self $\times$ Recent arrivals		Yes		No			No
Location dummies		Yes		Yes			Yes
R-squared		0.171		0.157			0.144
No. of observations		1801		1801			1801

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 9 categories. See footnote 11. (5) Location is classified into 5 categories (Accra, Ashanti, Central, Eastern, and Western).

Table 3.9: Appendix: Main Estimation Results (OLS) - Discrete Measure: Tanzania

Dependent variable:	log of earnings Tanzania						
	(a)		(b)		(c)		
	Sp. (3.14)		Sp. (3.15)		Sp. (3.16)		
	Wage x	Self x	Wage x	Self x	Wage x	Self x	
Recent arrivals (1992 to 2006 arrivals)	-0.394* (0.223)	-0.303* (0.168)	-0.440* (0.225)	-0.204 (0.156)	-0.466** (0.184)	-0.141 (0.156)	-
Age at migration (zero if native)	0.008 (0.005)	0.003 (0.005)	0.010* (0.005)	0.002 (0.005)	0.012*** (0.004)	0.000 (0.005)	-
Age (years)	0.004 (0.028)	0.043 (0.022)	0.014 (0.028)	0.043* (0.022)	-	-	0.038** (0.017)
Age squared $\times 10^{-2}$	0.001 (0.034)	-0.049* (0.026)	-0.010 (0.034)	-0.049* (0.026)	-	-	-0.039* (0.021)
Education (years)	0.202 (0.250)	-0.095** (0.047)	0.202 (0.240)	-0.084** (0.042)	-	-	-0.040 (0.058)
Education squared $\times 10^{-2}$	-0.570 (1.297)	0.725*** (0.273)	-0.570 (1.242)	0.668*** (0.243)	-	-	0.570* (0.304)
Male dummy	0.253*** (0.091)	0.353*** (0.095)	0.298*** (0.089)	0.340*** (0.085)	-	-	0.271*** (0.075)
Test scores							
Raven's test	0.000 (0.014)	-0.030* (0.016)	0.000 (0.014)	-0.033** (0.015)	-	-	-0.010 (0.010)
Language comprehension	0.035** (0.017)	-0.024 (0.017)	0.025 (0.017)	-0.024 (0.016)	-	-	-0.004 (0.013)
English reading test	0.098 (0.064)	0.173*** (0.058)	0.121* (0.061)	0.157*** (0.052)	-	-	0.146*** (0.042)
Numeracy test	0.000 (0.029)	0.024 (0.018)	0.007 (0.027)	0.030* (0.017)	-	-	0.023 (0.015)
Recent arrivals $\times$ Demeaned variables							
$\times$ Education (years)	0.079 (0.368)	0.089 (0.090)	-	-	-	-	-
$\times$ Education squared $\times 10^{-2}$	-0.401 (1.882)	-0.354 (0.543)	-	-	-	-	-
$\times$ Male dummy	0.236 (0.395)	-0.064 (0.228)	-	-	-	-	-
$\times$ Test scores							
Raven's test	0.007 (0.053)	-0.055 (0.047)	-	-	-	-	-
Language comprehension	-0.157** (0.075)	0.028 (0.031)	-	-	-	-	-
English reading test	0.094 (0.162)	-0.217* (0.123)	-	-	-	-	-
Numeracy test	0.107 (0.076)	0.055 (0.037)	-	-	-	-	-
Wage employed dummy	-0.907 (1.191)	-	-1.064 (1.152)	-	0.076 (0.106)	-	-
Constant		10.153*** (0.484)		10.090*** (0.476)			9.477*** (0.405)
Ethnic group dum. (EG)		No		No			Yes
EG $\times$ Wage/Self		Yes		Yes			No
EG $\times$ Wage/Self $\times$ Recent arrivals		Yes		No			No
Year dum. (YD)		No		No			Yes
YD $\times$ Wage/Self		Yes		Yes			No
YD $\times$ Wage/Self $\times$ Recent arrivals		Yes		No			No
Location dummies		Yes		Yes			Yes
R-squared		0.305		0.291			0.241
No. of observations		990		990			990

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 23 categories. See footnote 11. (5) Location is classified into 7 categories (Arusha, Dar es Salaam, Iringa, Morongoro, Tanga, Mwanza, and other).

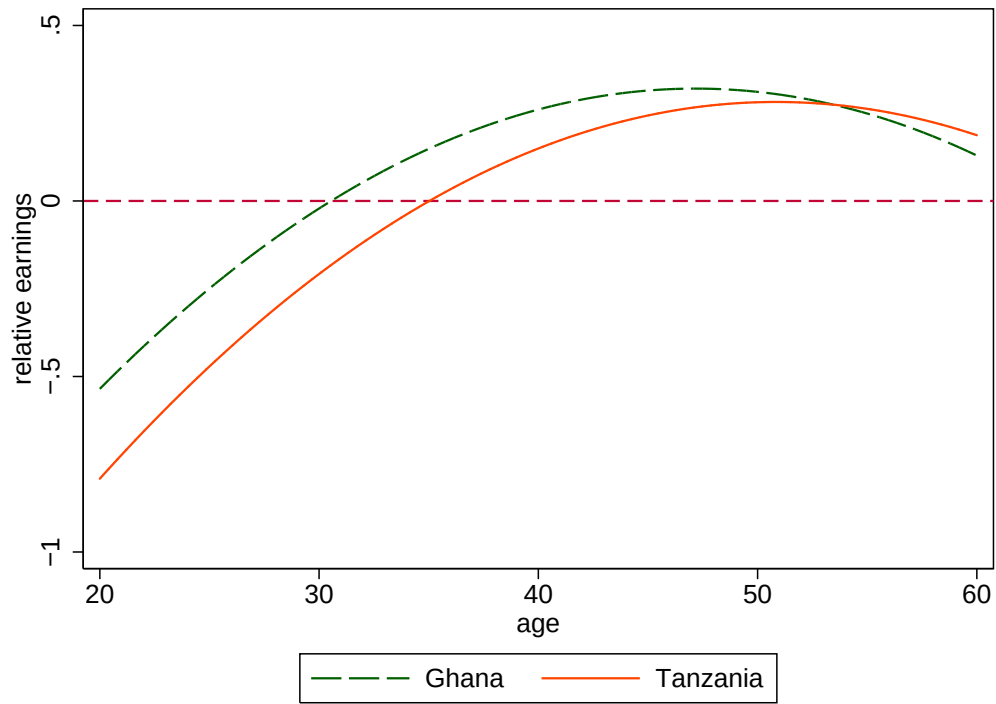


Figure 3.5: Appendix: Simulated Earnings of Wage-employed Migrants with an Arrival Age of 20 Relative to Otherwise Comparable Wage-employed Natives

Notes: (1) The relative earnings are the earnings of wage employed migrants whose arrival age is 20 relative to the earnings of wage-employed natives with similar observed characteristics (log of migrants' earnings less log of natives' earnings). (2) Simulations are based on estimates in columns (c) and (f) of Table 3.4.

Table 3.10: Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Ghana (1)

Dependent variables:	log of earnings	Asset dummies					
		Own saving	Bicycle	Fridge	Jewellery	Mobile phone	Motor vehicle
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
Wage employed	-0.402 (0.243)	-0.119 (0.074)	-0.062 (0.040)	-0.243** (0.104)	0.000 (0.046)	-0.184* (0.096)	-0.050 (0.041)
Age at migration	0.024 (0.040)	-0.005 (0.017)	0.001 (0.007)	-0.021 (0.017)	-0.004 (0.007)	-0.005 (0.017)	0.001 (0.011)
Age (years)	0.142** (0.063)	-0.023 (0.042)	-0.009 (0.019)	0.048 (0.033)	0.029* (0.015)	0.078*** (0.029)	0.005 (0.012)
Age squared $\times 10^{-2}$	-0.198** (0.090)	0.045 (0.067)	0.010 (0.025)	-0.006 (0.043)	-0.026 (0.023)	-0.092** (0.040)	-0.019 (0.015)
Education (years)	-0.001 (0.161)	0.005 (0.046)	0.012 (0.035)	0.014 (0.075)	0.012 (0.025)	0.004 (0.069)	0.006 (0.026)
Education squared $\times 10^{-2}$	0.122 (1.643)	-0.360 (0.401)	0.124 (0.372)	-0.334 (0.777)	-0.260 (0.279)	-0.049 (0.659)	-0.130 (0.251)
Male dummy	0.288 (0.191)	0.272** (0.110)	-0.002 (0.043)	0.061 (0.109)	-0.142*** (0.050)	0.010 (0.108)	-0.059 (0.047)
Test scores							
Raven's test	0.065*** (0.023)	0.022* (0.012)	-0.006 (0.004)	0.034*** (0.012)	0.016** (0.006)	0.017 (0.015)	0.001 (0.005)
Language comprehension	-0.029 (0.042)	-0.011 (0.022)	0.014 (0.008)	-0.013 (0.025)	-0.003 (0.010)	-0.003 (0.025)	-0.002 (0.011)
English reading	-0.127 (0.167)	0.061 (0.088)	-0.017 (0.036)	-0.002 (0.091)	0.019 (0.051)	-0.065 (0.084)	0.005 (0.031)
Numeracy test	0.040 (0.058)	-0.006 (0.020)	0.001 (0.009)	-0.004 (0.025)	0.003 (0.009)	0.008 (0.025)	0.021 (0.015)
Wage employed $\times$ Demeaned variables							
$\times$ Age at migration	-0.073 (0.072)	0.000 (0.023)	-0.005 (0.008)	0.003 (0.025)	0.001 (0.001)	-0.009 (0.025)	-0.015 (0.012)
$\times$ Age	-0.143 (0.118)	0.069 (0.063)	0.001 (0.008)	-0.074 (0.058)	-0.017 (0.025)	0.010 (0.060)	0.001 (0.016)
$\times$ Age squared $\times 10^{-2}$	0.308 (0.199)	-0.086 (0.098)	0.007 (0.034)	0.101 (0.083)	0.013 (0.040)	-0.013 (0.086)	0.012 (0.017)
$\times$ Education	-0.089 (0.242)	-0.000 (0.059)	-0.001 (0.036)	0.009 (0.088)	-0.045 (0.039)	0.081 (0.077)	-0.007 (0.026)
$\times$ Education squared $\times 10^{-2}$	1.017 (2.047)	0.485 (0.519)	-0.225 (0.386)	0.195 (0.868)	0.692* (0.396)	-0.635 (0.730)	0.160 (0.252)
$\times$ Male dummy	0.424 (0.478)	-0.164 (0.150)	0.020 (0.045)	-0.253 (0.164)	0.122 (0.086)	-0.118 (0.158)	0.069 (0.052)
$\times$ Test scores							
Raven's test	-0.065** (0.032)	-0.049*** (0.016)	0.003 (0.005)	-0.030 (0.020)	-0.001 (0.010)	-0.016 (0.019)	0.000 (0.005)
Language comprehension	0.128* (0.074)	0.017 (0.037)	-0.015 (0.011)	0.014 (0.033)	0.006 (0.013)	0.006 (0.036)	0.000 (0.011)
English reading	0.170 (0.267)	-0.075 (0.117)	0.006 (0.044)	-0.133 (0.129)	0.014 (0.074)	0.102 (0.119)	-0.029 (0.037)
Numeracy test	-0.057 (0.071)	0.040 (0.033)	0.012 (0.011)	0.034 (0.035)	-0.029* (0.015)	0.025 (0.036)	-0.021 (0.015)
Constant	9.691*** (1.508)	0.792 (0.867)	0.607 (0.461)	-0.433 (0.709)	-0.548 (0.372)	-1.206* (0.638)	-0.116 (0.265)
Ethnic group dum. (EG)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
EG $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dum. (YD)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YD $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.298	0.302	0.308	0.348	0.465	0.228	0.252
No. of obs.	215	215	215	215	215	215	215

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 9 categories. See footnote 11. (5) Location is classified into 5 categories (Accra, Ashanti, Central, Eastern, and Western).

Table 3.11: Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Ghana (2)

Dependent variables:	Asset dummies					
	Music System (h)	Radio (i)	Tele- phone (j)	Tele- vision (k)	Video recorder (l)	Watch (m)
Wage employed	-0.134 (0.083)	0.081 (0.118)	-0.027 (0.031)	-0.210** (0.103)	-0.218*** (0.066)	-0.073 (0.108)
Age at migration	0.019 (0.016)	0.002 (0.015)	-0.008 (0.009)	0.006 (0.016)	0.004 (0.014)	0.005 (0.014)
Age (years)	0.017 (0.029)	0.050* (0.030)	0.027* (0.014)	0.021 (0.032)	0.023 (0.031)	0.067** (0.032)
Age squared $\times 10^{-2}$	-0.018 (0.041)	-0.032 (0.040)	-0.028* (0.016)	-0.003 (0.039)	-0.024 (0.048)	-0.065 (0.040)
Education (years)	0.038 (0.051)	0.035 (0.071)	-0.022 (0.020)	0.024 (0.069)	-0.005 (0.040)	0.036 (0.055)
Education squared $\times 10^{-2}$	-0.741* (0.444)	-0.348 (0.729)	0.222 (0.158)	-0.298 (0.701)	-0.283 (0.321)	-0.532 (0.512)
Male dummy	0.154 (0.109)	0.061 (0.118)	-0.028 (0.026)	0.175 (0.112)	0.030 (0.089)	0.141 (0.109)
Test scores						
Raven's test	0.019* (0.011)	0.019 (0.012)	0.001 (0.003)	0.020* (0.011)	0.024** (0.011)	0.011 (0.109)
Language comprehension	-0.048** (0.021)	-0.030 (0.025)	0.000 (0.000)	-0.027 (0.027)	-0.027 (0.018)	-0.007 (0.027)
English reading	0.152* (0.088)	0.082 (0.091)	-0.022 (0.020)	0.057 (0.090)	0.024 (0.077)	0.009 (0.092)
Numeracy test	0.020 (0.020)	-0.019 (0.024)	-0.002 (0.007)	-0.001 (0.026)	0.035 (0.020)	-0.021 (0.026)
Wage employed $\times$ Demeaned variables						
$\times$ Age at migration	-0.025 (0.026)	-0.015 (0.025)	0.015 (0.010)	-0.038 (0.026)	0.001 (0.021)	-0.024 (0.024)
$\times$ Age	-0.001 (0.062)	-0.043 (0.052)	-0.056* (0.030)	-0.067 (0.060)	-0.078 (0.057)	-0.040 (0.057)
$\times$ Age squared $\times 10^{-2}$	0.010 (0.096)	0.077 (0.074)	0.071 (0.047)	0.135 (0.086)	0.105 (0.089)	0.078 (0.090)
$\times$ Education	-0.003 (0.062)	-0.026 (0.084)	0.043 (0.026)	0.037 (0.084)	0.040 (0.054)	0.029 (0.071)
$\times$ Education squared $\times 10^{-2}$	0.525 (0.536)	0.461 (0.799)	-0.380* (0.209)	-0.092 (0.801)	-0.086 (0.459)	0.152 (0.703)
$\times$ Male dummy	-0.093 (0.154)	-0.202 (0.187)	0.005 (0.045)	-0.234 (0.172)	0.011 (0.114)	-0.218 (0.184)
$\times$ Test scores						
Raven's test	-0.019 (0.017)	-0.037* (0.021)	-0.003 (0.004)	-0.028 (0.020)	-0.019 (0.016)	-0.021 (0.020)
Language comprehension	0.065** (0.028)	0.035 (0.030)	-0.005 (0.007)	0.047 (0.035)	0.044* (0.023)	-0.022 (0.033)
English reading	-0.209* (0.125)	-0.225 (0.137)	0.063 (0.043)	-0.145 (0.133)	-0.025 (0.112)	-0.073 (0.131)
Numeracy test	-0.003 (0.025)	0.047 (0.031)	0.007 (0.008)	0.036 (0.032)	-0.021 (0.026)	0.089*** (0.030)
Constant	-0.404 (0.556)	-0.980 (0.629)	-0.435* (0.221)	-0.490 (0.639)	-0.419 (0.565)	-0.779 (0.584)
Ethnic group dum. (EG)	Yes	Yes	Yes	Yes	Yes	Yes
EG $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes
Year dum. (YD)	Yes	Yes	Yes	Yes	Yes	Yes
YD $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.397	0.376	0.171	0.303	0.393	0.400
No. of obs.	215	215	215	215	215	215

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 9 categories. See footnote 11. (5) Location is classified into 5 categories (Accra, Ashanti, Central, Eastern, and Western).

Table 3.12: Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Tanzania (1)

Dependent variables:	log of earnings	Asset dummies					
		Own saving	Bicycle	Fridge	Jewellery	Mobile phone	Motor vehicle
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
Wage employed	-0.032 (0.184)	-0.028 (0.128)	-0.071 (0.060)	-0.259* (0.131)	0.058 (0.090)	-0.313*** (0.115)	-0.073 (0.066)
Age at migration	-0.011 (0.047)	0.032* (0.019)	-0.025 (0.018)	0.037 (0.030)	0.001 (0.018)	0.037 (0.023)	-0.005 (0.013)
Age (years)	0.318*** (0.107)	0.006 (0.005)	-0.037 (0.041)	-0.052 (0.057)	0.061 (0.049)	-0.001 (0.049)	0.022 (0.035)
Age squared $\times 10^{-2}$	-0.383** (0.149)	-0.026 (0.075)	0.060 (0.061)	0.026 (0.076)	-0.082 (0.068)	-0.042 (0.067)	-0.026 (0.041)
Education (years)	-0.128 (0.087)	0.009 (0.050)	-0.033 (0.049)	-0.002 (0.069)	0.015 (0.043)	-0.066 (0.058)	-0.024 (0.050)
Education squared $\times 10^{-2}$	0.741 (0.501)	0.162 (0.334)	0.218 (0.302)	0.190 (0.416)	0.201 (0.246)	0.652 (0.393)	0.126 (0.263)
Male dummy	0.358 (0.289)	0.231* (0.134)	0.180* (0.092)	0.267 (0.198)	0.403 (0.115)	0.254 (0.178)	0.137 (0.123)
Test scores							
Raven's test	-0.131** (0.060)	-0.028 (0.205)	0.030 (0.023)	0.002 (0.042)	-0.043** (0.021)	-0.013 (0.034)	-0.021 (0.027)
Language comprehension	0.002 (0.054)	-0.011 (0.019)	-0.013 (0.021)	-0.015 (0.038)	0.011 (0.022)	-0.021 (0.029)	0.006 (0.027)
English reading	0.197 (0.152)	0.047 (0.078)	-0.109* (0.064)	-0.037 (0.103)	-0.063 (0.090)	-0.137 (0.095)	0.014 (0.059)
Numeracy test	0.110** (0.053)	0.014 (0.024)	-0.007 (0.025)	0.002 (0.033)	-0.006 (0.028)	0.066** (0.029)	0.004 (0.021)
Wage employed $\times$ Demeaned variables							
$\times$ Age at migration	-0.011 (0.075)	0.003 (0.046)	0.027 (0.021)	-0.030 (0.054)	0.022 (0.032)	-0.022 (0.049)	0.016 (0.021)
$\times$ Age	-0.347 (0.286)	-0.263** (0.129)	0.008 (0.058)	-0.078 (0.128)	-0.253** (0.109)	-0.029 (0.091)	0.017 (0.076)
$\times$ Age squared $\times 10^{-2}$	0.506 (0.403)	0.331* (0.194)	-0.026 (0.078)	0.124 (0.191)	0.339** (0.148)	0.029 (0.130)	-0.051 (0.102)
$\times$ Education	0.666 (0.410)	0.174 (0.255)	0.012 (0.069)	0.324 (0.264)	0.157 (0.202)	0.237 (0.245)	0.197 (0.172)
$\times$ Education squared $\times 10^{-2}$	-3.222 (2.161)	-0.918 (1.408)	0.038 (0.411)	-1.584 (1.442)	-1.032 (1.047)	-1.228 (1.339)	-1.002 (0.897)
$\times$ Male dummy	0.171 (0.515)	0.073 (0.359)	-0.148 (0.126)	-0.189 (0.382)	0.086 (0.314)	-0.079 (0.333)	-0.111 (0.210)
$\times$ Test scores							
Raven's test	0.138 (0.087)	0.037 (0.046)	0.016 (0.025)	0.028 (0.044)	0.042 (0.038)	0.031 (0.044)	0.004 (0.024)
Language comprehension	-0.058 (0.111)	0.008 (0.065)	-0.046* (0.027)	0.015 (0.062)	0.003 (0.049)	0.024 (0.053)	0.009 (0.034)
English reading	0.180 (0.214)	0.067 (0.144)	-0.087 (0.077)	0.032 (0.139)	0.105 (0.135)	0.091 (0.146)	-0.009 (0.070)
Numeracy test	-0.107 (0.149)	-0.005 (0.073)	0.034 (0.036)	0.019 (0.080)	-0.010 (0.059)	-0.018 (0.061)	-0.010 (0.040)
Constant	5.051** (2.202)	-0.593 (1.150)	1.311 (0.924)	0.739 (1.322)	-1.191 (1.067)	-0.051 (1.106)	-0.258 (0.807)
Ethnic group dum. (EG)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
EG $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dum. (YD)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YD $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.725	0.551	0.705	0.561	0.617	0.614	0.333
No. of obs.	94	94	94	94	94	94	94

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 23 categories. See footnote 11. (5) Location is classified into 7 categories (Arusha, Dar es Salaam, Iringa, Morongoro, Tanga, Mwanza, and other).

Table 3.13: Appendix: Earnings and Asset Regressions of Recently Arrived Migrant Workers (OLS): Tanzania (2)

Dependent variables:	Asset dummies					
	Music System (h)	Radio (i)	Tele-phone (j)	Tele-vision (k)	Video recorder (l)	Watch (m)
Wage employed	-0.031 (0.045)	-0.005 (0.160)	-0.035 (0.047)	-0.045 (0.136)	-0.018 (0.040)	0.071 (0.154)
Age at migration	0.009 (0.008)	0.021 (0.034)	0.001 (0.004)	0.057** (0.025)	0.000 (0.003)	0.012 (0.031)
Age (years)	0.038 (0.032)	0.029 (0.063)	-0.002 (0.007)	0.108* (0.061)	0.000 (0.004)	-0.025 (0.052)
Age squared $\times 10^{-2}$	-0.057 (0.040)	-0.056 (0.094)	0.000 (0.011)	-0.203** (0.079)	-0.001 (0.007)	0.023 (0.069)
Education (years)	-0.010 (0.019)	-0.041 (0.075)	0.002 (0.009)	-0.066 (0.063)	0.001 (0.007)	0.042 (0.062)
Education squared $\times 10^{-2}$	0.049 (0.124)	0.400 (0.483)	-0.024 (0.050)	0.458 (0.439)	-0.022 (0.040)	-0.007 (0.403)
Male dummy	0.162 (0.112)	0.368** (0.174)	-0.003 (0.021)	0.483*** (0.142)	-0.004 (0.013)	0.295* (0.161)
Test scores						
Raven's test	-0.026 (0.021)	0.000 (0.047)	-0.004 (0.005)	-0.019 (0.032)	-0.003 (0.004)	-0.000 (0.039)
Language comprehension	0.006 (0.013)	-0.019 (0.037)	0.001 (0.005)	-0.007 (0.028)	0.002 (0.003)	-0.014 (0.032)
English reading	0.002 (0.035)	-0.092 (0.116)	0.004 (0.017)	0.048 (0.082)	0.007 (0.012)	-0.175* (0.102)
Numeracy test	0.011 (0.017)	0.015 (0.027)	-0.005 (0.007)	0.010 (0.027)	-0.005 (0.006)	0.015 (0.030)
Wage employed $\times$ Demeaned variables						
$\times$ Age at migration	0.071** (0.027)	0.001 (0.046)	0.036 (0.037)	-0.019 (0.045)	0.011 (0.016)	0.016 (0.044)
$\times$ Age	0.032 (0.075)	0.013 (0.182)	0.066 (0.071)	-0.362*** (0.123)	0.020 (0.061)	0.130 (0.191)
$\times$ Age squared $\times 10^{-2}$	-0.142 (0.103)	-0.039 (0.248)	-0.145 (0.115)	0.484** (0.172)	-0.049 (0.089)	-0.193 (0.257)
$\times$ Education	0.292* (0.172)	0.076 (0.338)	0.193 (0.261)	0.218 (0.231)	0.173 (0.174)	-0.040 (0.332)
$\times$ Education squared $\times 10^{-2}$	-1.428 (0.878)	-0.412 (1.764)	-0.895 (1.423)	-0.888 (1.306)	-0.882 (0.900)	0.103 (1.727)
$\times$ Male dummy	-0.248 (0.205)	-0.354 (0.431)	-0.129 (0.187)	-0.183 (0.339)	0.057 (0.107)	-0.180 (0.406)
$\times$ Test scores						
Raven's test	0.017 (0.029)	0.053 (0.068)	0.006 (0.028)	0.047 (0.039)	-0.011 (0.015)	0.023 (0.062)
Language comprehension	0.058* (0.033)	0.040 (0.087)	0.010 (0.044)	-0.004 (0.062)	0.013 (0.022)	0.019 (0.078)
English reading	-0.018 (0.080)	0.118 (0.232)	0.018 (0.075)	-0.084 (0.122)	-0.000 (0.045)	0.055 (0.216)
Numeracy test	-0.076* (0.041)	-0.114 (0.099)	-0.028 (0.041)	0.032 (0.071)	-0.003 (0.029)	-0.054 (0.101)
Constant	-0.749 (0.626)	-0.680 (1.369)	0.133 (0.302)	-2.336* (1.394)	-0.013 (0.091)	-0.044 (1.252)
Ethnic group dum. (EG)	Yes	Yes	Yes	Yes	Yes	Yes
EG $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes
Year dum. (YD)	Yes	Yes	Yes	Yes	Yes	Yes
YD $\times$ Wage	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.585	0.595	0.374	0.588	0.340	0.597
No. of obs.	94	94	94	94	94	94

Notes: (1) Unit of observations is a person-year. (2) Figures in () are standard errors. *** denotes at significance at 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across individuals and clustered residuals across years within each individual. (4) Ethnic group is classified into 23 categories. See footnote 11. (5) Location is classified into 7 categories (Arusha, Dar es Salaam, Iringa, Morongoro, Tanga, Mwanza, and other).

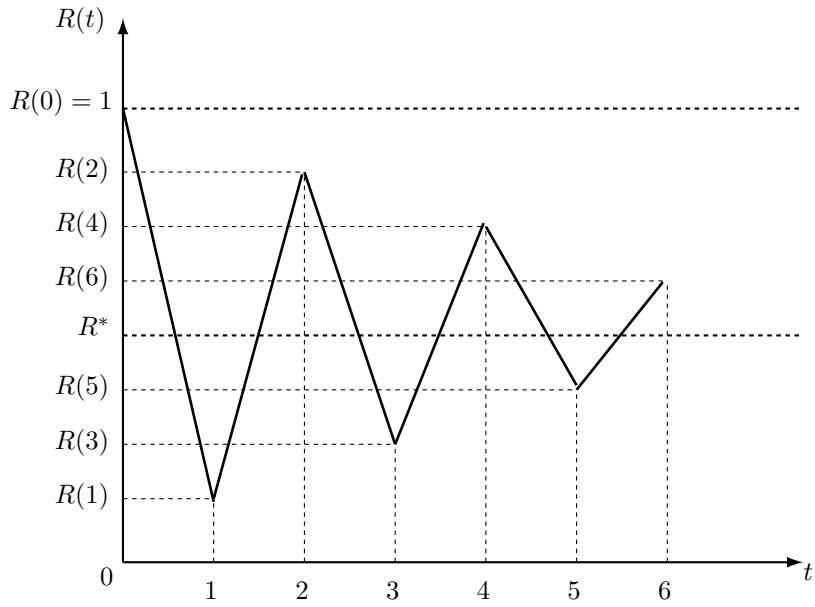


Figure 3.6: Appendix: Transition of Rural Population

Table 3.14: Appendix: Migrants' Urban Experience (years)

Ability group of migrants:	Period											
	1	2	3	4	5	6	7	8	9	10	11	12
$a_i \in$												
$[R(1), 1]$	1	-	-	-	-	-	-	-	-	-	-	-
$[R(2), 1]$	-	2	3	4	5	6	7	8	9	10	11	12
$[R(3), R(2)]$	-	-	1	-	-	-	-	-	-	-	-	-
$[R(4), R(2)]$	-	-	-	2	3	4	5	6	7	8	9	10
$[R(5), R(4)]$	-	-	-	-	1	-	-	-	-	-	-	-
$[R(6), R(4)]$	-	-	-	-	-	2	3	4	5	6	7	8
$[R(7), R(6)]$	-	-	-	-	-	-	1	-	-	-	-	-
$[R(8), R(6)]$	-	-	-	-	-	-	-	2	3	4	5	6
$[R(9), R(8)]$	-	-	-	-	-	-	-	-	1	-	-	-
$[R(10), R(8)]$	-	-	-	-	-	-	-	-	-	2	3	4
$[R(11), R(10)]$	-	-	-	-	-	-	-	-	-	-	1	-
$[R(12), R(10)]$	-	-	-	-	-	-	-	-	-	-	-	2

Chapter 4

Technology, Productivity and Size in African Manufacturing and Retailing Firms: Similar or Different?¹

Abstract

Using firm-level data of manufacturing and retailing drawn from the Enterprise Surveys conducted in seven Sub-Saharan African countries, this chapter attempts to improve our understanding of enterprise performance in urban Africa by investigating three aspects of firms' productive structure: technology, total factor productivity (TFP), and firm size. By estimating both Cobb-Douglas and translog production functions, we find that retailers can be modelled with similar production technology to manufacturers. Simple Cobb-Douglas function can adequately represents production technology in value added in both sectors. The homotheticity and constant returns to scale (CRS) are supported in both sectors. The homothetic technology implies that the difference in factor prices primarily explains firms' differing factor choices in each sector. This chapter parametrically estimates firms' TFP. We do not find any evidence to support that manufacturers have higher TFP than retailers, after controlling for the difference in technology between sectors. The similarity in the average level of TFP between sectors suggests that the heterogeneity within each sector, not across sectors, determines the aggregate level of TFP in an economy. Both TFP and labour productivity are distributed similarly between sectors. The distribution of technical efficiency is left-skewed in both sectors with great similarity. Given the similarity in the average level of TFP, the pattern of TFP distribution implies that a large mass of inefficient firms are prevailing in an overall economy. Size regressions find the size of retailers is, on average, half of manufacturers. The CRS production technology, however, rejects the efficiency advantage of manufacturers being larger than retailers. One implication from these findings is that the dispersion of factor prices may play a major role in characterising the industrial structure in urban Africa, not the difference in technology, TFP, and size across sectors.

4.1 Introduction

Urban industrial structure in Africa is characterised by relatively capital intensive manufacturing firms at one end, and labour intensive non-manufacturing firms at the other end. In Cameroon (2001), Mali (2004), and Zambia (2000), for example, value added (GDP) labour productivity in

¹ An earlier version of this chapter was presented at a seminar in Oxford. I am grateful to participants for helpful comments.

an industrial sector was higher than that in a service sector by 86%, 52% and 87%, respectively (Table 4.1).²

Using firm-level data of manufacturing and retailing from urban Africa, this chapter investigates firms' productive structure to improve our understanding of the sectoral difference in enterprise performance in urban Africa. The examination is important because the economic development is always associated with the expansion of high productive sectors and the contraction of low productive sectors. Investigating the difference in firms' productive structure across sectors would improve our understanding of how urban economies are evolving in Africa.

This chapter examines three key aspects: technology, total factor productivity (TFP), and firm size. Technology is important, because, if firms use homothetic technology for example, factor prices play a major role in driving their factor choices. We attempt to model firms' technology, by fitting both Cobb-Douglas and translog functional forms to data. This is the first attempt in modelling the production technology of retailing firms in the study of African firms. Whether retailers can be modelled with similar production technology to manufacturers is of great interest *per se*.

The TFP matters because it measures firms' underlying productivity that factor choices cannot explain. To the extent that the average level of TFP is similar between sectors, for instance, the heterogeneity in TFP within each sector, not across sectors, is likely to determine the aggregate level of TFP in an overall economy. Estimating TFP (or technical efficiency) has a long history in the literature (See Greene (1999) for an extensive summary). In this chapter, we parametrically estimate it by estimating a production function. We then compare the estimated TFP between sectors in terms of not only its average but also its distribution.

Firm size is another important dimension because it is related to many dimensions of firm performance, such as worker quality and factor prices. In addition, suppose firms enjoy economies of scale, there is an efficiency gain from being large. We compare the average firm size between sectors in this chapter.

The chapter exploits firm-level data drawn from the Enterprise Surveys conducted in seven

²These numbers are calculated by dividing (valued added GDP share/employment share) in an industrial sector by that in a service sector. We draw the GDP value added from the World Bank, World Development Indicator (<http://ddp-ext.worldbank.org/ext/DDPQQ/member.do?method=getMembers&userid=1&queryId=135>) and the employment number from International Labour Organisation (ILO), Key Indicators of the Labour Market (KILM), Sixth Edition.

Sub-Saharan African countries: Burkina Faso (2006, 2009), Cameroon (2006, 2009), Malawi (2005, 2009), Mali (2003, 2007), Niger (2005, 2009), Senegal (2003, 2007) and Zambia (2002, 2007). Since these surveys target the whole non-agricultural economy and exploit a standardised survey questionnaire with a uniform sampling methodology, these features enable us to investigate the performance of both manufacturing and retailing firms in a relatively easy and directly comparable way across countries. Another advantage of data arises from a maximum effort in the second round survey to recontact as many firms in the first round as possible, although the surveys are not intended to create complete panel data. Although rare, the effort was able to generate the panel components of manufacturers in each country's data. Pooling them across countries, we can construct panel data with a reasonable number of observations for our analysis of manufacturing firms.

For the countries of our interest, the service sector is the largest economy in terms of GDP contribution, and creates more jobs than the industrial sector does (Table 4.1). Despite the importance of the non-manufacturing sector in the economy, literature has disproportionally focused on the manufacturing sector (See Appendix 4.6 for a brief review on literature), leaving the performance of non-manufacturing firms in urban Africa unknown. By choosing one type of non-manufacturing business, retailing, and comparing it with manufacturing, this chapter opens the door to a new research of African firms.

The chapter proceeds as follows. Firstly, data are reviewed in Section 4.2. We then model the technology of both manufacturing and retailing firms in Section 4.3. Some of the key properties of production technology are tested in this section as well. Section 4.4 compares the mean level of TFP, firm size, and the pattern of productivity distribution between sectors. Section 4.5 provides our conclusion.

4.2 Data

4.2.1 Data Source and Structure

This chapter uses firm-level data from the Enterprise Surveys, which have been conducted in many countries for the last decade as part of the World Bank's initiative. For the analysis, we have constructed pooled data drawn from seven Sub-Saharan African countries, in which surveys were

conducted twice in the past: Burkina Faso (2006, 2009), Cameroon (2006, 2009), Malawi (2005, 2009), Mali (2003, 2007), Niger (2005, 2009), Senegal (2003, 2007) and Zambia (2002, 2007).

The surveys are designed to generate a sample representative of the whole non-agricultural private economy. A standardised survey questionnaire and a uniform sampling methodology allow us to directly compare data across countries and years, although they have some country-year specific characteristics. Sample firms are a stratified random sample by business sectors, firm size, and geographic locations and normally drawn from the list of eligible firms obtained from each country's statistical office. Depending upon the size of the economy (Gross National Income), the business sectors were stratified into several sectors, which usually comprise manufacturing including sub-sectors, retail, and other services. The level of firm size is stratified into small (5 to 19 of full-time permanent workers), medium (20 to 99) and large (100 or more), with micro firms (less than 5) in some countries. The overall sample size is determined so it is large enough for selected sectors to conduct statistically robust analyses with the level of precision at a minimum 7.5% in a 90 % confidence interval about the estimates of both the indicators of population proportion and the mean of log of sales at the sector level. The optimal size for selected sectors is, in principle, proportionally allocated into another level of strata (locations and size) according to the population distribution. However, the surveys tend to oversample large firms and this appears to be more noticeable in our data, because micro and small-sized firms are dominant in the class of economy we are investigating. For further details of sampling procedures, see *Enterprise Survey and Indicator Surveys, Sampling Methodology* (2009).³

While the original data contain the observations of non-retail non-manufacturers, we excluded them from the analysis because information on key variables required to model technology, such as capital, is not collected for those firms. Consequently, we exploited 2288 firm-year manufacturing and 829 retail observations pooled over countries and years in this chapter.⁴ However, as we will see soon, a large amount of missing or uncollected information on several variables substantially reduced the number of observations that could be used for a productivity analysis.

The second-round survey in each country extended a maximum effort to recontact as many firms in the first-round survey as possible, although the surveys are not intended to create complete panel

³https://www.enterprisesurveys.org/Documents/Sampling_Note.pdf .

⁴Retail observations do not contain wholesale firms.

data. The attempt generated 525 panel manufacturing firms pooled over sample countries. Due to the missing information problem, the number tenable for a productivity analysis is about 40%. In contrast to manufacturing, retailing observations are cross-sectional, drawn from the second-round survey in investigated countries.

Tables 4.2 and 4.3 summarise several characteristics in our data by sector and size. Firms are grouped into four categories (micro, small, medium and large) by using the same criteria as exploited when stratifying firms for a sampling purpose. While the manufacturing sector consists of several sub-industries, retailing is not. This is a potential concern in an empirical analysis because retailing firms normally vary in what and how they supply (McKinsey Global Institute, 1992; Baily and Solow, 2001; Baily and Zitzewitz, 2001). Retailers supply to customers a wide variety of goods, such as foods, clothing, consumer electronics, home appliances, home furnishings, etc. There are also many store formats from traditional mom-and-dad stores to mass market discounters and specialty chains, each of which offers different value propositions to customers. For example, mass merchandisers (e.g., Wal-Mart) compete by offering low prices, whereas in-town specialty chains (e.g., Gap, Benetton) and department stores produce intangible customer values by improving service quality, such as convenience and ambience. Therefore, our inability to make a distinction between retailers in these dimensions could make findings from the analysis less convincing. However, it appears that retailers in sample countries are quite different from modern retailing firms in advanced economies. Therefore, we do not expect this to be a serious concern in our analysis.

4.2.2 Measurement Issues

Following literature, output is measured by total annual sales. We use the number of full-time permanent employees as a measure of labour input. For the physical capital stock of manufacturers, we primarily exploited the net book value of machinery, vehicles, equipment, land, and buildings.⁵ Since the information is not available for retailers, total selling areas expressed by square meters are used as a measure of capital stock for them. Materials are raw materials and intermediate goods used in production for manufacturers, whereas finished goods and materials purchased to resell for

⁵In the second round survey in Malawi and Niger, information on the net book value was not collected. So, we used the replacement value in the second wave of these countries instead.

retailers. Indirect costs comprise the total annual costs of fuel and electricity in manufacturing. Only the latter is included in retailing.

All monetary values are firstly deflated into year 2000 prices by using either average consumer prices (for flow variables) or end-of-period consumer prices (for stock variables) by country.⁶ The deflated values are converted into US dollars (year 2000 prices) by using the annual average of daily market exchange rate (366 days) in each country.⁷ To all countries except Malawi and Zambia, the same exchange rate is applied because they are CFA franc countries.

There are a number of measurement concerns (see McKinsey Global Institute, 1992; Baily and Solow, 2001; Baily and Zitzewitz, 2001 for the extensive discussions of measuring service sector productivity). Firstly, retailing firms often rely on part-time workers, rather than full-time permanent employees, for supplying their daily services. Thus, using only the latter might fail to measure the true labour input. However, about 96% of 824 retailers in our data reported five or less as the number of part-time workers that they exploited in a surveyed year, and no part-time workers were employed in about 76%. It seems that this is not a crucial problem.

Another concern is the capital stock of retailing firms, which is usually the mix of real estate assets and information technology to control the operation. The real estate assets are to some degree embodied in the selling space of our capital measure, whereas the information technology is not. However, most retailing firms in our sample are not as modern as those in more advanced economies. It appears that the information technology does not matter very much for most retailing in our sample.

Price indexes and exchange rates, as they are not based on purchasing power parity, is a common measurement concern. This is because several factors, such as the change in tax policies and, in the case of exchange rate, capital flows, can easily distort the market-based variables.

Another problem is that input- and output-specific (producer) price indexes are not used in deflating monetary values. In the subsequent analysis, firms' TFP is estimated as regression residuals from an estimated production function. Then, if consumer prices grow faster than the input-specific (output-specific) price, this will result in systematically overestimating (underestimating) firms' underlying productivity.⁸

⁶International Monetary Fund, World Economic Outlook Database, April 2009 (<http://www.imf.org/external/pubs/ft/weo/2009/01/weodata/index.aspx>).

⁷OANDA (<http://www.oanda.com/currency/historical-rates>).

⁸More fundamentally, it is a question whether the residual based TFP can really measure firms' genuine pro-

Moreover, we are implicitly assuming that the level of inflation is the same between sectors by applying the same price indexes to both sectors. This is a separate issue from not using the input- and output-specific price deflators, whereby we are prone to overestimating (underestimating) TFP in a sector in which relevant sector-specific prices grow at the higher (lower) rate than consumer prices. If there is a serious gap in the level of the sector-specific price inflation between sectors, the inter-sectoral difference in our TFP estimates could be ascribed to the gap, rather than the genuine productivity difference.

Despite our expectation that consumer prices are a relatively appropriate price deflator (at least, for output) in retailing, using them in manufacturing can be a serious misleading factor in our attempt to compare TFP across sectors. Unfortunately, we cannot explicitly address this issue due to the lack of appropriate data. However, there was little inflation in CFA franc countries during the period of study relative to Malawi and Zambia.⁹ Therefore, if using consumer prices is indeed a problem, it is expected to be much more serious in Malawi and Zambia. As we will explain later, we conduct the same exercises as performed in Section 4.4, after dropping sample firms in Malawi and Zambia to check the sensitivity of our findings. We found that dropping them left implications from the main findings almost unchanged.

4.2.3 Descriptive Summary

With a few exceptions, Table 4.3 shows that labour productivity, factor inputs per employee, unit labour cost, firm age, export shares, and average schooling level of production workers monotonically rise over the size range in a manufacturing sector. These suggest that large firms are more labour productive, less labour intensive due to higher labour cost, and more likely to survive longer time and to exploit talented workers. The positive relationship between firm size and export shares is consistent with the well-studied dimensions of both self-selection and learning-by-exporting (Clerides et al., 1996, 1998; Delgado et al., 2002; Bigsten et al., 2004; Blalock and Gertler, 2004; Mengistae and Pattillo, 2004; Rankin et al., 2005; Van Biesebroeck 2005a), if large

ductivity, as economists usually exploit real sales revenue for output and real input expenditures for input, not the physical quantities of output and input. Katayama et al. (2009) formally argue that the resultant TFP might have little to do with firms' technical efficiency, as it captures several other effects, such as pass-through effects and mark-up effects.

⁹The annual growth rate of average consumer prices during the period of study was on average 3.0%, 13.6% and 25.1% for CFA franc countries, Malawi and Zambia, respectively (International Monetary Fund, World Economic Outlook Database, April 2009).

firms are more productive. Another point to note for manufacturers is that capital per worker has a larger dispersion than any other factor-labour ratios. This suggests either the substantial heterogeneity in the labour-to-capital cost within a manufacturing sector or large noise in our capital measure. Since collecting information on capital and indirect costs of micro firms is not satisfactory (only 11 and 12 observations in per-worker sense, respectively), the subsequent productivity analysis does not reflect the performance of micro manufacturers very much.

The monotonicity in retail is not as evident as in manufacturing. For example, the unit labour cost increases over the size range of micro to medium, but decline in large firms. In addition, the monotonic relationship cannot clearly be seen in factor-labour ratios. For instance, small firms use capital and materials per worker more than medium firms do. However, the violation, in part, appears to be attributed to the small number of observations in a particular size category. For micro and large firms, for example, we have only a few observations of capital and materials per worker. Accordingly, the subsequent productivity analysis primarily takes into account small and medium-sized firms in retailing.¹⁰ In contrast to manufacturing, materials per worker have the largest variance of factor-labour ratios in retail. This suggests large heterogeneity in material prices within a retailing sector.

We checked the equality of the mean between manufacturers and retailers by T-test for all variables except capital-labour ratio. The test assumed the unequal variance between two sectors and approximated the degree of freedom by Satterthwaite. With strong significance, all except material per worker differ between sectors at the mean level. While the difference is not as large as the macro-level difference between industry and services sectors shown in Table 4.1, manufacturers are more labour productive than retailers in both gross output and value added by 34% and 16%, respectively. Indirect costs per employee are larger in a manufacturing sector, but we expect that this is simply because of fuel costs, which are not included in retailers' indirect costs. Higher unit labour cost in a manufacturing sector implies that manufacturing workers are more skilful than retailing ones. Finally, manufacturers are on average much larger and have been operating businesses longer than retailers.

¹⁰The surveys collected information on export shares not only for manufacturers but also for retailers. However, what the export means in retailing businesses is seemingly unclear. As only 26 of 828 retailers reported non-zero export shares, we did not give careful consideration further to this point.

4.3 Modelling Technology

4.3.1 Model and Baseline Results

We begin with modelling firms' technology, which is the first perspective of an individual firm's productive structure, based on standard production theory. In the study of African manufacturing, Cobb-Douglas is an extensively used functional form (e.g., Biggs et al., 1995; Jones and Barr, 1996; Barr, 2000; Bigsten et al., 2000, 2004; Söderbom and Teal 2004; Van Biesebroeck, 2005a; Baptist and Teal, 2008), but estimating the production function of retailers is the first attempt that is often neglected in literature. In this study, we estimate not only the Cobb-Douglas function, but also a second-order transcendental logarithmic (translog) production function in both sectors, which is less restrictive than the Cobb-Douglas (Christensen et al., 1971, 1973). In the translog specification, we model the gross output Y of a firm i at time t in a sector $j \in (M, R)$, where M denotes manufacturing and R retail, as

$$\ln Y_{it}^j = a_i^j + \sum_q \beta_q^j \ln X_{qit} + \frac{1}{2} \sum_q \sum_s \beta_{qs}^j \ln X_{qit} \ln X_{sit} + \epsilon_{it}, \quad (4.1)$$

where $\beta_{qs}^j = \beta_{sq}^j$ for all $q \neq s$. The firm's TFP is expressed as a_i^j , X_q (X_s) is a factor input, which is either labour L , capital K , materials M or indirect costs C with notation q (s) $\in (l, k, m, c)$ corresponding to each factor input, and ϵ_{it} is a random error term with zero mean. While other covariates, such as firm age and dummies for country-region and country-year, are included in the actual estimation, we have ignored them for simplicity.

Following Söderbom and Teal (2004), we treat TFP as a firm-specific time-invariant effect by assuming that TFP is constant over the short time-span. Gross output is a natural measure of firm production, leaving value added essentially consistent with the behaviour of firms operating with constant return to scale (CRS) technology under perfectly competitive markets (Basu and Fernald, 1995, 1997). However, we also estimate the value added production function in the chapter, in addition to gross output. The translog specification is particularly useful in our attempt to estimate African retailers' technology for the first time because it allows us to test a range of interesting properties of production technology: (a) $\sum_s \beta_{qs}^j = 0$ for all q (homotheticity), (b) $\sum_s \beta_{qs}^j = 0$ for all q and $\sum_q \beta_q^j = 1$ (constant returns to scale (CRS)) and (c) $\beta_{qs}^j = 0$ for all q and s (Cobb-Douglas).

Throughout the analysis, we pool all firms drawn from seven different countries and put country-region dummies in all estimations. This procedure is based upon an implicit assumption that the sample firms are drawn from the same underlying population, after controlling for those dummies. In other words, those dummies are the only reason why firms differ across countries (and regions). Whilst this assumption had to be made to increase the number of observations, it is a question whether such an intercept can truly capture all (cultural, economic, political, and social) differences across seven different economies. However, excluding Cameroon (lower-middle-income economy), the other six countries were classified as low-income economies by criteria of the World Bank at the time of surveys.¹¹ In addition, Burkina Faso, Mali, Niger and Senegal are the member countries of the Economic Community of West African States (ECOWAS),¹² established in 1975 as a regional group of fifteen West African countries to promote economic integration in this region, and the West African Economic and Monetary Union (UEMOA), created in 1994 as a customs union and monetary union to encourage macro-economic convergence in this region. Moreover, Cameroon employs the Central African CFA franc as a national currency, which is effectively interchangeable with the West African CFA franc exploited in those four West African countries. Furthermore, as already mentioned in Section 4.2.2 and exactly shown in Appendix Section 4.9, the subsequent productivity analysis shows a number of findings almost insensitive to the inclusion of sample firms in non-CFA franc countries (i.e., Malawi and Zambia). All these characteristics might partly provide the justification of pooling sample firms across countries.

Table 4.4 reports the OLS estimation results of both translog and Cobb-Douglas production functions, based upon (4.1), by sector. The table presents the marginal effect of log of factor input on log of output, together with regression coefficients on other covariates (See Appendix Table 4.18 for full regression coefficients). While the marginal effect is exactly the estimated coefficient in the Cobb-Douglas function, it is evaluated at the sample mean of factor input in translog.¹³ There are several findings. Firstly, the marginal effect is similar between the translog and the Cobb-Douglas

¹¹<http://data.worldbank.org/about/country-classifications/>. Senegal is now classified as low-middle-income economy as of 2010.

¹²Niger was suspended from the membership of the ECOWAS in October, 2009 due to a political conflict. The second-round survey in Niger was conducted between May and October in 2009.

¹³The marginal effect of factor input $X_q \in (L, K, M, C)$ in translog is

$$\frac{\partial \ln Y_{it}^j}{\partial \ln X_{qit}} = \beta_q^j + \sum_{s \in (l, k, m, c)} \beta_{qs}^j \ln X_{sit}.$$

in both gross output and value added specifications in both sectors. This suggests that both the translog and the Cobb-Douglas fit the data in the same way around the mean of factor inputs. Secondly, it appears that retailing can be modelled with a standard production technology. In contrast to manufacturers, retailers are seemingly more material intensive, suggesting the importance of goods to sell in that sector. However, there is no remarkable difference in production function parameters between sectors. Finally, we tested some of key properties of production technology. Homotheticity and CRS are rejected in a manufacturing sector with an exception of homotheticity in column (c), whereas they are easily accepted in retail at any conventional significance level. Cobb-Douglas is rejected in all but value added in a retailing sector. However, these results are less reliable in the presence of the endogenous choice of factor inputs or measurement errors. In the next sub-sections, we carefully address these issues.

4.3.2 More on Panel Manufacturers

One of the most serious empirical concerns in estimating production functions is a simultaneity bias generated by the correlation between output determinants, observed by firms but not econometrician, and factor inputs. For example, top managers might alter factor inputs by responding to productivity shocks. In addition, it is certainly possible that some firm-specific unobserved characteristics, which may or may not be time-varying, simultaneously determine the level of factor inputs and the output. While OLS cannot address the problem, having panel data would enable us to find a better solution to it.

Before addressing the simultaneity bias, however, we firstly estimated the production function of panel manufacturers by OLS in Table 4.5. As in Table 4.4, the table reports the marginal effect, together with coefficients on other regressors (See Appendix Table 4.19 for full regression coefficients). The OLS estimation results of (4.1) are reported in columns (a), (b), (g) and (h). In contrast to the results in Table 4.4, the marginal effect of each factor input slightly changed, but the estimation results are not remarkably different between overall and panel observations. As already mentioned, the consistency of OLS estimates is doubtful. In multivariate regressions, it is generally impossible to draw inference about the direction of biases on the estimated coefficients. However, we expect that a capital coefficient is biased downwards, whereas coefficients on the other factor inputs are biased upwards as a rule of thumb. There are several reasons for this. For

example, the other factor inputs are more variable than capital stocks because managers are able to change the amount used for production year by year with greater ease. Then, it is possible that the difference in the magnitude of correlation with unobserved production disturbances between quasi-fixed capital stocks and the other factor inputs results in downward bias in a coefficient on the former and upward bias in coefficients on the latter (Levinsohn and Petrin, 2003). Attenuation bias arising from measurement errors is another explanation for the downward bias on a capital coefficient. Moreover, panel data usually contain firms that have survived during the period of study. If firms with larger capital stocks show greater propensity to continue in operation at the lower level of production efficiency, self-selection generated by firms' exit behaviour will lead to the negative bias in a capital coefficient (Olley and Pakes, 1996).

Within-group Estimation

Broadly speaking, two types of literature have separately been evolving to address the simultaneity bias, depending on underlying assumptions about unobservables that generate the problem (see Akerberg et al., 2006 for comprehensive discussions). One views time-invariant firm-specific fixed effects as a major source of bias. To see this clearly, let us re-write a production function as

$$\ln Y_{it}^j = a_i^j + \beta_j' \mathbf{z}_{it} + \epsilon_{it}, \quad (4.2)$$

where $\mathbf{z}_{it}$ is a vector of log of factor inputs, β_j is a vector of sector specific effects of $\mathbf{z}_{it}$, and ϵ_{it} is an error term such that $E(\epsilon_{it}) = 0$ for all t and $E(\epsilon_{is}\epsilon_{it}) = 0$ for all $s \neq t$ (no serial correlation). Assuming that the TFP can be decomposed into the sectoral mean (a_j) and time-invariant unobserved heterogeneity across firms within a sector (η_{ij}):

$$a_i^j = a_j + \eta_{ij}, \quad (4.3)$$

we can re-write (4.2) as

$$\ln Y_{it}^j = a_j + \beta_j' \mathbf{z}_{it} + \eta_{ij} + \epsilon_{it}. \quad (4.4)$$

The simultaneity bias arises from the potential correlation between the time-invariant firm-specific fixed effects (η_{ij}) and factor inputs ($\mathbf{z}_{it}$).

With another assumption: $E(\mathbf{z}_{is}\epsilon_{it}) = 0$ for all s and t (i.e., $\mathbf{z}_{it}$ is strictly exogenous with respect to the ϵ_{it} shocks), one of the simplest techniques to control the simultaneity problem is to subtract appropriate individual mean from variables:

$$\ln Y_{it}^j - \overline{\ln Y_i^j} = \beta_j'(\mathbf{z}_{it} - \bar{\mathbf{z}}_i) + (\epsilon_{it} - \bar{\epsilon}_i), \quad (4.5)$$

where $\overline{\ln Y_i^j} = \frac{1}{T} \sum_{h=1}^T \ln Y_{ih}^j$, $\bar{\mathbf{z}}_i = \frac{1}{T} \sum_{h=1}^T \mathbf{z}_{ih}$, and $\bar{\epsilon}_i = \frac{1}{T} \sum_{h=1}^T \epsilon_{ih}$. However, the literature recognises that this often yields unsatisfactory results, identifying a low and insignificant capital coefficient. The estimation results of (4.5) based upon two-period panel observations are reported in columns (c), (d), (i), and (j) in Table 4.5. Indeed, our within-group estimations traced the prediction, and the estimated marginal effect of capital stocks is even negative in columns (c) and (d). The most likely explanation for this unfavourable result is that while subtracting the mean can eliminate the fixed effects, it usually reduces the variability of covariates required to identify the coefficients of interest and, as a result, other specification errors, such as measurement errors in capital stocks, predominate (Griliches and Mairesse, 1998).¹⁴

GMM Estimation

Another common methodology to remove the fixed effects is to take the first-difference of equation (4.4):

$$\Delta \ln Y_{it}^j = \beta_j' \Delta \mathbf{z}_{it} + \Delta \epsilon_{it}, \quad (4.6)$$

where Δ denotes the difference between period t and $t - 1$. However, the strategy often generates the same problem (the reduced variation of covariates) as found in the within-group estimation and, moreover, it is possible that the first-difference of factor inputs ($\Delta \mathbf{z}_{it}$) has correlation with the first-difference of stochastic error ($\Delta \epsilon_{it}$), provided that $\mathbf{z}_{it}$ is contemporaneously correlated with ϵ_{it} (i.e., $E(\mathbf{z}_{it}\epsilon_{it}) \neq 0$), for example. Consequently, one would normally instrument the differenced

¹⁴Moreover, if true production functions are dynamic, the standard within-group estimators will be inconsistent (Nickell, 1981).

regressors by using the lagged levels with adequate assumptions. Given a new assumption that $E(\mathbf{z}_{is}\epsilon_{it}) = 0$ for all $s < t$ (i.e., $\mathbf{z}_{it}$ is endogenous with respect to the ϵ_{it} shocks), for example, $\mathbf{z}_{it-2}$ might be available as an instrument for $\Delta\mathbf{z}_{it}$ in equation (4.6), as long as $\mathbf{z}_{it}$ is serially correlated. This leads to the idea of the first-differenced GMM.

However, as factor inputs approach to follow random walks, the lagged levels tend to lose explanatory power for the current differences, making the first-differenced GMM estimates very imprecise. One way to solve this problem is to additionally use the lagged differences of factor inputs as instruments for the current levels. Again, to the extent that $E(\mathbf{z}_{is}\epsilon_{it}) = 0$ for all $s < t$, $\Delta\mathbf{z}_{it-1}$ can be used as an instrument for $\mathbf{z}_{it}$ in equation (4.4) with an assumption that $E(\Delta\mathbf{z}_{it-1}\eta_{ij}) = 0$ (e.g., the covariance between $\mathbf{z}_{it}$ and η_{ij} is constant over time). This is the so-called system GMM (Blundell and Bond, 1998, 2000; Bond, 2002).

Columns (e), (f), (k) and (l) in Table 4.5 report the results of the two-step system GMM based upon equations (4.4) and (4.6). Since the standard errors are often downward biased in small samples in the two-step GMM, we correct them by using the method proposed by Windmeijer (2005).¹⁵ Since our data consist of only two-period panel, we exploited $\Delta\mathbf{z}_{it}$ as an instrument for $\mathbf{z}_{it}$ in (4.4) and $\mathbf{z}_{it-1}$ as an instrument for $\Delta\mathbf{z}_{it}$ in (4.6) with the following moment conditions:¹⁶

$$E(\Delta\mathbf{z}_{it}(\eta_{ij} + \epsilon_{it})) = 0, \quad (4.7)$$

$$E(\mathbf{z}_{it-1}\Delta\epsilon_{it}) = 0, \quad (4.8)$$

where t and $t-1$ mean the first and second rounds in each country, respectively. For these moment conditions to be valid, we need the following assumptions: $E(\mathbf{z}_{is}\epsilon_{it}) = 0$ for all $s \leq t$ (i.e., $\mathbf{z}_{it}$ is predetermined) and $E(\mathbf{z}_{it}\eta_{ij}) = \tau_{ij}$ for all t (i.e., time-invariant covariance between $\mathbf{z}_{it}$ and η_{ij}). Since our panel data are short-run, the assumption of the constant covariance appears not absurd, whereas the first assumption is still a concern. In particular, it is certainly possible that $\mathbf{z}_{it}$ is contemporaneously correlated with the ϵ_{it} shocks, violating the exogeneity of instruments. The

¹⁵The GMM estimation is conducted by using the command "xtabond2" in Stata, which is a free contribution of Roodman (2006).

¹⁶In practice, the following adjustments are conducted for each estimation of the system GMM. Firstly, firm size three years ago drawn from a recall question is additionally included in instruments in (e) and (f). A dummy for whether a firm has its annual financial statements checked and certified by an external auditor in year t is additionally used in contemporaneous first-differenced equation in (e), (k) and (l). In addition, the number of power outages that a firm experienced in a typical month in year t is used in contemporaneous first-differenced equation in (e). Capital stocks and materials (inclusive of their interactions with both its own and the other factor inputs) are always used only in the first-differenced equation.

Hansen test, however, satisfied the overidentifying restriction in all GMM estimations (columns (e), (f), (k) and (l)), providing the justification of the moment conditions. Moreover, the instruments must be strong, as we are instrumenting $\mathbf{z}_{it}$ with $\Delta\mathbf{z}_{it}$ in (4.4) and $\Delta\mathbf{z}_{it}$ with $\mathbf{z}_{it-1}$ in (4.6).

In gross output specifications (e) and (f), a capital coefficient substantially increased and an indirect costs coefficient decreased from OLS. On the other hand, a labour coefficient substantially decreased from OLS in value added specifications (k) and (l), although all input coefficients disappointingly become insignificantly different from zero in (l). These corrections are exactly consistent with what we expected about the direction of potential biases in OLS estimates. The proportion of labour-to-capital coefficient is similar between gross output and value added specifications.

We again tested the properties of production technology. As opposed to the results in Table 4.4, homotheticity and CRS are now easily accepted in the system GMM. While Cobb-Douglas is still rejected in gross output specification,¹⁷ we can accept it in value added. As a whole, all the system GMM results appear to identify the technology parameters in a production function relatively well. In particular, the technology identified in our value added specification is remarkably similar to that of Ghanaian manufacturing firms identified by Söderbom and Teal (2004).

Levinsohn and Petrin Estimation

The second type of prevailing methodology attributes the source of the simultaneity bias to time-varying productivity shocks, which are observable or predictable by firms when they choose the current level of factor inputs. Taking into account firms' structural behaviour responding to the transmitted shocks, the methodology addresses the simultaneity through two steps;¹⁸ in the simplest two factor Cobb-Douglas production function of labour and capital, a labour coefficient is identified in the first step by putting in a proxy for transmitted productivity shocks in a production function. Assuming that such shocks evolve over time according to a first-order Markov process, the second step identifies a capital coefficient by using moment conditions that exploit

¹⁷Rejecting Cobb-Douglas does not necessarily mean that translog represents the underlying production technology more adequately. Since our translog estimation results accept homothetic technology, the key difference between translog and Cobb-Douglas is the elasticities of substitution between factor inputs and functional separability as a closely related concept. Cobb-Douglas is compatible with complete strong separability at every point in input space and its Allen partial elasticities are all one (Berndt and Christensen, 1973a, 1973b; Denny and Fuss, 1977). Then, the rejection of Cobb-Douglas (hence, complete strong separability) in gross output specification implies that grouping some of factor inputs (e.g., one index for materials and indirect costs) or further splitting them (e.g., skilled or unskilled workers for labour, and machinery, equipment, land, and buildings for capital) might improve the fitness to data. But, once we do so, Cobb-Douglas might in turn get accepted.

¹⁸If we are explicitly willing to control for firms' self-selection into exit, the approach needs three steps (Olley and Pakes, 1996).

an innovation component of shocks. Olley and Pakes's (1996) pioneering work uses the amount of investment as the proxy, whereas Levinsohn and Petrin's (2003) following study exploits an intermediate input, casting doubt on a strict monotonicity assumption regarding the investment. This is because the investment is often very lumpy due to its adjustment costs. While we followed Levinsohn and Petrin (2003) for Cobb-Douglas specification, the estimation results were rather unsuccessful. See Appendix 4.7 for a report on these estimation results.

4.3.3 More on Retailers

The same empirical concerns are applied to estimating production functions in a retailing sector. While we somehow attempt to control for the simultaneity in this sub-section, readers should recognise that just having single cross-sectional data disappointingly limits our ability to do so.

One important idea in the system GMM is that the lagged first-difference can be used as an instrument for the current level. As we have information on firm size at three points in time from a recall question (at present, three years ago, and at start-up), we exploited the information (past size growth) as an instrument to control for the simultaneity bias. In order to complement the instrument, we added on instruments the number of power outages that a firm experienced in a typical month in the current year and the average length of time (hours) in each power outage.

Table 4.6 presents a range of 2SLS estimation results of the Cobb-Douglas version of equation (4.1). Columns (a) and (f) provide the OLS estimation results in Table 4.4 for reference. We attempted several combinations of instrumented factor inputs and instruments. Capital stocks are always treated as an exogenous quasi-fixed factor due to the lack of strong instrument. Which factor inputs are instrumented and which variables are used as instruments are reported in the middle of the table. In gross output specification, controlling for the endogeneity reduced a coefficient on labour input from OLS, leaving coefficients on the other factor inputs almost unchanged. In columns (c) and (d), however, instruments are weak in explaining materials and indirect costs. In value added specification, there is no remarkable difference between OLS and 2SLS. This is consistent with either that the simultaneity bias is not serious or that instruments are also endogenous. In columns (b) and (g), however, Sargan test accepts the exclusion restriction of instruments. All estimation results support the CRS technology.

4.3.4 Summary of Technology

By estimating both Cobb-Douglas and translog production functions, this section showed that retailers could be modelled with similar production technology to manufacturers. While it appears that retailers are somewhat material intensive, suggesting the importance of goods to sell, there was no remarkable difference in technology between sectors. The results accepted homothetic and CRS technology in both sectors. In a value added production function, Cobb-Douglas adequately represented production technology in both sectors. The homotheticity implies that the difference in factor prices primarily explains the difference in firms' factor choices in each sector. The CRS suggests that there is no efficiency gain from being large in both sectors.

Finally, since our data accepted CRS in both sectors, we also estimated labour productivity after pre-imposing the CRS technology on a production function. Implication from the estimation results remained unchanged, providing further support for the CRS technology in both sectors. The results are described in Appendix 4.8.

4.4 Productivity and Size

4.4.1 Are Retailing Firms Less Productive?

Next, we investigate the other two aspects of firms' productive structure: TFP and firm size. Firstly, this sub-section asks the question of whether there is any difference in underlying productivity (TFP) between sectors. For this purpose, we pool manufacturers and retailers by using the production function (4.4) as

$$\ln Y_{it} = \text{constant} + (a_R - a_M)r_i + \beta'_M \mathbf{z}_{it} + (\beta'_R - \beta'_M) \cdot r_i(\mathbf{z}_{it} - E[\mathbf{z}_{it}]) + \eta_{ij} + \epsilon_{it}, \quad (4.9)$$

where r_i is a dummy which takes one if the firm i is a retailer. In practice, $E[\mathbf{z}_{it}]$, the population average of $\mathbf{z}_{it}$, is replaced with the sample mean. We expect that a coefficient on a retail dummy and those on interaction terms between the dummy and demeaned factor inputs capture the difference in the mean level of TFP and in technology between sectors, respectively. Note that a constant term of each firm in (4.9) is $a_M + (\beta'_R - \beta'_M)E[\mathbf{z}_{it}] \cdot r_i$. As long as technology is not markedly different across sectors, an estimated intercept in (4.9) should not be far away from an intercept

in separately estimated production function of manufacturers, which corresponds to columns (a) to (d) in Table 4.4 (in the absence of sub-industry dummies, strictly speaking).

The OLS estimation results of (4.9) are reported in Table 4.7. As before, we estimated both translog and Cobb-Douglas specifications. There are three points to note. Firstly, while not reported to save space, pooling regressions generate the marginal effect of factor inputs with great similarity to the marginal effect estimated by sector in Table 4.4 (See Appendix Table 4.20 for the marginal effect). Secondly, an intercept in each specification is not far away from a corresponding intercept in Table 4.4 (columns (a) to (d)), although somewhat large difference can be seen in value added specifications. These would relieve us from incorrectly pooling observations to identify the difference in TFP across sectors. Finally and most importantly, all but column (d) yield a coefficient on a retail dummy insignificantly different from zero, suggesting that once we control for the difference in technology, there is no difference in underlying productivity across sectors. Column (d) identifies a retail effect significantly different from zero, but it still rejects the TFP advantage of manufacturing firms. One normally tends to believe that manufacturers are more productive than retailers. Our estimations do not provide any support for this.

Jumping to the conclusion here is dangerous, because we do not control for unobserved difference between manufacturers and retailers.¹⁹ If there are any unobservables simultaneously determining both output and sectoral characteristics, our previous findings can be misleading. While we do not have instruments good enough to control for such a concern, our data give us some qualitative information on potential unobservables that we will explain below: (a) the characteristics of top manager and the largest owner, (b) the match quality between firms and workers, and (c) business environment.

Firstly, to the extent that manufacturing jobs are more skill-demanding, those who are working in that sector might be more talented. Then, a coefficient on a retail dummy in Table 4.7 might have been biased downwards. Since data do not contain information on employees' human capital, such as the average level of schooling, in a retailing sector, we cannot compare the worker quality between sectors. However, Tables 4.8 and 4.9 provide some information on a top manager and the

¹⁹Of course, there are more issues to consider. For example, the endogenous choice of factor inputs in a production function remains our concern. However, as our interest here is not in correctly estimating input coefficients but in identifying the difference in underlying productivity between manufacturers and retailers, it is a minor issue relative to the productivity difference. In addition, omitted variables bias is a general concern. However, it might be safe to say that we are relatively free from the problem due to the large number of R-squared.

largest owner. By looking at the proportion to the total number of observations in each sector, the tables examine several dimensions of the top management, such as gender, education and start-up reasons. Apart from the level of top managers' education, there is no remarkable difference between sectors. The proportion of a top manager holding graduate degree is somewhat larger in manufacturing than in retail, providing support for our view that manufacturing businesses require skills. Then, we additionally put in the dummies for a top manager's education in pooling regressions.²⁰ This treatment does not affect our previous findings. While we are not able to fully control for all possibilities from the aspect of the worker quality, the estimated coefficient on a retail dummy in Table 4.7 appears to be the lower bound, as long as the qualitative difference of workers between sectors generates a bias in the way that we expect. This confirms that it is certainly untrue that manufacturing firms have higher TFP than retailers.

The second concern is also related to the difference in skill requirement between sectors. Suppose that manufacturing is a more skill-demanding sector, the relatively small number of skilled workers in the economy of our interest might prevent manufacturing firms from finding good workers holding required skills. Consequently, the average quality of match between firms and workers might be lower in a manufacturing sector, generating an upward bias in a coefficient on a retail dummy in Table 4.7. As formally registered firms might be able to find better quality workers more easily, we firstly checked firms' formality at start-up in Table 4.10, and data show that there is no big difference across sectors. We then examined the extent to which the top management was feeling the quality of employed workers as a serious obstacle to their current operations. Approximately 20% of manufacturing firms are viewing inadequately educated workforce as a major or very severe obstacle, whereas only 11% of retailing firms view this issue. Accordingly, the proportion of firms feeling that it is not an obstacle is smaller in manufacturing than in retail by about 10%. These reveal a picture consistent with our expectation that firms and workers are better-matched in a retailing sector. Moreover, we compared the way firms usually find new employees between sectors, although the number of observations is somewhat small relative to the other dimensions because of the survey design. It seems that retailers largely rely on informal institutional mechanisms, such as family and friends, for seeking for workers, whereas manufacturers do more on formal institutional mechanisms, such as placement offices and public advertisement.

²⁰The results are not reported here due to space. They are available from the author if required.

If high quality workers are mainly supplied through the costly formal institutional mechanisms, the difference in the way of finding new employees provides some evidence that manufacturing is indeed skill-demanding and that manufacturers have more difficulty in finding talented workers.

Finally, Table 4.11 presents the proportion of the reported biggest business obstacle. If there is any systematic difference in business environment between sectors, it will help us to draw inference about the direction of potential bias in a coefficient on a retail dummy in pooling regressions. Both sectors draw a similar picture, showing access to finance, the supply of electricity, intense competition with informally operating establishments, and tax rates as major problems. No remarkable difference in self-recognised business environment between sectors provides support for pooling regression results in Table 4.7.

In summary, pooling OLS provided evidence to reject the TFP advantage of manufacturers. In particular, three pooling regressions out of four showed that there was no significant difference in underlying productivity between sectors, once we allowed the technology to differ across sectors. The similarity in the average level of TFP suggests that the inter-sectoral difference in TFP does not play a major role in explaining the aggregate level of TFP in an overall economy. The heterogeneity within each sector, not across sectors, primarily explains it. The difficulty in controlling for the unobserved difference between sectors, however, prevents us from having a definitive conclusion. Some qualitative information provided us with evidence to suggest that the quality of match between firms and workers was higher in a retailing sector. This means that the inter-sectoral difference in the match quality might have made a coefficient on a retail dummy biased upwards enough to eliminate the genuine TFP advantage of manufacturers. Nevertheless, our findings are intriguing because they cast doubt on unquestionably accepting a likely belief that manufacturing is more productive than non-manufacturing.

4.4.2 Are Retailing Firms Smaller?

In order to test if there is any difference in the average firm size between sectors, we model the size (L) of firm i at time t as

$$\ln L_{it} = \alpha_0 + \alpha_1 r_i + \alpha_2' \mathbf{v}_{it} + \alpha_3' \cdot r_i (\mathbf{v}_{it} - E[\mathbf{v}_{it}]) + v_{it}, \quad (4.10)$$

where $\mathbf{v}_{it}$ is a vector of regressors that were exploited in estimating production functions, other than factor inputs, and v_{it} is an error term with mean zero. As long as we are correctly pooling observations and estimating the model, we expect that α_1 captures the difference in average firm size between sectors.

Column (a) in Table 4.12 reports the OLS estimation result of (4.10) without interaction terms, and column (b) does with it. A coefficient on a retail dummy is negative with 1% significance, suggesting that the size of retailing firms is on average half of manufacturing firms, after controlling for observed characteristics such as firm age. As before, unobservables (in particular, the better quality of match between firms and workers in a retailing sector) are a potential concern in these OLS estimations. If the better match helps firms expand the size, however, it must have made a coefficient on a retail dummy biased upwards. Then, this provides further support for our finding. Another concern is sample selection. We realised that the estimations in Table 4.12 contained almost no micro manufacturers due to missing information on technology license and export shares. We decided to drop those variables so that the estimations could include micro-manufacturers. The estimation results were, however, almost unaffected by the treatment.²¹ Similarly, the estimations in Table 4.12 do not contain large retailers due to the sampling design. As long as the surveys are designed to construct a sample representative of non-agricultural private economy, however, it seems true that retailers are smaller than manufacturers on average.

Do manufacturing firms enjoy efficiency gain from being larger than retailers? The analysis in Section 4.3 accepted the CRS technology in both sectors. In this case, the answer appears to be no.

4.4.3 How are Productivities Distributed in Each Sector?

The preceding analysis provided evidence to suggest the primary importance of within-sector heterogeneity in the level of TFP for the aggregate level of TFP in an overall economy. Then, this sub-section investigates how productivities are distributed in each sector.

The analysis uses two productivity measures: labour productivity, measured by output per employee, and TFP. Estimating TFP has a long history in the literature. The literature has

²¹The coefficient on a retail dummy was -0.702 (std. 0.036) in a specification that has no interaction with a retail dummy and -0.767 (std. 0.040) in a specification with the interaction.

invented a wide range of methodologies, which can be categorised into several dimensions, such as parametric or non-parametric framework, deterministic or stochastic frontier, and econometric or mathematical programming approach. One of the oldest measures to make a comparison between firms is index numbers. For example, Caves et al. (1982a) proposed a technique to create the bilateral measure of input, output and productivity by using the deterministic structure of translog production. This allows for arbitrary returns to scale and does not require us to estimate the parameters of production structure. Using a slightly more restrictive framework, which imposes CRS on the translog structure, Caves et al. (1982b) extended the bilateral measure to multilateral indexes for input, output and productivity. While we have not made a distinction between the average frontier (an oxymoron, though) and the “best practice” frontier, TFP is often calculated as a measure of technical efficiency, which conventionally refers to the ratio of actual output to its potential (Farrell, 1957). Data envelope analysis (DEA), initiated by Charnes et al. (1978), is one way of measuring the efficiency. This relies on deterministic mathematical programming and does not parameterise production technology (see Seiford and Thrall (1990) for a recent overview, for example). On the other hand, a stochastic frontier approach fully parameterises a production function by assuming the error component composed of an (non-negative) inefficiency term and a stochastic error (e.g., Aigner et al., 1977; Meeusen and Broeck, 1977; Jondrow et al., 1982; Battese and Coelli, 1988) and usually estimates the technical efficiency by using a maximum likelihood technique.

Each methodology has its own pros and cons, and delineating them is outside the scope of this chapter (see Greene (1999) for elaborate discussions, for example). We are, however, unwilling to rely on the DEA and the stochastic frontier approach. A productivity estimate based on DEA is essentially sensitive to outliers because the technique compares each firm to all the other firms in constructing the productivity measure. Since the survey-oriented micro-level data often include lots of measurement errors, the technique does not attract us very much.²² On the other hand, the stochastic frontier approach has relatively much difficulty in controlling for the endogeneity of factor choices in estimating a production function, which makes the approach less attractive to us. In the next paragraph, we describe our strategy to measure TFP. Whatever methodology is selected,

²²Another reason for not using DEA is that the mathematical programming approach does not give us statistical properties about the estimated parameters.

however, it might be worth knowing the finding in Van Biesebroeck (2008), which investigated the extent to which the estimated TFP was sensitive to the choice of estimation methods. Using data drawn from manufacturing firms in Columbia and Zimbabwe, he showed that both nonparametric (index numbers, DEA) and parametric methods (stochastic frontier model, system GMM, Olley and Pakes, Levinsohn and Petrin) yielded surprisingly similar TFP estimates, although input coefficients in production functions were sensitive to the choice of methods.

In this chapter, we rely on parametric regressions for estimating TFP. Following Söderbom and Teal (2004), the chapter assumes that TFP can be measured as firm-specific fixed effects over the short time-span. For panel manufacturers, we extract the fixed effects by regressing residuals (plus a constant term) in columns (f) and (l) in Table 4.5 (system GMM, Cobb-Douglas) on firm dummies and use it as TFP. On the other hand, for retailers, we simply exploit residuals (plus a constant term) in columns (e) and (g) in Table 4.6 (2SLS, Cobb-Douglas) as TFP, because our data of retailers are not panel.²³ While we use these TFP estimates as our benchmark,²⁴ we checked the sensitivity to the choice of estimation methods in Tables 4.13 (manufacturers) and 4.14 (retailers), before moving ahead. Table 4.13 tests the equality of twelve demeaned TFP distributions, calculated based on a variety of estimation results in Table 4.5. Each letter in row and column stands for: "Manufacturing (M)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "System GMM (G)" or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth. FE means the extracted fixed effects. Our benchmark TFP estimates correspond to MGGC-FE and MVGC-FE. Similarly, the equality of six demeaned TFP distributions of retailers based on the estimation results in Table 4.6 is tested in Table 4.14, in which each letter in row and column stands for: "Retail (R)" in the first and "2SLS (2)" or "OLS (O)" in the third.²⁵ Our benchmark TFP estimates for retailers are RG2C and RV2C. In each table, the number in the bottom-left refers to the p-values of two-sample Wilcoxon rank-sum (Man-Whitney) test, whereas the number in the top-right refers to that of Kolmogorov-Smirnov test. Here, the null is that two independent demeaned TFP distributions are drawn from the same

²³According to the terminology of technical efficiency literature, it is possible to view our TFP measurement strategy as the mix of stochastic and deterministic frontier approaches. This is because, for manufacturers, we attempt to separate TFP from a random error component, such as transitory shocks and measurement errors, by treating TFP as time-invariant firm-specific effects (stochastic frontier). On the other hand, the deviation from a theoretical maximum is solely attributed to the inefficiency of each firm for retailers (deterministic frontier).

²⁴In the subsequent analysis, we dropped the five largest TFP firms in both sectors due to our concern about outliers.

²⁵The second and fourth letters are the same as in manufacturing.

distribution of population.

With a few exceptions, Kolmogorov-Smirnov test accepts the null in both sectors for any two TFP distributions within each output measure (gross output or value added), but not across output measures. On the other hand, Wilcoxon rank-sum (Man-Whitney) test accepts the equality for any two TFP distributions. Therefore, it appears that our TFP estimates are insensitive to the choice of estimation methods within each output measure.

Figures 4.1 and 4.2 show the empirical cumulative distribution functions of both our benchmark TFP and labour productivity, where the mean of productivity in each sector is normalised to zero. The figures reveal remarkable similarity about the pattern of productivity distributions across sectors. Table 4.15 formally tested the equality of these demeaned productivity distributions between sectors by using two-sample Wilcoxon rank-sum (Man-Whitney) test (bottom-left) and Kolmogorov-Smirnov test (top-right). With one exception of Kolmogorov-Smirnov test for TFP in gross output, all reported p-values accept the null that two independent groups of manufacturers and retailers are drawn from the same distribution of population, given the mean productivity being equal. Not only the average TFP but also the pattern of its dispersion is seemingly identical between sectors.

To interpret our TFP as a measure of technical efficiency, we constructed an efficiency score of a firm i in a sector j by rescaling TFP (let us denote it as θ) as

$$\text{Efficiency score} = \exp \left(\theta_i^j - \theta_{max}^j \right), \quad (4.11)$$

where θ_{max}^j is the sample maximum of estimated TFP in a sector j .²⁶ Here, we measure the technical efficiency by assuming that a firm with the largest TFP is achieving a technology frontier, which is common across firms. Since this index is bounded in the interval $(0, 1]$, which gets closer to one as the firm gets more efficient, we are able to directly compare the average technical efficiency between sectors. Figures 4.3 (gross output) and 4.4 (value added) show the histograms of technical efficiency based on our benchmark TFP estimates. The distribution of efficiency score is left-skewed with greater similarity between sectors. To the extent that there is indeed no remarkable difference in underlying productivity between sectors on average, this suggests the dominance

²⁶This is the same exercise as what Söderbom and Teal (2004) did to measure the technical efficiency of Ghanaian manufacturing firms.

of less productive firms in an overall economy. Table 4.16 reports the average efficiency scores in these figures, together with efficiency scores based on other TFP estimates. Our benchmark average scores based on a gross output production function are 0.303 for manufacturing and 0.207 for retailing, whereas 0.201 and 0.122 in value added, respectively. The lower efficiency score in retailing might result from the intrinsic nature of deterministic frontier approach. This is because the approach attributes all the deviation from the theoretical maximum, including transitory shocks and measurement errors, solely to the inefficiency of the firm.

However, the estimated efficiency scores appear to be somewhat smaller than those in preceding studies. For Ghanaian three manufacturing industries during 1972-1973, for example, Page Jr. (1980) estimated the average technical efficiency to be approximately 0.9 by OLS, although it seems that small non-random sample data and no consideration to bias correction in estimating a production function are problematic. Using well-known RPED survey data during 1992 to 1993, Biggs et al. (1995) took a stochastic frontier approach. The estimated average technical efficiency of manufacturers was 0.43 for Ghana, 0.49 for Kenya, and 0.49 for Zimbabwe. Lundvall and Battese (2000) also adopted a stochastic frontier approach and estimated the average efficiency score to be about 0.75 for Kenyan manufacturing firms by using the RPED data during 1993 to 1995. Moreover, extending the covered period of RPED data to 1991-1997, Söderbom and Teal (2004) estimated the average efficiency score as 0.53 for Ghanaian manufacturers. The differences in investigated countries, in the level of industry aggregation, and in the approach of measuring technical efficiency make it difficult for us to directly compare our efficiency scores to those in the literature. But, it seems that our estimates are far from accepting Tybout's (2000) summary of previous studies; the literature does not provide evidence to support that the average deviation from the efficiency frontier is larger in developing economies than in developed. However, the efficiency score is often sensitive to outliers in data. Accordingly, eliminating the top 3% of firms with large TFP in each sector raised the average efficiency score, based on gross output, up to 0.428 in manufacturing and 0.351 in retailing sectors, for example. Rather than showing the absolute number of average score, we emphasise that the technical efficiency, calculated by processing regression residuals, produced the pretty similar distribution pattern between sectors, providing evidence to suggest that less productive firms are prevailing in an overall economy.

Whilst we do not describe here to save space, we re-conducted the preceding productivity

analysis (i.e., comparing the average TFP and the pattern of productivity distribution between sectors), after dropping sample firms in Malawi and Zambia. As these two countries use different currency from the other countries, we checked the sensitivity of our findings to the exchange rate used in this study by dropping sample firms in these countries. In addition, as mentioned in Section 4.2.2, these two countries have experienced higher inflation than the other countries during the period of study. If using consumer prices as a price deflator was a misleading factor in the preceding analysis, the problem might have been more serious in sample firms in Malawi and Zambia. Dropping sample firms in these countries, however, left our findings almost unchanged, providing further support. Appendix 4.9 reports these results.

4.5 Conclusion

The previous studies of African firms have exclusively focused on manufacturing and, as such, African non-manufacturing has been treated as a “black box” in the literature. By picking up one type of non-manufacturing business, retailing, this chapter went into the unknowns of African non-manufacturing for the first time. Using firm-level data drawn from the Enterprise Surveys conducted in seven Sub-Saharan African countries, this chapter attempted to improve our understanding of the sectoral difference in enterprise performance in urban Africa by investigating three aspects of firms’ productive structure: technology, TFP, and firm size.

By estimating both Cobb-Douglas and translog production functions, we found that retailers could be modelled with similar production technology to manufacturers. While it seems that retailers are somewhat more material intensive than manufacturers, suggesting the strong importance of goods to sell in that sector, the analysis showed that there was no remarkable difference in technology between sectors. In both sectors, the translog production function was similar to the Cobb-Douglas production function at the mean level of factor inputs. In value added, the simple Cobb-Douglas adequately represented production technology in both sectors. We tested some of key properties of production technology. Our analysis provided evidence to support homothetic and CRS technology in both sectors. The homotheticity implies that the difference in factor prices primarily explains the firms’ differing factor choices in each sector.

This chapter has parametrically estimated firms’ TFP. In contrast to a likely belief that manu-

facturing is more productive than retail, we did not find any evidence to support the TFP advantage of manufacturing firms. The similarity in the average level of TFP between sectors suggests that the heterogeneity in the level of TFP within each sector is a major determinant of the aggregate level of TFP, not the heterogeneity across sectors. We found that both TFP and labour productivity were distributed in a highly similar way between sectors. The distribution of technical efficiency was left-skewed in both sectors with great similarity. Taking into account no remarkable difference in the average level of TFP between sectors, the pattern of TFP distribution suggests that a large mass of inefficient firms are prevailing in an overall economy.

The analysis provided evidence that the size of retailing firms was on average half of manufacturers. To the extent that firms enjoy economies of scale, this suggests that manufacturers can gain efficiency by being larger than retailers. However, the CRS production technology rejects manufacturers' efficiency advantage arising from the difference in firm size.

Our findings suggest that manufacturers and retailers are remarkably similar. Technology, the average level of TFP, and the distribution pattern of TFP and labour productivity are similar between sectors. Then, the similarity eventually points to an answer for the question of what primarily characterises the industrial structure in urban Africa. Evidence suggests the importance of factor prices, not the difference in technology, TFP, and firm size across sectors. This is a very surprising conclusion because it contradicts what one normally believes is important.

We conclude the chapter with a few words of caution. First of all, consumer prices were used to deflate the monetary values of both input and output in both sectors. Thus, we cannot completely rule out the possibility that regression residuals might be capturing something different from firms' genuine productivity (factor prices, mark-up, etc. See Katayama et al. (2009) for a formal discussion regarding the measurement issues of TFP based upon regression residuals, for example) and, therefore, that firms' TFP might not accurately be compared between sectors in this chapter. Secondly, we do not have information on the human capital of retailing workers (e.g., education). In addition, the analysis provided some qualitative evidence to support that firms and workers were better matched in retail than in manufacturing. Our inability to control for intangibles in these dimensions might have prevented us from finding the difference in underlying productivity between sectors. Finally, the productivity analysis did not rely much on micro and, in the case of retailing, large firms due to either missing information (e.g., unasked questions for micro

firms) or sampling design. Thus, our findings should be treated with caution, for generalisation.

4.6 Appendix: Literature Review on Productivity and Size in African Manufacturing

Productivity and size are the relatively well-studied dimensions of African manufacturing. Studying the choice of factor inputs (or technology) that determine labour productivity is a natural starting point of research, and collecting micro-level data has enabled researchers to provide insightful evidence that would not have been visible by a traditional macro- or sector-level analysis.²⁷ For example, Pack (1976) described, by visiting and interviewing Kenyan plants, that there existed considerable variation in the choice of technology between core and peripheral activities of production process.²⁸ At the peripheral activities, labour-intensive technology was initially adopted, and once it was selected, little scope for efficient substitution of capital for labour was left given the current volume of output (demand). In a similar vein, Pack (1982) argued that it was possible to simultaneously achieve the higher level of labour productivity and employment by efficiently substituting labour for capital in less developed countries. The recent study of Söderbom and Teal (2004) showed that the diversity of factor choices across firms in a Ghanaian manufacturing sector could be explained by not non-homothetic technology but the dispersion of factor prices, and that eliminating such factor price differentials could create employment by improving allocative efficiency in the market.

While there is no doubt that factor prices matter in explaining firms' factor choices (hence, labour productivity), the relatively inelastic firms' response to structural adjustment programmes, which had partly aimed at eliciting manufacturers' sizable investment by eliminating price distortions in a market, encouraged the examination of the determinants of labour productivity, operating other than through factor prices.²⁹ This has a link with a large body of subsequent research on total factor productivity (TFP) or technical efficiency. An attempt of measuring the average level

²⁷Panel data collected in Regional Program on Enterprise Development (RPED) surveys, which were conducted in the 1990's on the initiative of the World Bank, are among the most well-known micro-level data of African manufacturing firms.

²⁸He decomposed the operation of plants into five activities: material receiving, processing, material handling among processes, packaging and storage of finished products. The core activity refers to processing, whereas the peripheral activities do to the others.

²⁹The inelastic response is not unusual. For example, as firms' equipment is already installed and there is little possibility for factor substitution given the current level of output (demand), the change in relative factor prices is unlikely to have an effect on firms' factor choices and the resultant labour productivity. One motivation of an enormous amount of the subsequent literature that is investigating the determinants of TFP is that the improvement of firms' labour productivity can be achieved by directly raising their technical efficiency in the short-run, in addition to correcting price distortions in a market.

of technical efficiency, proposed by Farrell (1957), of African manufacturers is of a natural interest and has generated a number of studies as we saw in Section 4.4 (Page Jr., 1980; Biggs et al., 1995; Lundvall and Battese, 2000; Söderbom and Teal, 2004). It is not easy to compare the absolute level of efficiency across countries by looking at the calculated efficiency score because the reference best practice frontier is country-specific. However, Biggs et al. (1995) found the relatively low level of average efficiency by using micro-level data from Ghana, Kenya and Zimbabwe during 1992 to 1993, suggesting that the dispersion of technical efficiency is left-skewed in African manufacturing. On the other hand, extending the covered period to 1991-1997, Söderbom and Teal (2004) found that the average technical efficiency in a Ghanaian manufacturing sector, which was derived after addressing the endogeneity in estimating a production function more carefully than previous studies, was similar to that found in other developing economies (and slightly below that for OECD countries).

Given the large dispersion of TFP and manufacturers' inelastic response to price incentives, it is expected that lifting each firm's TFP to the most efficient level of domestic firms crucially contributes to the growth of labour productivity in the country. Pack (1976) argued that labour productivity could be increased without capital deepening due to considerable "disembodied" gains arising from the improvement of worker and organisation quality, which allows managers to choose more efficient technology. Promoting the diffusion of knowledge through technology licensing or technical aid from foreign manufacturers and a central consulting group of domestic technicians is one policy suggestion for Sub-Saharan Africa by Pack (1993). The extensive study of Biggs et al. (1995) explicitly related African manufacturers' technical efficiency to their "technological capabilities" and shed light on the importance of firms' learning mechanisms in improving their TFP. All these discussions have subsequently given rise to a large amount of rigorous empirical studies of TFP determinants: learning-by-doing (Jones and Barr, 1996), learning-by-exporting (Bigsten et al., 2004; Mengistae and Pattillo, 2004; Rankin et al., 2005; Van Biesebroeck 2005a), knowledge spillover (Barr, 2000), and agglomeration externality (Fafchamps and Hamine, 2004).

Firm size is another important dimension in the study of African manufacturing. The large body of studies in this aspect has partly been motivated by the growing share of micro and small enterprises in the economy and the public belief that such a small business sector is important in creating jobs, in particular, for unskilled workers. One likely consensus about the relationship

between firm size and growth is that the growth pattern is far from accepting Gibrat's law.³⁰ For instance, Mead and Liedholm (1998)'s study of micro and small enterprises in developing countries, including Botswana, Kenya, Malawi, Swaziland and Zimbabwe, showed that small (and young) firms grew faster. Similarly, using the retrospective data of manufacturing firms collected in Côte d'Ivoire during 1995 to 1996, Sleuwaegen and Goedhuys (2002) found the negative relationship between firm growth and size (and age). The relationship was also observed in McPherson (1996) examining five southern African countries (Botswana, Lesotho, South Africa, Swaziland, and Zimbabwe). On the other hand, exploiting data drawn from nine Sub-Saharan African countries, Van Biesebroeck (2005b) provided evidence that cast doubt on such a view and found that large firms grew more rapidly than small ones, conditional on previous performance and age effects.³¹

The question of which of large or small firms grow faster has a close link with how African manufacturing sector is evolving. The contribution of small firms to the process of industrial churning is quite different in Africa from the theoretical prediction, which suggests that newly entered small firms either grow or fail more rapidly than large firms, depending upon its efficiency (Jovanovic, 1982; Ericson and Pakes, 1995). Evidence from the studies of firm growth and survival suggests that small firms rarely, if ever, graduate into large scale operations, whereas large firms start large and keep their size, seemingly benefiting from being large in several dimensions, such as the access to credit³² (Frazer, 2005; Van Biesebroeck, 2005b; Sandefur, 2008). Söderbom et al. (2006) more explicitly focused on the relationship among size, productivity, and firm survival by using the data of manufacturing firms in Ghana, Kenya and Tanzania and found that while there existed selection on efficiency, it occurred among the relatively large firms but not among small firms.

³⁰According to Gibrat's law, firm growth and firm size are stochastically independent.

³¹There are also other studies investigating the determinants of firm growth, such as formality and location of enterprises, proprietors' gender, and entrepreneurs' ethnicity (McPherson, 1996; Ramachandran and Shah, 1999; Sleuwaegen and Goedhuys, 2002; Biggs and Shah, 2006)

³²The first econometric study of firm survival in Africa is McPherson (1995). Using data of micro and small enterprises in Swaziland, Botswana, Malawi and Zimbabwe, he found no strong correlation between firm survival and firm size (level) but found a positive relationship between firm survival and firm growth.

4.7 Appendix: Levinsohn and Petrin Estimations for Panel Manufacturers

In this section, we estimate the Cobb-Douglas production function of panel manufacturers by following Levinsohn and Petrin (2003) (L&P). The estimation results are reported in Appendix Table 4.21. In columns (a) to (d), materials are used as a proxy for productivity shocks in the first step, whereas indirect costs are used in columns (e) to (h). In columns (b), (d), (f) and (h), we pre-imposed CRS on an output production function to directly estimate labour productivity. The standard errors were estimated by bootstrapping with 50 replications.

In regressions based upon gross output, a coefficient on materials substantially declined in columns (a) and (b) from OLS in Table 4.5. Although a capital coefficient slightly increased in these columns, it is still very low. In columns (e) and (f), a capital coefficient declined from OLS. In regressions based upon value added, both labour and capital coefficients substantially decreased from OLS, and the CRS technology is strongly rejected with 1% significance. After estimating production functions, we checked the correlation between the estimated productivity shocks and an intermediate input (productivity proxy variables) by regressing the former on the latter, together with capital stocks, dummies for country-region, country-year and industries. Conditional on capital stocks (per worker), we found that the estimated productivity shocks were positively correlated with either materials (per worker) or indirect costs (per worker), depending upon which variables are used as a proxy. While marginal, this provides some support for a monotonicity assumption needed to justify the validity of those proxy variables.

Unfortunately, it seems that input coefficients are poorly identified in our L&P estimations. While we do not investigate why L&P is not working well in our data, one critique for L&P is a collinearity problem (Akerberg et al., 2006). The main purpose of the first step in L&P is to identify a coefficient on variable inputs which are not used as a proxy for productivity shocks. However, if these variables (labour input, for example) are also adjusted by top managers responding to productivity shocks, we cannot correctly identify any coefficients in the first step. Akerberg et al. (2006) seriously take the point to suggest an alternative approach, which builds upon the idea in Olley and Pakes (1996) and Levinsohn and Petrin (2003).

4.8 Appendix: Labour Productivity Specification

This section estimates a production function by pre-imposing CRS on output. Then, we model the labour productivity y (output per worker) of a firm i at time t as

$$\ln y_{it}^j = \alpha_i^j + \sum_q \beta_q^j \ln x_{qit} + \frac{1}{2} \sum_q \sum_s \beta_{qs}^j \ln x_{qit} \ln x_{sit} + \epsilon_{it}, \quad (4.12)$$

which can be derived by pre-imposing CRS on equation (4.1), and x_q (x_s) is a factor input per employee, which is either capital per worker k , material per worker m and indirect costs per worker c , with notation q (s) $\in (k, m, c)$ corresponding to each factor-labour ratio. The TFP and coefficients β in (4.12) correspond to those in (4.1).

The OLS estimation results of both translog and Cobb-Douglas based on (4.12) are presented in Appendix Table 4.22 by sector. The table reports the marginal effect of log of factor input per worker on log of labour productivity, together with regression coefficients on other covariates. While the marginal effect is exactly the estimated coefficient in Cobb-Douglas, it is evaluated at the sample mean of factor input in translog.³³ Each column in Appendix Table 4.22 corresponds to that in Table 4.4. We also estimated (4.12) for panel manufacturers, addressing the endogeneity of factor choice. Appendix Table 4.23 presents the results, in which each column corresponds to that in Table 4.5. All estimation results are marvelously similar between (4.1) and (4.12), despite that we pre-imposed CRS in the latter. This provides strong support for accepting the CRS technology in both sectors.

We also estimated the labour productivity version of equation (4.9) in Appendix Table 4.24. Each column in the table corresponds to that in Table 4.7. As in Table 4.7, we do not find any evidence to suggest the TFP advantage of manufacturing firms.

³³The marginal effect of factor input per worker $x_q \in (k, m, c)$ is

$$\frac{\partial \ln y_{it}^j}{\partial \ln x_{qit}} = \beta_q^j + \sum_{s \in (k, m, c)} \beta_{qs}^j \ln x_{sit}.$$

4.9 Appendix: Sensitivity of Productivity Analysis

Malawi and Zambia use different currency from the other five countries (i.e., CFA franc countries). In addition, these two countries have experienced higher inflation than the other countries during the period of study. Thus, the distinction in these dimensions between these two countries and the other countries might have generated the misleading results in the productivity analysis. In this section, we check the sensitivity of our findings by re-conducting the main productivity analysis, after dropping sample firms in these two countries.

Appendix Table 4.25 reports the OLS estimation results of (4.9). The marginal effect is presented in Appendix Table 4.26. The sensitivity of TFP estimates to the choice of estimation methods is checked in Appendix Tables 4.27 and 4.28. Appendix Figures 4.5 and 4.6 show the empirical cumulative distribution functions of benchmark TFP and labour productivity. Appendix Table 4.29 tests the equality of these demeaned productivity distributions between sectors. The histogram of benchmark technical efficiency is drawn in Appendix Figures 4.7 and 4.8. Appendix Table 4.30 reports the average efficiency scores across estimations. All the results are similar to the results presented in the main body of the chapter. Our findings on productivity do not appear to be sensitive to the inclusion of sample firms in Malawi and Zambia.

Table 4.1: Industrial Structure and Labour Productivity

Country group	Value added (% of GDP)				Employment (% of total employment)				Labour productivity (Services are 1.00)	
	Year	Agri-culture	Industry	Services	Year	Agri-culture	Industry	Services	Agri-culture	Industry
Sample countries										
Burkina Faso	2006	33.00	22.00	44.00	-	-	-	-	-	-
Cameroon	2001	22.00	33.00	45.00	2001	60.60	9.10	23.10	0.19	1.86
Malawi	2008	34.00	21.00	45.00	-	-	-	-	-	-
Mali	2004	36.00	24.00	40.00	2004	41.50	16.50	41.90	0.91	1.52
Niger	2003	40.00	17.00	43.00	-	-	-	-	-	-
Senegal	2008	15.00	23.00	62.00	-	-	-	-	-	-
Zambia	2000	22.00	25.00	52.00	2000	71.60	5.80	22.60	0.13	1.87
Other Sub-Saharan Africa										
Ethiopia	2005	47.00	13.00	40.00	2005	80.20	6.60	13.20	0.19	0.65
Ghana	1999	36.00	25.00	39.00	1999	55.00	14.00	31.00	0.52	1.42
Tanzania	2006	45.00	17.00	37.00	2006	74.60	5.00	20.30	0.33	1.87
Uganda	2003	26.00	24.00	49.00	2003	61.80	10.30	27.60	0.24	1.31
South Asia										
Bangladesh	2005	20.00	27.00	53.00	2005	48.10	14.50	37.40	0.29	1.31
Nepal	2001	40.00	22.00	38.00	2001	65.70	13.40	20.10	0.32	0.87
Pakistan	2005	35.00	27.00	51.00	2007	43.00	20.30	36.60	0.58	0.95
East Asia										
Cambodia	2001	36.00	23.00	40.00	2001	70.20	10.50	19.10	0.24	1.05
Indonesia	2007	14.00	47.00	39.00	2007	41.20	18.80	39.90	0.35	2.56
Philippines	2007	14.00	32.00	54.00	2007	36.10	15.10	48.80	0.35	1.92
Thailand	2007	12.00	46.00	43.00	2007	41.70	20.70	37.40	0.25	1.93

Sources: (1) Value added - World Bank, World Development Indicator (<http://ddp-ext.worldbank.org/ext/DDPQQ/member.do?method=getMembers&userId=135>), (2) Employment - International Labour Organization, Key Indicators of the Labour Market (KILM), Sixth Edition.

Notes: (1) Country group is by criteria of the World Bank - LIE and LMIE stand for low-income and lower-middle-income economies, respectively. (2) The labour productivity of sector $j \in$ (agriculture, industry) is computed as sector j 's (valued added share/employment share) divided by service sector's (valued added share/employment share).

Table 4.2: Sample Composition

	Manufacturing					Retail				
	Micro	Small	Medium	Large	Total	Micro	Small	Medium	Large	Total
(1) Country										
Burkina Faso	0.03	0.07	0.05	0.06	0.06	0.03	0.19	0.13	0.60	0.15
Cameroon	0.03	0.07	0.12	0.17	0.10	0.02	0.14	0.59	0.25	0.15
Malawi	0.35	0.28	0.13	0.05	0.20	-	-	-	-	-
Mali	0.00	0.03	0.15	0.19	0.10	0.33	0.25	0.01	-	0.24
Niger	0.04	0.10	0.06	0.02	0.07	-	-	-	-	-
Senegal	0.22	0.26	0.18	0.18	0.22	0.33	0.20	0.06	-	0.22
Zambia	0.29	0.14	0.27	0.30	0.22	0.28	0.20	0.18	0.15	0.22
No. of obs.	111	1060	719	398	2288	242	480	87	20	829
(2) Industry										
Chemicals and pharmaceuticals	0.03	0.04	0.10	0.08	0.06	-	-	-	-	-
Electronics	-	0.00	0.00	0.01	0.00	-	-	-	-	-
Foods	0.05	0.27	0.28	0.32	0.27	-	-	-	-	-
Garments	0.38	0.21	0.04	0.01	0.13	-	-	-	-	-
Leather	0.00	-	0.00	-	0.00	-	-	-	-	-
Metals and machinery	0.17	0.09	0.12	0.06	0.10	-	-	-	-	-
Non-metallic and plastic materials	0.00	0.02	0.06	0.08	0.04	-	-	-	-	-
Textiles	0.04	0.02	0.03	0.08	0.03	-	-	-	-	-
Wood and furniture	0.02	0.02	0.03	0.00	0.02	-	-	-	-	-
Other manufacturing	0.24	0.29	0.31	0.32	0.30	-	-	-	-	-
No. of obs.	111	1060	719	396	2286	-	-	-	-	-
(3) Technology licensed from a foreign-owned company										
No. of obs.	-	0.07	0.13	0.26	0.13	-	-	-	-	-
(4) Legal Status	22	984	693	390	2089	-	-	-	-	-
Shareholding company with shares	0.08	0.08	0.04	0.08	0.07	-	0.00	-	-	0.00
traded in the stock market										
Shareholding company with non-traded	0.12	0.22	0.47	0.55	0.35	0.08	0.17	0.31	0.47	0.16
shares or shares traded privately										
Sole proprietorship	0.69	0.57	0.23	0.11	0.39	0.87	0.74	0.49	0.15	0.74
Partnership	-	0.07	0.11	0.05	0.08	0.03	0.05	0.10	-	0.05
Limited partnership	0.10	0.02	0.07	0.12	0.05	0.00	0.00	0.04	0.15	0.01
Other	-	0.01	0.04	0.06	0.03	-	0.02	0.04	0.21	0.02
No. of obs.	110	1059	711	382	2262	242	479	87	19	827

Notes: (1) Unit of observations is a firm-year. (2) The size of the firm is classified by the number of full-time permanent employees, where a micro firm has less than 5, a small firm has 5 to 19, a medium firm has 20 to 99, and a large firm has 100 or more. (3) The number is the proportion to the total number of observations in each item and group.

Table 4.3: Descriptive Statistics

	Manufacturing					Retail				
	Micro	Small	Medium	Large	Total	Micro	Small	Medium	Large	Total
(1) Log of output per employee (US dollars as of 2000)										
Mean	7.54	8.38	9.25	9.53	8.81	7.98	8.46	8.55	9.54	8.51***
Std.	(1.15)	(1.36)	(1.26)	(1.27)	(1.42)	(1.05)	(1.42)	(1.48)	(1.07)	(1.28)
Median	7.34	8.17	9.18	9.57	8.76	7.96	8.26	8.42	9.82	8.38
No. of obs.	104	985	675	362	2126	241	471	85	19	816
(2) Log of value added per employee (US dollars as of 2000)										
Mean	7.81	7.65	8.40	8.60	8.05	7.11	7.88	7.96	9.40	7.90**
Std.	(1.77)	(1.27)	(1.20)	(1.32)	(1.33)	(1.72)	(1.23)	(1.31)	(1.22)	(1.26)
Median	8.01	7.38	8.40	8.62	7.90	7.44	7.74	7.87	9.75	7.77
No. of obs.	9	838	544	270	1661	8	407	60	7	482
(3) Log of capital per employee (US dollars as of 2000 for manufacturing and square meters of selling areas for retails)										
Mean	6.82	6.87	7.70	7.98	7.34	1.49	2.07	1.57	0.87	1.97
Std.	(2.00)	(1.81)	(1.72)	(1.86)	(1.85)	(0.94)	(1.03)	(1.31)	(2.28)	(1.13)
Median	6.91	6.96	7.87	8.28	7.46	1.20	2.06	1.51	1.38	1.98
No. of obs.	11	848	589	315	1763	13	453	77	12	555
(4) Log of materials per employee (US dollars as of 2000)										
Mean	6.44	7.24	8.23	8.53	7.73	5.86	7.77	7.32	8.35	7.69
Std.	(1.39)	(1.58)	(1.60)	(1.67)	(1.71)	(1.45)	(1.49)	(1.97)	(1.87)	(1.58)
Median	6.21	7.03	8.34	8.73	7.73	5.85	7.66	7.35	8.78	7.63
No. of obs.	97	904	615	324	1940	8	414	61	6	489
(5) Log of indirect costs per employee (US dollars as of 2000)										
Mean	5.16	5.00	5.76	6.12	5.44	4.11	3.94	3.93	3.84	4.00***
Std.	(0.93)	(1.45)	(1.29)	(1.50)	(1.48)	(1.01)	(1.14)	(1.31)	(1.40)	(1.13)
Median	5.25	4.92	5.90	6.31	5.49	4.06	3.80	3.84	3.51	3.89
No. of obs.	12	870	587	303	1772	216	443	71	12	742

Manufacturing						Retail				
	Micro	Small	Medium	Large	Total	Micro	Small	Medium	Large	Total
(6) Log of labour cost per employee (US dollars as of 2000)										
Mean	6.11	6.47	7.07	7.22	6.76	6.04	6.31	6.67	6.37	6.27***
Std.	(0.89)	(0.88)	(1.02)	(1.17)	(1.04)	(0.73)	(0.85)	(1.00)	(1.48)	(0.87)
Median	5.95	6.45	7.10	7.24	6.73	6.00	6.29	6.84	6.83	6.26
No. of obs.	100	967	645	337	2049	237	451	83	18	789
(7) Firm size (1 + full-time permanent employees)										
Mean	4.00	10.75	45.26	442.66	93.50	3.37	8.31	33.18	118.35	12.00***
Std.	(1.00)	(4.04)	(20.92)	(877.26)	(391.30)	(1.05)	(2.93)	(17.39)	(26.20)	(19.70)
Median	4.00	10.00	41.00	201.00	19.00	3.00	7.00	25.00	106.00	7.00
No. of obs.	111	1060	719	379	2269	242	480	87	19	828
(8) Firm age (years)										
Mean	11.12	12.51	18.84	24.35	16.57	8.54	12.66	16.71	25.12	12.10***
Std.	(9.51)	(9.22)	(15.32)	(17.57)	(13.92)	(5.73)	(9.72)	(9.97)	(23.72)	(9.72)
No. of obs.	107	1000	697	392	2196	242	478	84	16	820
(9) Government / State own (percentage)										
Mean	2.72	0.81	1.32	4.40	1.68	0.20	0.00	0.22	0.00	0.08***
Std.	(16.28)	(8.73)	(10.18)	(17.16)	(11.55)	(2.31)	(0.00)	(2.14)	(0.00)	(1.43)
No. of obs.	111	1058	719	396	2284	242	480	87	19	828
(10) Export share of total sales (percentage)										
Mean	4.56	5.06	10.64	25.94	10.31	1.07	0.60	0.45	3.68	0.79***
Std.	(18.25)	(16.69)	(25.46)	(34.68)	(24.68)	(8.33)	(4.03)	(3.37)	(11.16)	(5.81)
No. of obs.	109	1049	702	385	2245	242	480	87	19	828
(11) Average schooling level of production workers (years)										
Mean	6.69	6.37	8.27	9.04	7.30	-	-	-	-	-
Std.	(3.79)	(3.55)	(3.22)	(2.81)	(3.54)	-	-	-	-	-
No. of obs.	90	764	415	212	1481	-	-	-	-	-

Notes: (1) Unit of observations is a firm-year. (2) The size of the firm is classified by the number of full-time permanent employees, where a micro firm has less than 5, a small firm has 5 to 19, a medium firm has 20 to 99, and a large firm has 100 or more. (3) Monetary value is expressed in US dollars as of 2000. (4) T-tests on the equality of means between manufacturers and retailers are conducted except for log of capital per employee, assuming unequal variances. The degree of freedom is approximated by Satterthwaite. *** denotes the rejection of equality at significance 1%, ** at 5%, and * at 10%.

Table 4.4: Baseline Production Functions and Size Regressions (OLS)

Dependent variables:	Manufacturing				Retail			
	log of output		log of employees		log of value added		log of employees	
	Translog (a)	Cobb-Douglas (b)	Translog (c)	Cobb-Douglas (d)	Translog (e)	Cobb-Douglas (f)	Translog (g)	Cobb-Douglas (h)
Marginal effects								
Log of employees	0.257*** [0.000]	0.256*** [0.000]	0.919*** [0.000]	0.879*** [0.000]	-	0.177*** [0.019]	0.194*** [0.001]	0.944*** [0.000]
Log of capital	0.028*** [0.012]	0.028*** [0.015]	0.224*** [0.000]	0.217*** [0.000]	-	0.023 [0.389]	0.012 [0.601]	0.158*** [0.034]
Log of materials	0.611*** [0.000]	0.587*** [0.000]	-	-	-	0.718*** [0.000]	0.696*** [0.000]	-
Log of indirect costs	0.147*** [0.000]	0.157*** [0.000]	-	-	-	0.076*** [0.002]	0.105*** [0.000]	-
Regression coefficients								
Log of firm age	0.028 (0.017)	0.026 (0.019)	0.064* (0.037)	0.075** (0.037)	0.308*** (0.035)	0.036 (0.040)	0.050 (0.047)	0.297*** (0.079)
Technology license (dummy)	-0.007 (0.050)	0.008 (0.050)	0.223*** (0.095)	0.228*** (0.094)	0.423*** (0.082)	-	-	-
Export share of total sales (percentage)	0.000 (0.007)	0.000 (0.000)	0.004*** (0.001)	0.003** (0.001)	0.014*** (0.001)	0.006 (0.005)	0.004 (0.006)	0.027* (0.015)
Government / State own (percentage)	-0.000 (0.001)	-0.000 (0.001)	-0.005 (0.003)	-0.005 (0.003)	0.006** (0.002)	-	-	-
Constant	4.191*** (0.456)	3.181*** (0.256)	6.910*** (0.601)	6.277*** (0.351)	2.461*** (0.271)	5.767*** (1.260)	3.949*** (0.434)	8.228*** (0.658)
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	NA	NA	NA
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test [p-values]								
Homotheticity	[0.000]	-	[0.126]	-	-	[0.703]	-	[0.932]
CRS	[0.000]	[0.088]	[0.001]	[0.002]	-	[0.766]	-	[0.928]
Cobb-Douglas	[0.000]	-	[0.055]	-	-	[0.000]	-	[0.686]
R-squared	0.955	0.944	0.802	0.799	0.512	0.886	0.874	0.463
No. of obs	1342	1342	1320	1320	1759	476	476	477
								816

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) Marginal effects are evaluated at the sample mean of inputs in translog specifications, whereas they are exactly estimated coefficients in Cobb-Douglas specifications. (5) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status and industry dummies.

Table 4.5: Production Functions of Panel Manufacturers

Dependent variables:	Log of output						Log of value added					
	OLS			Within-Group			OLS			Within-Group		
	Trans- log (a)	Cobb- Douglas (b)	Trans- log (c)	Cobb- Douglas (d)	Trans- log (e)	Cobb- Douglas (f)	Trans- log (g)	Cobb- Douglas (h)	Trans- log (i)	Cobb- Douglas (j)	Trans- log (k)	Cobb- Douglas (l)
Marginal effects												
Log of employees	0.277*** [0.000]	0.240*** [0.000]	0.211 [0.153]	0.270** [0.038]	0.271* [0.050]	0.268 [0.236]	0.954*** [0.000]	0.933*** [0.000]	0.686*** [0.004]	0.566*** [0.008]	0.808*** [0.001]	0.760 [0.190]
Log of capital	0.015 [0.497]	0.016 [0.507]	-0.055 [0.324]	-0.015 [0.799]	0.074 [0.301]	0.093 [0.078]	0.240*** [0.000]	0.238*** [0.000]	0.055 [0.523]	0.046 [0.584]	0.189* [0.066]	0.245 [0.589]
Log of materials	0.597*** [0.000]	0.569*** [0.000]	0.550*** [0.000]	0.401*** [0.000]	0.616*** [0.000]	0.543*** [0.165]	-	-	-	-	-	-
Log of indirect costs	0.174*** [0.000]	0.217*** [0.000]	0.107*** [0.030]	0.164** [0.023]	0.091 [0.354]	0.130 [0.237]	-	-	-	-	-	-
Regression coefficients												
Log of firm age	0.081** (0.040)	0.060 (0.042)	0.002 (0.100)	0.056 (0.091)	-0.016 (0.140)	0.054 (0.130)	-0.007 (0.079)	-0.009 (0.080)	0.302* (0.156)	0.181 (0.139)	0.143 (0.171)	0.064 (0.125)
Technology license (dummy)	-0.050 (0.098)	-0.055 (0.096)	0.054 (0.186)	-0.001 (0.169)	-0.148 (0.145)	-0.195 (0.131)	0.121 (0.176)	0.118 (0.176)	0.172 (0.331)	0.039 (0.341)	-0.192 (0.292)	-0.150 (0.346)
Export share (percentage)	-0.000 (0.001)	0.000 (0.001)	0.004 (0.006)	0.002 (0.006)	0.005 (0.005)	0.002 (0.006)	0.003 (0.002)	0.003 (0.002)	0.022** (0.009)	0.014 (0.010)	0.006 (0.013)	0.001 (0.017)
Government / State own (percentage)	-0.001 (0.003)	-0.001 (0.003)	-0.013 (0.008)	-0.009 (0.008)	0.000 (0.006)	-0.000 (0.006)	-0.009 (0.009)	-0.009 (0.008)	-0.033*** (0.012)	-0.030** (0.012)	-0.020* (0.011)	-0.017 (0.016)
Constant	4.440*** (0.820)	3.177*** (0.367)	7.624*** (2.262)	5.746*** (1.246)	4.397 (3.209)	2.976* (1.626)	5.493*** (1.035)	6.290*** (0.526)	5.554*** (2.625)	9.531*** (1.448)	4.173 (3.172)	7.161** (3.427)
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test [p-values]												
Homotheticity	[0.107]	-	[0.246]	-	[0.714]	-	[0.346]	-	[0.009]	-	[0.131]	-
CRS	[0.136]	[0.167]	[0.227]	[0.201]	[0.813]	[0.732]	[0.006]	[0.002]	[0.000]	[0.098]	[0.191]	[0.985]
Cobb-Douglas	[0.000]	-	[0.000]	-	[0.000]	-	[0.545]	-	[0.017]	-	[0.160]	-
Hansen Test [p-values]	-	-	-	-	[0.356]	[0.579]	-	-	-	-	[0.891]	[0.769]
R-squared	0.930	0.913	-	-	-	-	0.732	0.731	-	-	-	-
No. of obs	487	487	487	487	487	487	476	476	476	476	476	476

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) Standard errors in the system GMM is corrected by the method proposed by Windmeijer (2005). (5) The system GMM uses the first-lag as instruments for the contemporaneous first-difference and the contemporaneous first-difference as instruments for the current level with its own adjustments across columns; firm size three years ago drawn from a recall question is additionally included in instruments in (e) and (f). A dummy for whether a firm has its annual financial statements checked and certified by an external auditor in year t is additionally used in contemporaneous first-differenced equation in (e), (k) and (l). In addition, the number of power outages that a firm experienced in a typical month in year t is used in contemporaneous first-differenced equation in (e). Capital stocks and materials (inclusive of their interactions with both its own and the other factor inputs) are always used only in the first-differenced equation. (6) Marginal effects are evaluated at the sample mean of inputs in translog specifications, whereas they are exactly estimated coefficients in Cobb-Douglas specifications. (7) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status and industry dummies.

Table 4.6: Production Functions of Retailers

Dependent variables:	log of output					log of value added	
	OLS (a)	2SLS (b)	2SLS (c)	2SLS (d)	2SLS (e)	OLS (f)	2SLS (g)
Log of employees	0.194*** (0.060)	0.137 (0.222)	0.122 (1.424)	0.135 (0.257)	0.140 (0.366)	0.915*** (0.116)	0.885** (0.407)
Log of capital	0.012 (0.024)	0.026 (0.046)	0.022 (0.344)	0.025 (0.088)	0.018 (0.056)	0.141** (0.069)	0.158 (0.161)
Log of materials	0.696*** (0.025)	0.693*** (0.022)	0.718 (2.486)	0.691*** (0.154)	0.722** (0.302)	-	-
Log of indirect costs	0.105*** (0.028)	0.119*** (0.043)	0.106 (1.228)	0.125 (0.538)	0.100 (0.563)	-	-
Log of firm age	0.050 (0.047)	0.049 (0.050)	0.044 (0.530)	0.049 (0.065)	0.044 (0.074)	0.296*** (0.079)	0.265** (0.107)
Export share of total sales (percentage)	0.004 (0.006)	-0.002 (0.006)	-0.002 (0.037)	-0.002 (0.009)	-0.002 (0.008)	0.027* (0.015)	0.015 (0.013)
Constant	3.949*** (0.434)	3.898*** (0.617)	3.827 (6.953)	3.351* (1.902)	3.823* (1.964)	8.228*** (0.658)	7.645*** (1.296)
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	NA	NA	NA	NA	NA	NA	NA
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrumented variables							
Log of employees	-	Yes	Yes	Yes	Yes	-	Yes
Log of capital	-	No	No	No	No	-	No
Log of materials	-	No	Yes	No	Yes	-	-
Log of indirect costs	-	No	No	Yes	Yes	-	-
Instruments							
$\Delta \ln L_t$	-	Yes	Yes	Yes	Yes	-	Yes
$\Delta \ln L_{t-1}$	-	No	No	No	Yes	-	No
Power outage (no.)	-	Yes	Yes	Yes	Yes	-	Yes
Power outage (average duration)	-	No	No	No	No	-	Yes
Joint significance of instruments in the 1st stage							
Log of employees							
F-statistic	-	15.64	16.21	15.23	20.98	-	10.95
p-values	-	[0.000]	[0.000]	[0.000]	[0.000]	-	[0.000]
Log of materials							
F-statistic	-	-	0.54	-	6.07	-	-
p-values	-	-	[0.584]	-	[0.000]	-	-
Log of indirect costs							
F-statistic	-	-	-	0.75	4.09	-	-
p-values	-	-	-	[0.473]	[0.007]	-	-
Sargan test [p-values]	-	[0.991]	-	-	-	-	[0.674]
F-test [p-values]: CRS	[0.871]	[0.874]	[0.955]	[0.925]	[0.932]	[0.631]	[0.890]
R-squared	0.874	-	-	-	-	0.463	-
No. of obs.	476	408	408	408	408	477	409

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms in OLS. (4) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status dummies. (5) $\Delta \ln L_t$ is the difference in firm size (1 + full-time permanent employees) between the present and three years ago and $\Delta \ln L_{t-1}$ is the difference between three years ago and start-up. (6) Power outage (no.) is the number of power outages that a firm experienced in a typical month in year t and power outage (average duration) is the average length of time (hours) in each power outage.

Table 4.7: TFP Comparison - Pooling Regressions (OLS)

Dependent variables:	Lof of output				log of value added			
	(a) Translog		(b) Cobb-Douglas		(c) Translog		(d) Cobb-Douglas	
	Retail		Retail		Retail		Retail	
	$\times$ Demeaned		$\times$ Demeaned		$\times$ Demeaned		$\times$ Demeaned	
Retail dummy	-0.438 (0.650)		0.160 (0.172)		-0.406 (1.330)		1.512*** (0.331)	
Log of employees ($\ln L$)	1.022*** (0.203)	-0.098 (0.855)	0.262*** (0.028)	-0.105 (0.069)	1.460*** (0.190)	-0.023 (0.611)	0.870*** (0.038)	0.100 (0.110)
Log of capital (manufacturing: $\ln K_1$)	-0.001 (0.074)		0.026** (0.011)		-0.076 (0.138)		0.230*** (0.018)	
Log of capital (retail: $\ln K_2$)	0.128 (0.188)		0.030 (0.022)		0.382** (0.167)		0.208*** (0.058)	
Log of materials ($\ln M$)	-0.055 (0.119)	0.289 (0.266)	0.577*** (0.026)	0.145*** (0.035)				
Log of indirect costs ($\ln C$)	0.307*** (0.079)	-0.286 (0.214)	0.161*** (0.018)	-0.062* (0.033)				
$\ln^2 L \times 0.5$	0.145*** (0.051)	0.011 (0.169)			-0.007 (0.051)	-0.010 (0.221)		
$\ln^2 K_1 \times 0.5$	0.007 (0.010)				0.045** (0.020)			
$\ln^2 K_2 \times 0.5$	-0.017 (0.019)				0.003 (0.053)			
$\ln^2 M \times 0.5$	0.183*** (0.024)	-0.105*** (0.034)						
$\ln^2 C \times 0.5$	0.071*** (0.016)	0.011 (0.037)						
$\ln L \times \ln K_1$	0.012 (0.020)				-0.050** (0.025)			
$\ln L \times \ln K_2$	0.016 (0.050)				-0.079 (0.101)			
$\ln L \times \ln M$	-0.136*** (0.022)	0.086* (0.047)						
$\ln L \times \ln C$	0.016 (0.023)	-0.122* (0.067)						
$\ln K_1 \times \ln M$	-0.016 (0.010)							
$\ln K_2 \times \ln M$	-0.004 (0.021)							
$\ln K_1 \times \ln C$	0.009 (0.010)							
$\ln K_2 \times \ln C$	-0.001 (0.022)							
$\ln M \times \ln C$	-0.084*** (0.015)	0.063** (0.025)						
Log of firm age	0.031** (0.015)		0.028 (0.017)		0.119*** (0.034)		0.124*** (0.033)	
Technology license (dummy)	-0.019 (0.050)		0.003 (0.050)		0.205** (0.094)		0.209** (0.093)	
Export share of total sales (percentage)	0.000 (0.000)		0.001 (0.000)		0.003** (0.001)		0.003** (0.001)	
Government / State own (percentage)	-0.001 (0.001)		-0.001 (0.001)		-0.008** (0.003)		-0.008** (0.003)	
Constant	4.547*** (0.429)		3.161*** (0.232)		6.029*** (0.546)		5.443*** (0.311)	
Legal status dum.	Yes		Yes		Yes		Yes	
Industry dum.	No		No		No		No	
Country-region dum.	Yes		Yes		Yes		Yes	
Country-year dum.	Yes		Yes		Yes		Yes	
F-test [p-values]								
Manufacturing								
Homotheticity	[0.000]		-		[0.103]		-	
CRS	[0.000]		[0.118]		[0.000]		[0.001]	
Cobb-Douglas	[0.000]		-		[0.025]		-	
Retail								
Homotheticity	[0.753]		-		[0.406]		-	
CRS	[0.860]		[0.864]		[0.159]		[0.076]	
Cobb-Douglas	[0.009]		-		[0.523]		-	
R-squared	0.946		0.936		0.748		0.745	
No. of obs	1818		1818		1797		1797	

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status dummies.

Table 4.8: Top Manager's Characteristics

	Manufac- turing	Retail
(1) Top manager's gender		
Female	0.10	0.16
No. of obs.	340	257
(2) Education of top manager		
No education	0.09	0.16
Primary School (whether complete or not)	0.15	0.21
Secondary School (whether complete or not)	0.13	0.16
Vocational Training	0.15	0.13
Some university training	0.11	0.08
Graduate degree	0.35	0.24
No. of obs.	2180	823

Notes: (1) Unit of observations is a firm-year. (2) The number is the proportion to the total number of observations in each item and sector.

Table 4.9: The Largest Owner's Characteristics

	Manufac- turing	Retail
(1) The reason for the largest owner's start-up or takeover		
To take advantage of a business opportunity	0.86	0.83
Jobs or opportunities were absent or not satisfactory	0.13	0.16
No. of obs.	335	212
(2) The way the largest owner acquired ownership		
Started the businesses on their or with partners	0.79	0.89
Purchased an existing business	0.12	0.05
Joined an existing family business	0.04	0.03
Joined an existing not family business	0.02	0.01
Other	0.00	-
No. of obs.	339	255
(3) The largest owner's involvement in decision making		
On his/her own	0.43	0.64
In consultation with other owners	0.19	0.18
Delegate to other partner owners	0.02	0.03
Appoint a manager who makes decision	0.13	0.05
A board of directors makes decision	0.18	0.08
Other	0.02	0.03
No. of obs.	339	255
(4) The largest owner's gender		
Female	0.20	0.16
No. of obs.	574	84
(5) The largest owner's nationality		
African	0.80	0.87
Indian	0.03	0.06
Lebanese / Middle Eastern	0.06	0.02
Other Asian	0.00	0.00
European / Caucasian	0.07	0.02
Other	0.00	0.00
No. of obs.	1573	827

Notes: (1) Unit of observations is a firm-year. (2) The number is the proportion to the total number of observations in each item and sector.

Table 4.10: Match Quality between Firms and Workers

	Manufac- turing	Retail
(1) Formally registered at start-up		
Yes	0.80	0.77
No. of obs.	1453	825
(2) Inadequately educated workforce is		
No obstacle	0.43	0.53
Minor obstacle	0.19	0.20
Moderate obstacle	0.15	0.12
Major obstacle	0.14	0.08
Very severe obstacle	0.05	0.03
No of obs.	2284	829
(3) How to usually find new employees		
Through family/friends	0.37	0.51
Public placement office	0.08	0.02
Private placement office	0.17	0.09
Public announcement / advertisement	0.16	0.11
School-related network	0.02	0.05
Other	0.17	0.19
No. of obs.	222	257

Notes: (1) Unit of observations is a firm-year. (2) The number is the proportion to the total number of observations in each item and sector.

Table 4.11: The Biggest Business Obstacle

	Manufac- turing	Retail
Access to finance	0.20	0.21
Access to land	0.05	0.07
Business licensing and permits	0.01	0.01
Corruption	0.04	0.03
Courts	0.00	0.00
Crime, theft and disorder	0.02	0.04
Customs and trade regulations	0.01	0.03
Electricity	0.21	0.14
Inadequately educated workforce	0.02	0.01
Labour regulations	0.00	0.00
Political instability	0.00	0.00
Practices of competitors in the informal sector	0.11	0.15
Tax administration	0.09	0.07
Tax rates	0.13	0.12
Transport	0.03	0.04
No. of obs.	1939	827

Notes: (1) Unit of observations is a firm-year. (2) The number is the proportion to the total number of observations in each sector.

Table 4.12: Size Comparison - Pooling Regressions (OLS)

Dependent variable:	log of firm size	
	(a)	(b)
Retail dummy	-0.684*** (0.037)	-0.785*** (0.051)
Log of firm age	0.264*** (0.027)	0.287*** (0.034)
Technology license (dummy)	0.489*** (0.083)	0.482*** (0.082)
Export share of total sales (percentage)	0.013*** (0.001)	0.014*** (0.001)
Government / State own (percentage)	0.007*** (0.002)	0.007*** (0.002)
Retail $\times$ Demeaned variables		
Log of firm age	-	-0.090* (0.048)
Technology license (dummy)	-	-
Export share of total sales (percentage)	-	-0.007 (0.004)
Government / State own (percentage)	-	-0.032** (0.013)
Constant	2.642*** (0.226)	2.571*** (0.231)
Legal status dummies	Yes	Yes
Industry dummies	No	No
Country-region dummies	Yes	Yes
Country-year dummies	Yes	Yes
R-squared	0.584	0.585
No. of obs.	2576	2576

Notes: (1) Unit of observations is a firm-year. (2) Figures in () are standard errors. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status dummies.

Table 4.13: Sensitivity Check of (Demeaned) TFP Distributions across Estimations - Manufacturing

	MGOT	MGOC	MGGT	MGGC	MGGT-FE	MGGC-FE	MVOT	MVOC	MVGT	MVGC	MVGT-FE	MVGC-FE
MGOT	-	0.860	0.107	0.159	0.297	0.099	0.000	0.000	0.000	0.000	0.000	0.000
MGOC	0.877	-	0.084	0.301	0.745	0.194	0.000	0.000	0.000	0.000	0.000	0.000
MGGT	0.458	0.616	-	0.411	0.607	0.283	0.003	0.002	0.000	0.000	0.001	0.001
MGGC	0.985	0.843	0.622	-	0.887	0.963	0.000	0.000	0.000	0.000	0.000	0.000
MGGT-FE	0.836	0.954	0.670	0.873	-	0.687	0.001	0.000	0.000	0.000	0.000	0.000
MGGC-FE	0.831	0.667	0.518	0.897	0.776	-	0.000	0.000	0.000	0.000	0.000	0.000
MVOT	0.999	0.972	0.887	0.986	0.975	0.992	-	1.000	0.188	0.369	0.783	0.725
MVOC	0.969	0.926	0.849	0.957	0.948	0.959	0.999	-	0.488	0.520	0.926	0.972
MVGT	0.635	0.629	0.475	0.542	0.567	0.544	0.616	0.629	-	0.808	0.989	0.546
MVGC	0.753	0.813	0.952	0.876	0.870	0.864	0.966	0.991	0.696	-	0.538	0.729
MVGT-FE	0.734	0.732	0.529	0.662	0.647	0.668	0.647	0.658	0.996	0.689	-	0.876
MVGC-FE	0.979	0.930	0.836	0.955	0.918	0.963	0.855	0.864	0.764	0.855	0.766	-

Notes: (1) P-values are reported. (2) The number in the bottom left refers to a two-sample Wilcoxon rank-sum (Man-Whitney) test and the number in the top right refers to a Kolmogorov-Smirnov test for the hypothesis that two independent samples are drawn from the same distribution of populations. (3) Each letter in row and column stands for: "Manufacturing (M)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "System GMM (G)" or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth. FE means the extracted fixed effects. For example, TFP estimates in MGGC-FE and MVGC-FE are calculated based upon regression results in columns (f) and (l) in Table 4.5, respectively.

Table 4.14: Sensitivity Check of (Demeaned) TFP Distributions across Estimations - Retail

	RGOT	RGOC	RG2C	RVOT	RVOC	RV2C
RGOT	-	0.414	0.714	0.000	0.000	0.000
RGOC	0.523	-	0.996	0.000	0.000	0.000
RG2C	0.607	0.912	-	0.000	0.000	0.000
RVOT	0.254	0.300	0.344	-	1.000	1.000
RVOC	0.243	0.291	0.324	0.995	1.000	1.000
RV2C	0.357	0.422	0.447	0.976	0.967	-

Notes: (1) P-values are reported. (2) The number in the bottom left refers to a two-sample Wilcoxon rank-sum (Man-Whitney) test and the number in the top right refers to a Kolmogorov-Smirnov test for the hypothesis that two independent samples are drawn from the same distribution of populations. (3) Each letter in row and column stands for: "Retail (R)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "2SLS (2)" or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth. For example, TFP estimates in RG2C and RV2C are calculated based upon regression results in columns (e) and (g) in Table 4.6, respectively.

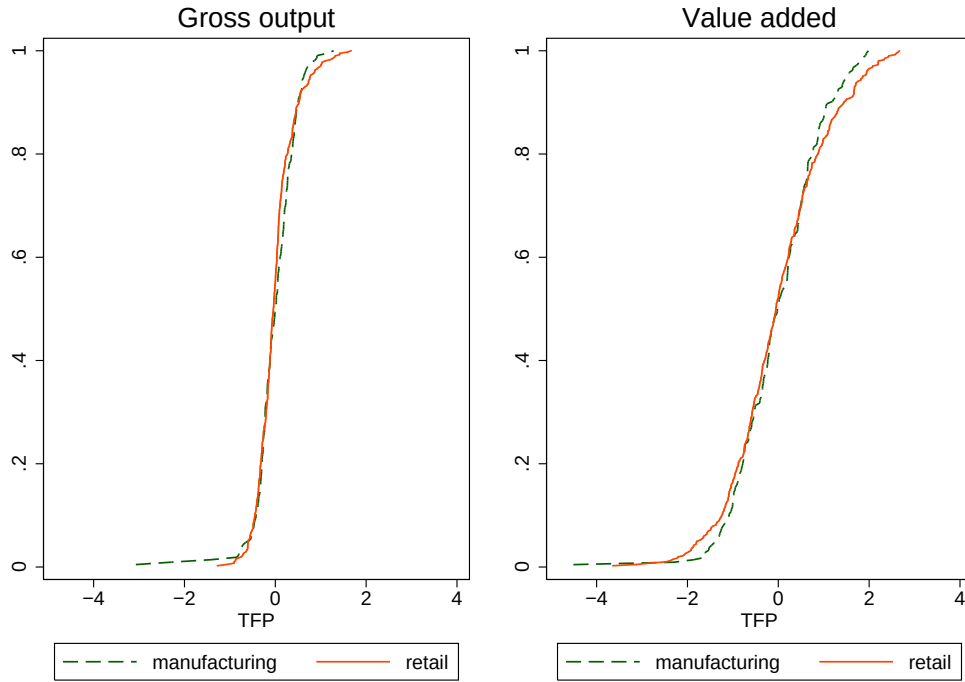


Figure 4.1: Empirical Cumulative Distribution Functions of (Demeaned) TFP

Notes: (1) TFP for manufacturing is estimated based upon regression results in columns (f) (MGGC-FE) and (l) (MVGCFE) in Table 4.5. (2) TFP for retail is estimated based upon regression results in columns (e) (RG2C) and (g) (RV2C) in Table 4.6. (3) The mean of TFP in each sector is normalized to be zero.

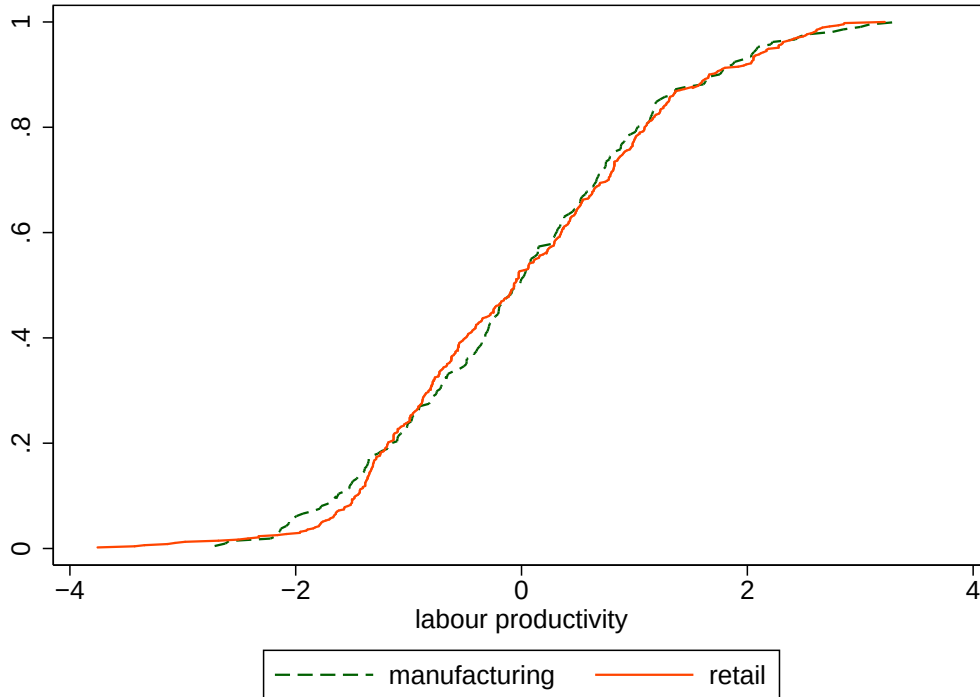


Figure 4.2: Empirical Cumulative Distribution Functions of (Demeaned) Labour Productivity

Note: The mean of labour productivity in each sector is normalized to be zero.

Table 4.15: Test for the Equality of (Demeaned) Productivity Distributions between Sectors

	Manufac- turing	Retail
(A) TFP (Gross output)		
Manufacturing	-	0.032
Retail	0.236	-
(B) TFP (Value Added)		
Manufacturing	-	0.687
Retail	0.723	-
(C) Labour productivity		
Manufacturing	-	0.454
Retail	0.358	-

Notes: (1) P-values are reported. (2) The number in the bottom left refers to a two-sample Wilcoxon rank-sum (Man-Whitney) test and the number in the top right refers to a Kolmogorov-Smirnov test for the hypothesis that two independent samples are drawn from the same distribution of populations. (3) TFP for manufacturing is estimated based upon regression results in columns (f) (MGGC-FE) and (l) (MVGC-FE) in Table 4.5. (4) TFP for retail is estimated based upon regression results in columns (e) (RG2C) and (g) (RV2C) in Table 4.6.

Table 4.16: TFP Efficiency Scores

Manufacturing			Retail		
	Mean	Std.		Mean	Std.
MGOT	0.309	(0.123)	RGOT	0.229	(0.117)
MGOC	0.290	(0.124)	RGOC	0.221	(0.118)
MGGT	0.230	(0.128)	RG2C	0.207	(0.114)
MGGC	0.145	(0.093)	RVOT	0.116	(0.153)
MGGT-FE	0.354	(0.163)	RVOC	0.110	(0.147)
MGGC-FE	0.303	(0.130)	RV2C	0.122	(0.157)
MVOT	0.150	(0.146)			
MVOC	0.154	(0.150)			
MVGT	0.083	(0.123)			
MVGC	0.086	(0.114)			
MVGT-FE	0.155	(0.162)			
MVGC-FE	0.201	(0.191)			

Note: Each letter stands for: either "Manufacturing (M)" or "Retail (R)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "System GMM (G)", "2SLS (2)", or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth. FE means the extracted fixed effects. For example, TFP estimates in MGGC-FE and MVGC-FE are calculated based upon regression results in columns (f) and (l) in Table 4.5 and TFP estimates in RG2C and RV2C are based upon regression results in columns (e) and (g) in Table 4.6.

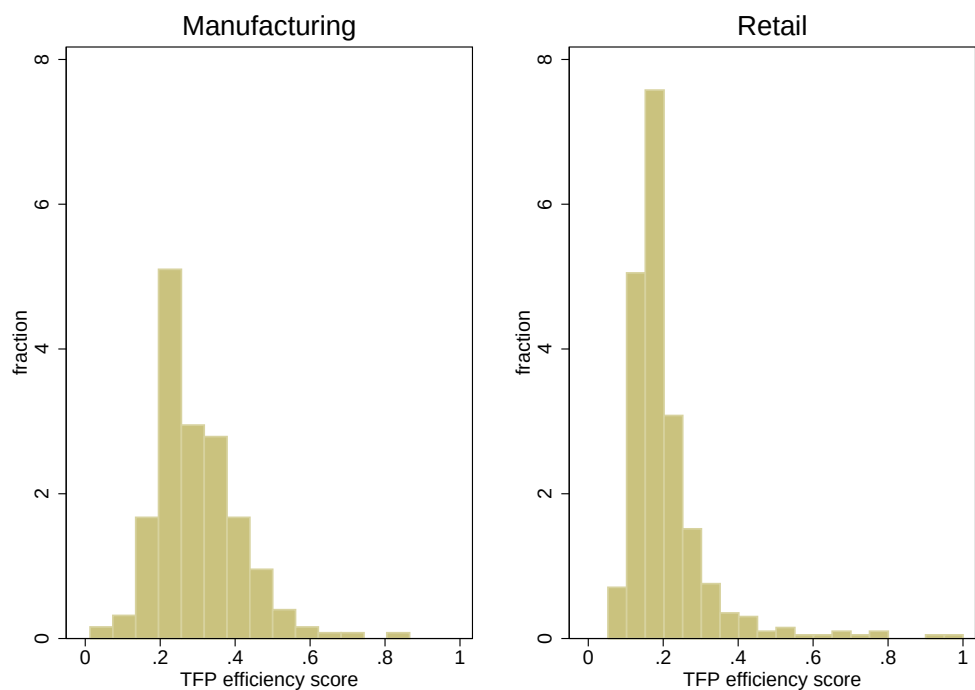


Figure 4.3: Histogram of TFP Efficiency Score Based upon Gross Output

Note: TFP is estimated based upon a regression result in column (f) (MGGC-FE) in Table 4.5 for manufacturing and that in column (e) (RG2C) in Table 4.6 for retail.

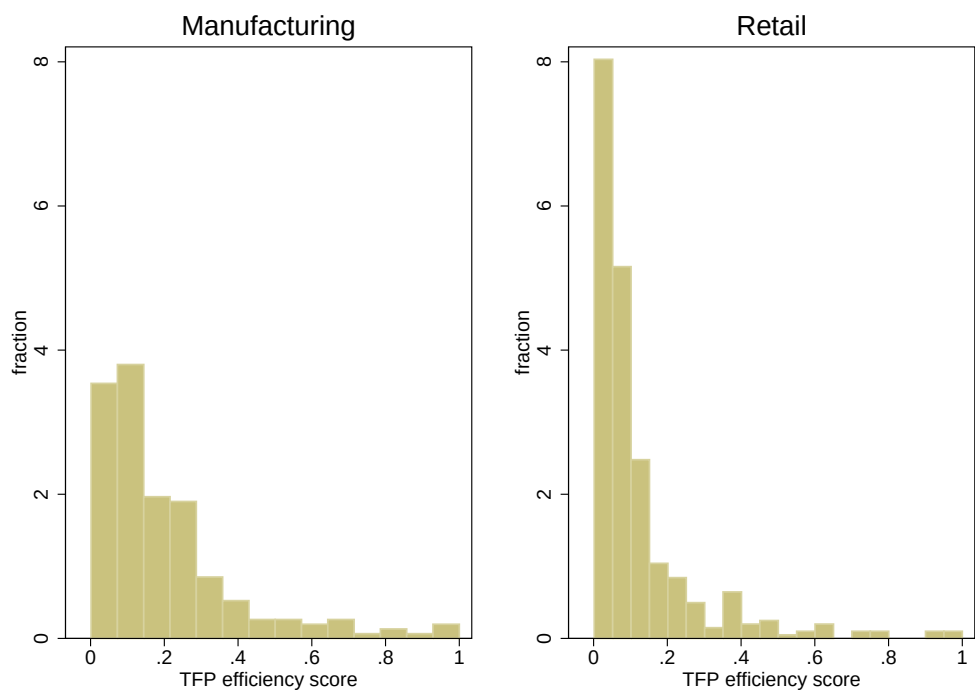


Figure 4.4: Histogram of TFP Efficiency Score Based upon Value Added

Note: TFP is estimated based upon a regression result in column (l) (MVGC-FE) in Table 4.5 for manufacturing and that in column (g) (RV2C) in Table 4.6 for retail.

Table 4.17: Appendix: Transition of GDP Share by Sector (%)

	1970				1985				2003			
	Agri- culture	Indus- try	Ser- vice		Agri- culture	Indus- try	Ser- vice		Agri- culture	Indus- try	Ser- vice	
Burkina Faso	31	23	46		35	20	45		35	22	43	
Cameroon	31	19	50		22	36	42		21	31	48	
Malawi	44	17	39		43	22	35		37	19	44	
Mali	66	12	22		40	16	44		38	24	38	
Niger	65	7	28		37	21	42		40	17	43	
Senegal	-	-	-		23	21	56		18	24	58	
Zambia	11	60	29		15	47	39		22	27	51	
Sub-Saharan Africa	20	31	49		18	34	48		19	30	51	

Source: World Bank, World Development Indicator (<http://ddp-ext.worldbank.org/ext/DDPQQ/member.do?method=getMembers&userid=1&queryId=135>).

Table 4.18: Appendix: Full Regression Results of Table 4.4 (OLS)

Dependent variables:	Manufacturing						Retail					
	log of output		log of value added		log of employees		log of output		log of value added		log of employees	
	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas	Translog	Cobb-Douglas
	(a)	(b)	(c)	(d)	(e)		(f)	(g)	(h)	(i)	(j)	
Log of employees (ln L)	1.010*** (0.194)	0.256*** (0.028)	1.426*** (0.199)	0.879*** (0.037)	-	-	0.902* (0.480)	0.194*** (0.060)	0.965 (0.609)	0.915*** (0.116)	-	-
Log of capital (ln K)	0.023 (0.076)	0.028** (0.011)	-0.083 (0.153)	0.217*** (0.019)	-	-	0.080 (0.192)	0.012 (0.024)	0.311 (0.201)	0.141** (0.069)	-	-
log of materials (ln M)	-0.015 (0.117)	0.587*** (0.025)	-	-	-	-	0.189 (0.237)	0.696*** (0.025)	-	-	-	-
Log of indirect costs (ln C)	0.314*** (0.078)	0.157*** (0.018)	-	-	-	-	-0.081 (0.197)	0.105*** (0.028)	-	-	-	-
ln ² L × 0.5	0.150*** (0.049)	-	-0.019 (0.050)	-	-	-	0.073 (0.169)	-	-0.064 (0.248)	-	-	-
ln ² K × 0.5	0.007 (0.010)	-	0.042* (0.021)	-	-	-	-0.028 (0.020)	-	-0.051 (0.060)	-	-	-
ln ² M × 0.5	0.178*** (.022)	-	-	-	-	-	0.088*** (0.023)	-	-	-	-	-
ln ² C × 0.5	0.072*** (0.016)	-	-	-	-	-	0.090*** (0.032)	-	-	-	-	-
ln L × ln K	0.013 (0.019)	-	-0.042 (0.025)	-	-	-	0.069 (0.050)	-	0.029 (0.117)	-	-	-
ln L × ln M	-0.138*** (0.022)	-	-	-	-	-	-0.063 (0.042)	-	-	-	-	-
ln L × ln C	0.015 (0.022)	-	-	-	-	-	-0.088 (0.059)	-	-	-	-	-
ln K × ln M	-0.015 (0.010)	-	-	-	-	-	-0.014 (0.022)	-	-	-	-	-
ln K × ln C	0.006 (0.010)	-	-	-	-	-	0.009 (0.021)	-	-	-	-	-
ln M × ln C	-0.083*** (0.014)	-	-	-	-	-	-0.024 (0.018)	-	-	-	-	-
Log of firm age	0.028 (0.017)	0.026 (0.019)	0.064* (0.037)	0.075** (0.037)	0.308*** (0.035)	-	0.036 (0.040)	0.050 (0.047)	0.297*** (0.079)	0.296*** (0.079)	0.182*** (0.036)	-
Technology license (dummy)	-0.007 (0.050)	0.008 (0.050)	0.223*** (0.095)	0.228*** (0.094)	0.423*** (0.082)	-	-	-	-	-	-	-
Export share of total sales (percentage)	0.000 (0.007)	0.000 (0.000)	0.004*** (0.001)	0.003*** (0.001)	0.014*** (0.001)	-	0.006 (0.005)	0.004 (0.006)	0.026* (0.015)	0.027* (0.015)	0.007* (0.004)	-
Government / State own (percentage)	-0.000 (0.001)	-0.000 (0.001)	-0.005 (0.003)	-0.005 (0.003)	0.006** (0.002)	-	-	-	-	-	-0.025* (0.013)	-
Constant	4.191*** (0.456)	3.181*** (0.256)	6.910*** (0.601)	6.277*** (0.351)	2.461*** (0.271)	-	5.767*** (1.260)	3.949*** (0.434)	7.751*** (1.174)	8.228*** (0.658)	1.817*** (0.162)	-
Legal status dummies	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-
Industry dummies	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-
Country-region dummies	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-
Country-year dummies	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-
F-test [n-values]	[0.000]	[0.088]	[0.126]	[0.002]	-	-	[0.703]	[0.871]	[0.932]	[0.631]	-	-
CRS	[0.000]	-	[0.001]	-	-	-	[0.766]	-	[0.928]	-	-	-
Cobb-Douglas	[0.000]	-	[0.055]	-	-	-	[0.000]	-	[0.686]	-	-	-
R-squared	0.955	0.944	0.802	0.799	0.512	-	0.886	0.874	0.465	0.463	0.396	-
No. of obs	1342	1342	1320	1320	1759	-	476	476	477	477	816	-

Note: See notes in Table 4.4.

Table 4.19: Appendix: Full Regression Results of Table 4.5

Dependent variables:	Log of output						Log of value added					
	OLS			System GMM			OLS			Within-Group		
	Trans- log	Cobb- Douglas	(a)	Trans- log	Cobb- Douglas	(e)	Trans- log	Cobb- Douglas	(h)	Trans- log	Cobb- Douglas	(j)
Log of employees (ln <i>L</i>)	1.125*** (0.378)	0.240*** (0.048)	0.940 (0.894)	0.270** (0.129)	0.268 (0.226)	1.189* (0.660)	1.342*** (0.394)	0.933*** (0.067)	0.566*** (0.213)	2.185*** (0.982)	0.566*** (0.213)	2.762*** (0.860)
Log of capital (ln <i>K</i>)	0.142 (0.146)	0.016 (0.025)	-0.087 (0.232)	-0.015 (0.058)	0.093 (0.078)	0.456 (0.321)	0.255 (0.241)	0.238*** (0.035)	0.046 (0.085)	0.167 (0.270)	0.046 (0.085)	0.205 (0.561)
log of materials (ln <i>M</i>)	-0.301* (0.165)	0.569*** (0.041)	-0.802** (0.347)	0.401*** (0.101)	0.543*** (0.165)	-0.467 (0.392)	-	-	-	-	-	-
Log of indirect costs (ln <i>C</i>)	0.484*** (0.145)	0.217*** (0.035)	1.083*** (0.251)	0.164** (0.071)	0.458 (0.110)	0.458 (0.485)	-	-	-	-	-	-
ln ² <i>L</i> × 0.5	0.134 (0.100)	-	-0.028 (0.206)	-	-	-0.024 (0.286)	-0.006 (0.093)	-	-	-0.576*** (0.206)	-	-0.319 (0.255)
ln ² <i>K</i> × 0.5	-0.014 (0.018)	-	-0.050* (0.028)	-	-	-0.006 (0.041)	0.009 (0.032)	-	-	-0.028 (0.042)	-	0.020 (0.043)
ln ² <i>M</i> × 0.5	0.206*** (0.031)	-	0.300*** (0.032)	-	-	0.224*** (0.062)	-	-	-	-	-	-
ln ² <i>C</i> × 0.5	-0.013 (0.042)	-	-0.062 (0.046)	-	-	-0.011 (0.080)	-	-	-	-	-	-
ln <i>L</i> × ln <i>K</i>	0.015 (0.030)	-	0.066 (0.053)	-	-	0.097 (0.091)	-0.031 (0.041)	-	-	0.057 (0.085)	-	-0.066 (0.058)
ln <i>L</i> × ln <i>M</i>	-0.160*** (0.033)	-	-0.220*** (0.067)	-	-	-0.147* (0.084)	-	-	-	-	-	-
ln <i>L</i> × ln <i>C</i>	0.041 (0.045)	-	0.131 (0.105)	-	-	-0.017 (0.160)	-	-	-	-	-	-
ln <i>K</i> × ln <i>M</i>	-0.031* (0.016)	-	-0.017 (0.021)	-	-	-0.067** (0.032)	-	-	-	-	-	-
ln <i>K</i> × ln <i>C</i>	0.037* (0.021)	-	0.058 (0.041)	-	-	0.014 (0.059)	-	-	-	-	-	-
ln <i>M</i> × ln <i>C</i>	-0.063** (0.028)	-	-0.128*** (0.046)	-	-	-0.029 (0.052)	-	-	-	-	-	-
Log of firm age	0.081** (0.040)	0.060 (0.042)	0.002 (0.100)	0.056 (0.091)	0.054 (0.130)	-0.016 (0.140)	-0.007 (0.079)	-0.009 (0.080)	0.181 (0.139)	0.302* (0.156)	0.181 (0.139)	0.143 (0.171)
Technology license (dummy)	-0.050 (0.098)	-0.055 (0.096)	0.054 (0.186)	-0.001 (0.169)	-0.195 (0.131)	-0.148 (0.145)	0.121 (0.176)	0.118 (0.176)	0.039 (0.341)	0.172 (0.331)	0.039 (0.341)	-0.150 (0.346)
Export share of total sales (percentage)	-0.000 (0.001)	0.000 (0.001)	0.004 (0.006)	0.002 (0.006)	0.002 (0.006)	0.005 (0.005)	0.003 (0.002)	0.003 (0.002)	0.014 (0.010)	0.022** (0.009)	0.014 (0.010)	0.006 (0.013)
Government / State own (percentage)	-0.001 (0.003)	-0.001 (0.003)	-0.013 (0.008)	-0.009 (0.008)	-0.000 (0.006)	0.000 (0.006)	-0.009 (0.009)	-0.009 (0.008)	-0.030*** (0.012)	-0.033*** (0.012)	-0.030*** (0.012)	-0.017 (0.016)
Constant	4.440*** (0.820)	3.177*** (0.367)	7.624*** (2.262)	5.746*** (1.246)	2.976* (1.626)	4.397 (3.209)	5.493*** (1.035)	6.290*** (0.526)	9.531*** (1.448)	5.554*** (2.625)	4.173 (3.172)	7.161*** (3.427)
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test [p-values]	[0.107]	[0.167]	[0.246]	[0.201]	[0.732]	[0.714]	[0.346]	[0.002]	-	[0.009]	-	[0.131]
CRS	[0.136]	-	[0.227]	-	[0.813]	[0.813]	[0.006]	-	[0.098]	[0.000]	-	[0.985]
Cobb-Douglas	[0.000]	-	[0.000]	-	[0.000]	[0.000]	[0.545]	-	-	[0.017]	-	[0.160]
Hansen test [p-values]	-	-	-	-	[0.579]	-	-	-	-	-	-	-
R-squared	0.930	0.913	-	-	-	-	0.732	0.731	-	-	-	-
No. of obs	487	487	487	487	487	487	476	476	476	476	476	476

Note: See notes in Table 4.5.

Table 4.20: Appendix: Marginal Effects in Table 4.7

Dependent variables:	log of output		log of value added	
	Translog	Cobb-Douglas	Translog	Cobb-Douglas
	(a)	(b)	(c)	(d)
Marginal effects of Manufacturing				
Log of employees	0.259*** [0.000]	0.262*** [0.000]	0.908*** [0.000]	0.870*** [0.000]
Log of capital	0.024** [0.020]	0.026** [0.020]	0.236*** [0.000]	0.230*** [0.000]
Log of materials	0.602*** [0.000]	0.577*** [0.000]	-	-
Log of indirect costs	0.145*** [0.000]	0.161*** [0.000]	-	-
Marginal effects of Retail				
Log of employees	0.197** [0.018]	0.157** [0.012]	1.060*** [0.000]	0.971*** [0.000]
Log of capital	0.043* [0.090]	0.030 [0.168]	0.220*** [0.000]	0.208*** [0.000]
Log of materials	0.740*** [0.000]	0.722*** [0.000]	-	-
Log of indirect costs	0.070*** [0.006]	0.098*** [0.000]	-	-

Notes: (1) Marginal effects are evaluated at the sample mean of inputs in translog specifications, whereas they are exactly estimated coefficients in Cobb-Douglas specifications. (2) See also notes in Table 4.7.

Table 4.21: Appendix: Levinsohn and Petrin Estimation Results

Dependent variables:	Levinsohn and Petrin (1)				Levinsohn and Petrin (2)			
	log of output	log of output per employee	log of value added	log of value added per employee	log of output	log of output per employee	log of value added	log of value added per employee
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Log of employees	0.226** (0.188)	-	0.539*** (0.061)	-	0.253*** (0.046)	-	0.570*** (0.064)	-
Log of capital	0.038 (0.044)	0.023 (0.049)	0.139* (0.075)	0.121 (0.081)	0.000 (0.036)	0.000 (0.093)	0.104 (0.072)	0.112 (0.081)
log of materials	0.349** (0.160)	0.393*** (0.127)	-	-	0.573*** (0.044)	0.571*** (0.038)	-	-
Log of indirect costs	0.188*** (0.029)	0.188*** (0.030)	-	-	0.136 (0.089)	0.134 (0.093)	-	-
Log of firm age	0.064 (0.041)	0.084** (0.037)	0.029 (0.064)	0.061 (0.060)	0.059 (0.043)	0.080** (0.036)	0.007 (0.064)	0.056 (0.070)
Technology license (dummy)	-0.067 (0.117)	-0.069 (0.075)	0.050 (0.164)	0.037 (0.144)	-0.057 (0.092)	-0.038 (0.098)	0.062 (0.163)	0.109 (0.168)
Export share of total sales (percentage)	-	-	-0.001 (0.002)	0.001 (0.002)	-	-	0.000 (0.002)	0.002 (0.002)
Government / State own (percentage)	-0.001 (0.004)	-0.001 (0.004)	-0.002 (0.009)	-0.003 (0.007)	-0.000 (0.003)	-0.000 (0.003)	-0.008 (0.008)	-0.007 (0.008)
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test [p-values]: CRS	[0.134]	-	[0.000]	-	[0.662]	-	[0.000]	-
R-squared	-	-	-	-	-	-	-	-
No. of obs	487	487	469	469	487	487	475	475

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) The standard errors are estimated by bootstrapping with 50 replications. (4) Materials and capital stocks (per worker in labour productivity specifications) are used as a proxy for productivity shocks in columns (a) to (d), whereas indirect costs and capital stocks (per worker in labour productivity specifications) are in column (e) to (h). (5) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status and industry dummies.

Table 4.22: Appendix: Baseline Production Functions and Size Regressions (OLS) - Labour Productivity Specification

Dependent variables:	Manufacturing						Retail					
	log of output per employee			log of value added per employee			log of output per employee			log of value added per employee		
	Translog		Cobb-Douglas		(c)	(d)	Translog		Cobb-Douglas		(f)	(g)
	(a)	(b)	(b)	(b)			(f)	(g)	(g)	(h)	(i)	(j)
Marginal effects												
Log of capital per employee	0.030** [0.012]	0.029** [0.013]	0.230*** [0.000]	0.221*** [0.000]	-	0.031 [0.224]	0.012 [0.607]	0.148** [0.037]	0.137** [0.039]	-		
Log of materials per employee	0.608*** [0.000]	0.588*** [0.000]	-	-	-	0.716*** [0.000]	0.696*** [0.000]	-	-	-		
Log of indirect costs per employee	0.149*** [0.000]	0.158*** [0.000]	-	-	-	0.075*** [0.003]	0.105*** [0.000]	-	-	-		
Regression coefficients												
Log of firm age	0.038** (0.017)	0.034* (0.018)	0.101*** (0.035)	0.104*** (0.035)	0.308*** (0.035)	0.040 (0.040)	0.052 (0.046)	0.306*** (0.080)	0.304*** (0.081)	0.182*** (0.036)		
Technology license (dummy)	-0.011 (0.050)	0.016 (0.050)	0.253*** (0.094)	0.259*** (0.094)	0.423*** (0.082)	-	-	-	-	-		
Export share of total sales (percentage)	0.001 (0.000)	0.001 (0.000)	0.004*** (0.001)	0.005*** (0.001)	0.014*** (0.001)	0.005 (0.005)	0.004 (0.006)	0.028* (0.015)	0.027* (0.015)	0.007* (0.004)		
Government / State own (percentage)	-0.000 (0.001)	-0.000 (0.001)	-0.005 (0.003)	-0.004 (0.003)	0.006** (0.002)	-	-	-	-	-0.025* (0.013)		
Constant	4.747*** (0.501)	3.239*** (0.259)	7.458*** (0.532)	6.524** (0.342)	2.461*** (0.271)	5.989*** (0.947)	3.969*** (0.390)	8.202*** (0.623)	8.357*** (0.576)	1.817*** (0.162)		
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Industry dummies	Yes	Yes	Yes	Yes	Yes	NA	NA	NA	NA	NA		
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
F-test [p-values]												
Cobb-Douglas	[0.000]	-	[0.056]	-	-	[0.000]	-	[0.280]	-	-		
R-squared	0.870	0.845	0.428	0.422	0.512	0.844	0.828	0.232	0.230	0.396		
No. of obs	1342	1342	1320	1320	1759	476	476	477	477	816		

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) In all estimations, constant returns to scale are pre-imposed on output production functions. (5) Marginal effects are evaluated at the sample mean of inputs in translog specifications, whereas they are exactly estimated coefficients in Cobb-Douglas specifications. (6) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status and industry dummies.

Table 4.23: Appendix: Production Functions of Panel Manufacturers - Labour Productivity Specification

Dependent variables:	Log of output per employee																	
	OLS			Within-Group			System GMM			Log of value added per employee								
	OLS			Within-Group			System GMM			OLS			Within-Group			System GMM		
	Trans-log (a)	Cobb-Douglas (b)	Trans-log (c)	Cobb-Douglas (d)	Trans-log (e)	Cobb-Douglas (f)	Trans-log (g)	Cobb-Douglas (h)	Trans-log (i)	Cobb-Douglas (j)	Trans-log (k)	Cobb-Douglas (l)						
Marginal effects																		
Log of capital per employee	0.016 [0.492]	0.016 [0.502]	-0.036 [0.479]	-0.008 [0.884]	0.097 [0.128]	0.124 [0.202]	0.245*** [0.000]	0.245*** [0.000]	0.059 [0.444]	0.056 [0.481]	0.211* [0.055]	0.254 [0.294]						
Log of materials per employee	0.608*** [0.000]	0.573*** [0.000]	0.561*** [0.000]	0.408*** [0.098]	0.552*** [0.000]	0.529*** [0.001]	- [0.000]	- [0.000]	- [0.001]	- [0.001]	- [0.001]	- [0.001]						
Log of indirect costs per employee	0.175*** [0.000]	0.219*** [0.035]	0.103*** [0.026]	0.158** [0.074]	0.124 [0.103]	0.140 [0.197]	- [0.000]	- [0.000]	- [0.000]	- [0.000]	- [0.000]	- [0.000]						
Regression coefficients																		
Log of firm age	0.100** (0.039)	0.077* (0.041)	-0.035 (0.088)	0.034 (0.092)	-0.072 (0.125)	0.324 (0.597)	0.057 (0.076)	0.058 (0.076)	0.156 (0.147)	0.131 (0.146)	-0.002 (0.274)	0.277 (0.595)						
Technology license (dummy)	-0.055 (0.096)	-0.040 (0.095)	0.040 (0.194)	-0.000 (0.177)	0.001 (0.004)	-0.223 (0.175)	0.181 (0.174)	0.180 (0.175)	0.048 (0.353)	0.050 (0.352)	-0.242 (0.343)	-0.309 (0.419)						
Export share (percentage)	0.000 (0.001)	0.000 (0.001)	0.001 (0.005)	-0.000 (0.005)	-0.000 (0.007)	0.000 (0.008)	0.005** (0.002)	0.005** (0.002)	0.007 (0.008)	0.007 (0.008)	0.001 (0.007)	-0.001 (0.011)						
Government / State own (percentage)	-0.000 (0.001)	-0.001 (0.003)	-0.009 (0.007)	-0.006 (0.007)	-0.550 (0.412)	-0.000 (0.006)	-0.007 (0.009)	-0.007 (0.009)	-0.023** (0.011)	-0.023** (0.011)	-0.026** (0.012)	-0.013 (0.013)						
Constant	5.180*** (0.937)	3.229*** (0.372)	7.630*** (2.073)	5.051*** (0.880)	7.366** (3.632)	2.403 (2.144)	6.711*** (0.939)	6.617*** (0.529)	7.292*** (1.244)	7.898*** (0.848)	14.798** (6.447)	6.652** (2.971)						
Legal status dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
Country-region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
Country-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
F-test [p-values]:Cobb-Douglas	[0.000]	-	[0.000]	-	[0.000]	-	[0.912]	-	[0.586]	-	[0.212]	-						
Hansen test [p-values]	-	-	-	-	[0.463]	[0.513]	-	-	-	-	[0.368]	[0.728]						
R-squared	0.831	0.794	-	-	-	-	0.362	0.362	-	-	-	-						
No. of obs	487	487	487	487	487	487	476	476	476	476	476	476						

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) Standard errors in the system GMM is corrected by the method proposed by Windmeijer (2005). (5) In all estimations, constant returns to scale are pre-imposed on output production functions. (6) The system GMM uses the first-lag as instruments for the contemporaneous first-difference and the contemporaneous first-difference as instruments for the current level with its own adjustments across columns; a dummy for whether a firm has its annual financial statements checked and certified by an external auditor in year t is additionally used as an instrument in contemporaneous first-differenced equation in (f), (k) and (l). In addition, the number of power outages that a firm experienced in a typical month in year t is used in contemporaneous first-differenced equation in (e). Capital stocks and materials per worker (inclusive of their interactions with both its own and the other factor input per worker) are always used only in the first-differenced equation. (7) Marginal effects are evaluated at the sample mean of inputs in translog specifications, whereas they are exactly estimated coefficients in Cobb-Douglas specifications. (8) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status and industry dummies.

Table 4.24: Appendix: TFP Comparison - Pooling Regressions (OLS) - Labour Productivity Specification

Dependent variable:	log of output per employee				log of value added per employee			
	(a) Translog		(b) Cobb-Douglas		(c) Translog		(d) Cobb-Douglas	
	Retail		Retail		Retail		Retail	
	× Demeaned		× Demeaned		× Demeaned		× Demeaned	
Retail dummy	-0.182 (0.255)		0.160 (0.109)		0.417 (0.451)		1.409*** (0.172)	
Log of capital per employee (manufacturing: $\ln k_1$)	0.043 (0.075)		0.027** (0.011)		-0.062 (0.131)		0.236*** (0.017)	
Log of capital per employee (retail: $\ln k_2$)	0.106 (0.181)		0.029 (0.021)		0.226 (0.138)		0.199*** (0.056)	
Log of materials per employee ($\ln m$)	-0.112 (0.123)	0.340 (0.230)	0.578*** (0.026)	0.143*** (0.035)				
Log of indirect costs per employee ($\ln c$)	0.351*** (0.078)	-0.454** (0.203)	0.163*** (0.018)	-0.065** (0.033)				
$\ln^2 k_1 \times 0.5$	0.011 (0.010)							
$\ln^2 k_2 \times 0.5$	-0.014 (0.010)							
$\ln^2 m \times 0.5$	0.177*** (0.024)	-0.100*** (0.033)						
$\ln^2 c \times 0.5$	0.073*** (0.017)	0.008 (0.036)						
$\ln^2 k_1 \times 0.5$	-0.023** (0.011)				0.042** (0.019)			
$\ln^2 k_2 \times 0.5$	-0.005 (0.021)				-0.013 (0.048)			
$\ln k_1 \times \ln c$	0.015 (0.010)							
$\ln k_2 \times \ln c$	0.001 (0.021)							
$\ln m \times \ln c$	-0.091*** (0.015)	0.071*** (0.024)						
Log of firm age	0.036** (0.015)		0.033** (0.017)		0.152*** (0.033)		0.154*** (0.033)	
Technology license (dummy)	-0.021 (0.049)		0.012 (0.050)		0.234** (0.093)		0.242*** (0.092)	
Export share of total sales (percentage)	0.001* (0.000)		0.001* (0.007)		0.004*** (0.001)		0.005*** (0.001)	
Government / State own (percentage)	-0.001 (0.001)		-0.001 (0.001)		-0.007** (0.003)		-0.007* (0.003)	
Constant	4.949*** (0.476)		3.209*** (0.239)		6.658*** (0.492)		5.696*** (0.304)	
Legal status dum.	Yes		Yes		Yes		Yes	
Industry dum.	No		No		No		No	
Country-region dum.	Yes		Yes		Yes		Yes	
Country-year dum.	Yes		Yes		Yes		Yes	
F-test [p-values]								
Manufacturing Cobb-Douglas	[0.000]		-		[0.031]		-	
Retail Cobb-Douglas	[0.001]		-		[0.784]		-	
R-squared	0.858		0.835		0.330		0.325	
No. of obs	1818		1818		1797		1797	

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) In all estimations, constant returns to scale are pre-imposed on output production functions. (5) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status dummies.

Table 4.25: Appendix: TFP Comparison - Pooling Regressions (OLS) (Exclude Malawi and Zambia)

Dependent variable:	Lof of output				log of value added			
	(a) Translog		(b) Cobb-Douglas		(c) Translog		(d) Cobb-Douglas	
	Retail		Retail		Retail		Retail	
	× Demeaned		× Demeaned		× Demeaned		× Demeaned	
Retail dummy	-0.587 (0.901)		-0.016 (0.254)		0.194 (2.102)		0.917** (0.455)	
Log of employees (ln L)	1.378*** (0.273)	0.041 (0.631)	0.396*** (0.044)	-0.321*** (0.085)	1.542*** (0.232)	0.419 (0.801)	1.046*** (0.051)	-0.172 (0.156)
Log of capital (manufacturing: ln K ₁)	0.017 (0.095)		0.021 (0.016)		-0.041 (0.165)		0.197*** (0.022)	
Log of capital (retail: ln K ₂)	-0.079 (0.259)		0.039 (0.034)		-0.267 (0.404)		0.210** (0.087)	
Log of materials (ln M)	-0.212 (0.138)	0.343 (0.310)	0.534*** (0.038)	0.183*** (0.048)				
Log of indirect costs (ln C)	0.320*** (0.096)	-0.300 (0.249)	0.157*** (0.023)	-0.050 (0.042)				
ln ² L × 0.5	0.262*** (0.091)	-0.525** (0.244)			0.022 (0.082)	-0.0330 (0.506)		
ln ² K ₁ × 0.5	0.001 (0.015)				0.040 (0.025)			
ln ² K ₂ × 0.5	-0.003 (0.038)							
ln ² M × 0.5	0.206*** (0.028)	-0.114*** (0.038)						
ln ² C × 0.5	0.057** (0.026)	0.025 (0.048)						
ln L × ln K ₁	0.022 (0.034)				-0.053 (0.034)			
ln L × ln K ₂	0.065 (0.088)				-0.072 (0.214)			
ln L × ln M	-0.181*** (0.027)	0.123** (0.057)						
ln L × ln C	-0.009 (0.029)	-0.043 (0.079)						
ln K ₁ × ln M	-0.021 (0.014)							
ln K ₂ × ln M	0.004 (0.028)							
ln K ₁ × ln C	0.017 (0.016)							
ln K ₂ × ln C	-0.007 (0.032)							
ln M × ln C	-0.076*** (0.020)	0.045 (0.029)						
Log of firm age	0.039* (0.021)		0.044* (0.025)		0.134*** (0.045)		0.136*** (0.045)	
Technology license (dummy)	0.057 (0.069)		0.074 (0.073)		0.151 (0.126)		0.154 (0.125)	
Export share of total sales (percentage)	0.001* (0.001)		0.002* (0.001)		0.006*** (0.001)		0.006*** (0.001)	
Government / State own (percentage)	0.000 (0.004)		-0.000 (0.003)		-0.005 (0.008)		-0.004 (0.008)	
Constant	4.751*** (0.582)		3.311*** (0.300)		5.356*** (0.670)		4.975*** (0.398)	
Legal status dum.	Yes		Yes		Yes		Yes	
Industry dum.	No		No		No		No	
Country-region dum.	Yes		Yes		Yes		Yes	
Country-year dum.	Yes		Yes		Yes		Yes	
F-test [p-values]								
Manufacturing								
Homotheticity	[0.000]		-		[0.281]		-	
CRS	[0.000]		[0.000]		[0.000]		[0.000]	
Cobb-Douglas	[0.000]		-		[0.185]		-	
Retail								
Homotheticity	[0.148]		-		[0.533]		-	
CRS	[0.160]		[0.346]		[0.658]		[0.533]	
Cobb-Douglas	[0.000]		-		[0.295]		-	
R-squared	0.943		0.926		0.754		0.751	
No. of obs	1133		1133		1135		1135	

Notes: (1) Unit of observations is a firm-year. (2) Figures in () and [] are standard errors and p-values, respectively. *** denotes at significance 1%, ** at 5%, and * at 10%. (3) Standard errors are robust to heteroskedasticity across firms and clustered residuals across years within each firm. (4) Country-region and country-year are classified into 23 and 14 categories, respectively. See Table 4.2 for the classification of legal status dummies.

Table 4.26: Appendix: Marginal Effects in Table 4.25

Dependent variables:	log of output		log of value added	
	Translog	Cobb-Douglas	Translog	Cobb-Douglas
	(a)	(b)	(c)	(d)
Marginal effects of Manufacturing				
Log of employees	0.395*** [0.000]	0.396*** [0.000]	1.082*** [0.000]	1.046*** [0.000]
Log of capital	0.009 [0.532]	0.021 [0.203]	0.200*** [0.000]	0.197*** [0.000]
Log of materials	0.564*** [0.000]	0.534*** [0.000]	-	-
Log of indirect costs	0.139*** [0.000]	0.157*** [0.000]	-	-
Marginal effects of Retail				
Log of employees	0.214** [0.024]	0.074 [0.311]	0.998*** [0.000]	0.874*** [0.000]
Log of capital	0.050 [0.155]	0.039 [0.246]	0.227*** [0.005]	0.210** [0.087]
Log of materials	0.731*** [0.000]	0.718*** [0.000]	-	-
Log of indirect costs	0.078*** [0.008]	0.106*** [0.002]	-	-

Notes: (1) Marginal effects are evaluated at the sample mean of inputs in translog specifications, whereas they are exactly estimated coefficients in Cobb-Douglas specifications. (2) See also notes in Table 4.25.

Table 4.27: Appendix: Sensitivity Check of (Demeaned) TFP Distributions across Estimations - Manufacturing (Exclude Malawi and Zambia)

	MGOT	MGOC	MGGT	MGGC	MGGT-FE	MGGC-FE	MVOT	MVOC	MVGT	MVGC	MVGT-FE	MVGC-FE
MGOT	-	0.504	0.239	0.066	0.312	0.058	0.001	0.001	0.000	0.000	0.001	0.001
MGOC	0.910	-	0.349	0.225	0.614	0.249	0.002	0.004	0.000	0.000	0.005	0.006
MGGT	0.742	0.796	-	0.199	0.921	0.170	0.052	0.052	0.002	0.006	0.018	0.054
MGGC	0.489	0.452	0.486	-	0.255	0.994	0.067	0.122	0.011	0.005	0.028	0.009
MGGT-FE	0.976	0.971	0.819	0.519	-	0.284	0.006	0.008	0.001	0.000	0.008	0.004
MGGC-FE	0.436	0.418	0.419	0.950	0.458	-	0.037	0.037	0.001	0.001	0.004	0.002
MVOT	0.632	0.623	0.734	0.938	0.715	0.949	-	1.000	0.426	0.358	0.793	0.687
MVOC	0.667	0.654	0.711	0.959	0.675	0.959	0.994	-	0.428	0.445	0.696	0.685
MVGT	0.370	0.421	0.463	0.557	0.420	0.587	0.634	0.652	-	0.678	0.993	0.678
MVGC	0.907	0.935	0.911	0.953	0.905	0.981	0.908	0.882	0.792	-	0.626	0.922
MVGT-FE	0.701	0.722	0.684	0.843	0.692	0.894	0.798	0.812	0.879	0.861	-	0.915
MVGC-FE	0.798	0.822	0.742	0.929	0.809	0.974	0.821	0.802	0.860	0.807	0.995	-

Notes: (1) P-values are reported. (2) The number in the bottom left refers to a two-sample Wilcoxon rank-sum (Man-Whitney) test and the number in the top right refers to a Kolmogorov-Smirnov test for the hypothesis that two independent samples are drawn from the same distribution of populations. (3) Each letter in row and column stands for: "Manufacturing (M)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "System GMM (G)" or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth. FE means the extracted fixed effects.

Table 4.28: Appendix: Sensitivity Check of TFP Distributions across Estimations - Retail (Exclude Malawi and Zambia)

	RGOT	RGOC	RG2C	RVOT	RVOC	RV2C
RGOT	-	0.131	0.053	0.000	0.000	0.000
RGOC	0.469	-	0.352	0.000	0.000	0.000
RG2C	0.599	0.959	-	0.000	0.000	0.000
RVOT	0.200	0.251	0.350	-	1.000	0.964
RVOC	0.308	0.377	0.464	0.848	-	0.975
RV2C	0.501	0.586	0.665	0.850	0.984	-

Notes: (1) P-values are reported. (2) The number in the bottom left refers to a two-sample Wilcoxon rank-sum (Man-Whitney) test and the number in the top right refers to a Kolmogorov-Smirnov test for the hypothesis that two independent samples are drawn from the same distribution of populations. (3) Each letter in row and column stands for: "Retail (R)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "2SLS (2)" or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth.

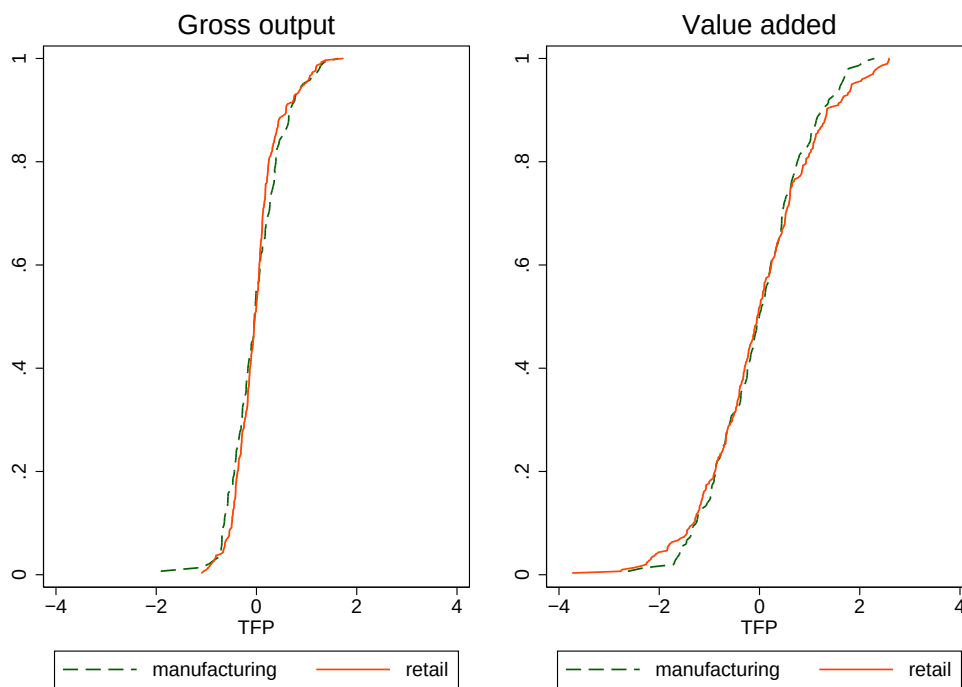


Figure 4.5: Appendix: Empirical Cumulative Distribution Functions of (Demeaned) TFP (Exclude Malawi and Zambia)

Notes: (1) TFP for manufacturing is estimated based upon MGGC-FE and MVGC-FE. (2) TFP for retail is estimated based upon RG2C and RV2C. (3) The mean of TFP in each sector is normalized to be zero.

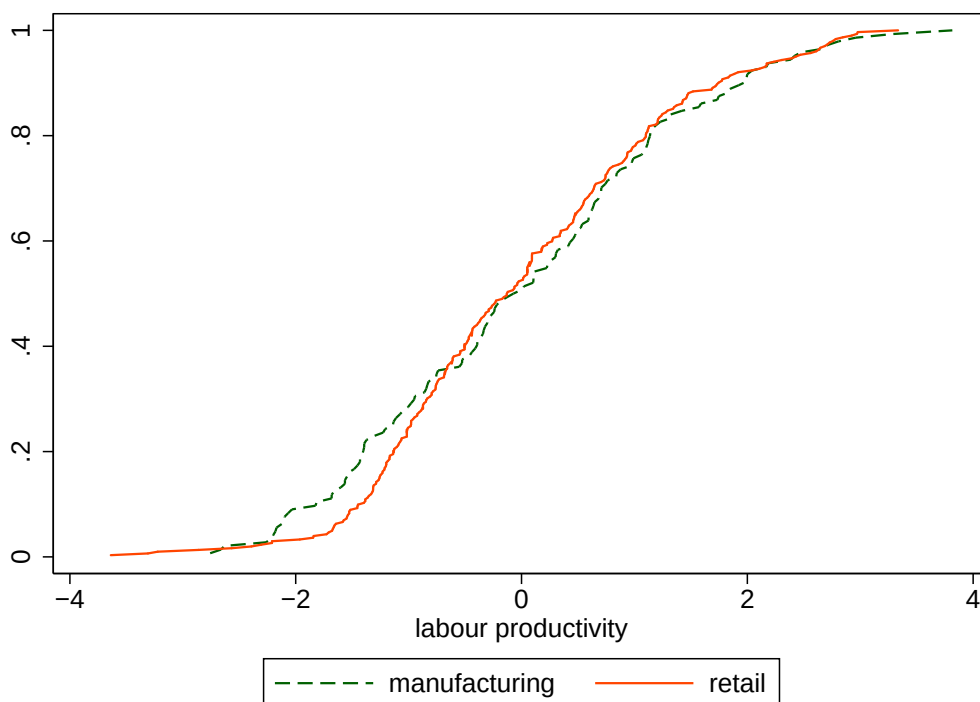


Figure 4.6: Appendix: Empirical Cumulative Distribution Functions of (Demeaned) Labour Productivity (Exclude Malawi and Zambia)

Note: The mean of labour productivity in each sector is normalized to be zero.

Table 4.29: Appendix: Test for the Equality of (Demeaned) Productivity Distributions (Exclude Malawi and Zambia)

	Manufac- turing	Retail
(A) TFP (Gross output)		
Manufacturing	-	0.177
Retail	0.975	-
(B) TFP (Value Added)		
Manufacturing	-	0.983
Retail	0.875	-
(C) Labour productivity		
Manufacturing	-	0.107
Retail	0.674	-

Notes: (1) P-values are reported. (2) The number in the bottom left refers to a two-sample Wilcoxon rank-sum (Man-Whitney) test and the number in the top right refers to a Kolmogorov-Smirnov test for the hypothesis that two independent samples are drawn from the same distribution of populations. (3) TFP for manufacturing is estimated based upon MGGC-FE and MVGC-FE. (4) TFP for retail is estimated based upon RG2C and RV2C.

Table 4.30: Appendix: Sensitivity Check of TFP Efficiency Scores (Exclude Malawi and Zambia)

	Manufacturing			Retail	
	Mean	Std.		Mean	Std.
MGOT	0.327	(0.145)	RGOT	0.242	(0.119)
MGOC	0.272	(0.139)	RGOC	0.234	(0.122)
MGGT	0.276	(0.166)	RG2C	0.200	(0.118)
MGGC	0.199	(0.157)	RVOT	0.061	(0.101)
MGGT-FE	0.310	(0.168)	RVOC	0.126	(0.167)
MGGC-FE	0.227	(0.144)	RV2C	0.137	(0.179)
MVOT	0.172	(0.171)			
MVOC	0.180	(0.177)			
MVGT	0.089	(0.130)			
MVGC	0.046	(0.100)			
MVGT-FE	0.150	(0.163)			
MVGC-FE	0.156	(0.164)			

Note: Each letter stands for: either "Manufacturing (M)" or "Retail (R)" in the first; either "Gross output (G)" or "Value added (V)" in the second; "System GMM (G)", "2SLS (2)", or "OLS (O)" in the third; and "Translog (T)" or "Cobb-Douglas (C)" in the fourth. FE means the extracted fixed effects.

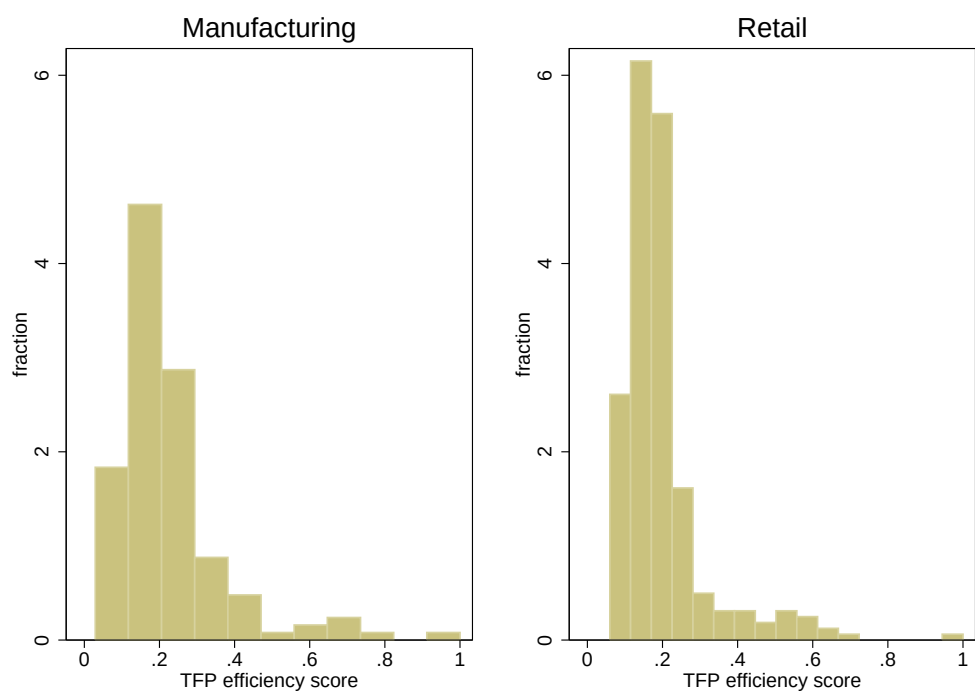


Figure 4.7: Appendix: Histogram of TFP Efficiency Score Based upon Gross Output (Exclude Malawi and Zambia)

Note: TFP is estimated based upon MGGC-FE for manufacturing and RG2C for retail.

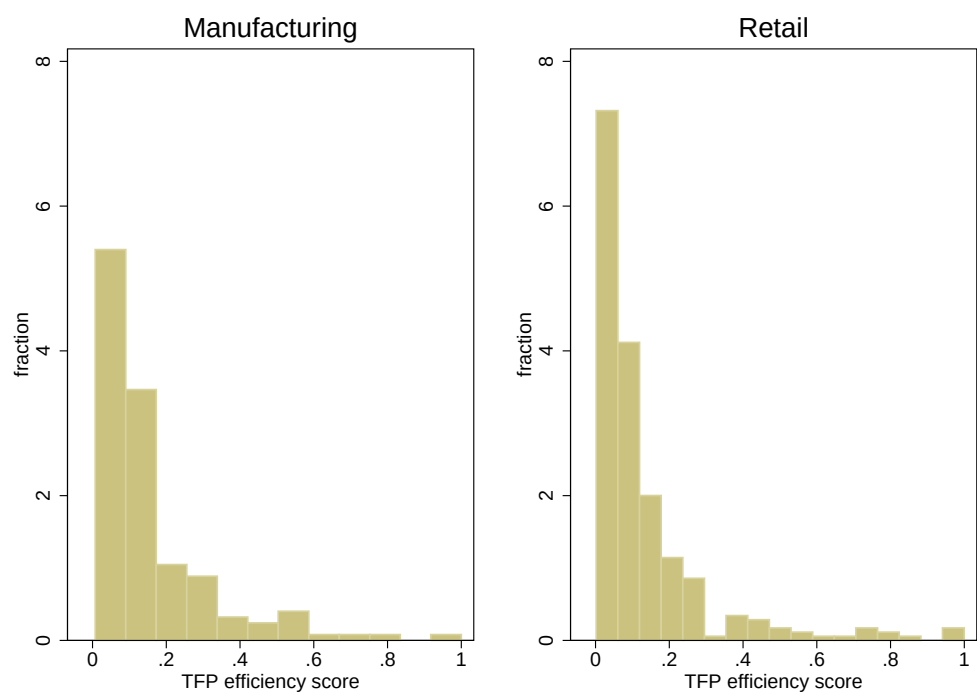


Figure 4.8: Appendix: Histogram of TFP Efficiency Score Based upon Value Added (Exclude Malawi and Zambia)

Note: TFP is estimated based upon MVGC-FE for manufacturing and RV2C for retail.

Chapter 5

Conclusion

As industrialisation and urbanisation are common phenomena in the process of economic development, the exploration of the pattern of economic growth is more or less equivalent to the examination of the pattern of sectoral and spatial transformations in the economy. In many African countries, we usually observe a substantial spatial mobility without industrialisation. Then, whether or not Africa is on the growth path of the economy is a fundamental question. The thesis implicitly attempted to improve our understanding of the question by exploring the income determination process in labour markets from these spatial and sectoral dimensions. The thesis comprised three independent but thematically related core chapters. Chapters 2 and 3 shed light on the spatial aspect, whereas chapter 4 focused on the sectoral aspect.

According to the recent international policy debates, establishing credible institutions is the mainstay of spatially targeted policies. These institutions usually include the provision of basic social services, law and order, and the establishment of well-functioning land markets. Chapter 2 implicitly asked the question of how education, one of these institutions, could impact on the pattern of spatial transformation by using long-run household panel data from rural Tanzania. If the gains from migration are entirely attributed to the higher return to schooling at the destination, those who cannot afford the cost of schooling (as well as the cost of migration) cannot reap the benefits of migration. Then, an ability to have access to education substantially affects the pattern of spatial transformation, influencing the welfare distribution across people and across regions. We found that the pre-migration level of schooling raised the probability to migrate, but that education

was not the main part of the story in explaining the gains from migration. This suggests that some mechanisms, operating other than through education, prevent individuals from exploiting the existing higher income and consumption opportunities at the destination. To improve our understanding of the pattern of spatial transformation and link it with an effective policy design, further research is required for these unexplored mechanisms.

Urbanisation often draws attention to its negative impact, such as the declining productivity associated with congestion costs, the grime and crime in rapidly expanding cities and the proliferating urban slums. Consequently, the policy response often heads for curbing the process and controlling urban growth. Nevertheless, it is not well understood how rural-to-urban migrants are performing in urban labour markets, due to the lack of good migration and labour force data. Chapter 3 attempted to provide some new findings on migrants' performance in urban labour markets by using data drawn from urban household panel surveys in Ghana and Tanzania. Specifically, the chapter investigated how rural-to-urban migrants' earnings compared with those of natives in urban labour markets by attempting to identify the process of their assimilation at the destination. The novelty in the chapter was to make a distinction between wage-dependent and self-employed migrants. There is a prevailing preconception that wage jobs are better than self-employed jobs in developing economies. However, the chapter found that wage employed migrants would achieve higher lifetime earnings if they entered a self-employed sector from their arrival, after controlling for individuals' traits and allowing for the varying returns to those traits across urban residents. The evidence points to a view that those who do not have physical and social capitals enough to start a business cannot enter a potentially profitable self-employed sector.

How the sectoral activities are interacting with each other is an important question to improve our understanding of the pattern of economic development. This is because the growth of economies is always associated with the expansion of high productive sectors and the contraction of low productive sectors. Using firm-level data of manufacturing and retailing from seven Sub-Saharan African countries, chapter 4 attempted to improve our understanding of the sectoral difference in enterprise performance in urban Africa by investigating three aspects of firms' productive structure: technology, TFP, and firm size. We found that technology was similar between sectors, that retailing firms were smaller and less capital intensive but not, on average, ones with lower TFP, and that TFP differences are primarily within sectors, all pointing to the primary importance of

factor prices in characterising the urban industrial structure.

Reallocating labour through internal migration is a key part of the income determination process in the economy. The heterogeneity in enterprise performance also plays a major role in determining the income distribution in an urban economy. We expected that education was an important factor in explaining migrants' gains and, as such, limiting the internal migration of the less educated. Our priors regarding urban migrants' activities were that wage jobs were preferable to self-employed jobs and that a self-employed sector was a free-entry sector. We anticipated that the heterogeneity in technology, TFP, and firm size *across* sectors primarily characterised the urban industrial structure. However, our findings in this thesis contradicted these priors, pointing to the importance of migration constraints operating other than through education, capital constraints for migrants to start a business, and the heterogeneity in firm performance *within* sectors.

We acknowledge that these findings can be challenged on the ground of the difficulty in controlling for potential endogeneity. The endogenous selection into migration and schooling was the key econometric concern in chapter 2. We addressed the problem by taking two approaches: the IV and CF approaches. Chapter 3 addressed the quality difference across arrival cohorts, associated with some selection mechanisms in urban labour markets, by directly controlling for the reasoning ability and cognitive skills of individuals, measured by an individual's test scores. In chapter 4, the simultaneously determined factor choices and output were the main econometric challenge in estimating a production function. In addition, the unobserved sectoral characteristics (worker quality, match quality, etc.) were a potential problem in comparing TFP and firm size between sectors. However, having instruments, particularly based on panel data in the case of manufacturing, and examining some qualitative information that characterises unobserved inter-sectoral differences enabled us to somehow address these issues in that chapter.

Is the pattern of spatial and sectoral transformations observed in Africa a part of the growth process? Our findings only begin to answer this question. The spatial transformation not associated with the sectoral transformation might be a particular concern for policymakers and a puzzling question for researchers. In order to provide policymakers with clear guidance on policy priorities and effective policy interventions, the research must be continued.

Bibliography

Abdulai, Awudu and Anna CroleRees, “Determinants of Income Diversification Amongst Rural Households in Southern Mali,” *Food Policy*, 2001, 26, 437–452.

Akerberg, Daniel A., Kevin Caves, and Garth Frazer, “Structural Identification of Productivity Functions,” December 2006. mimeo.

Aigner, Dennis, C. A. Knox Lovell, and Peter Schmidt, “Formulation and Estimation of Stochastic Frontier Production Function Models,” *Journal of Econometrics*, 1977, 6, 21–37.

Akerlof, George A., “Labor Contracts as Partial Gift Exchange,” *Quarterly Journal of Economics*, November 1982, 97 (4), 543–569.

—, “Gift Exchange and Efficiency-Wage Theory: Four Views,” *American Economic Review*, May 1984, 74 (2), 79–83. Papers and Proceedings of the Ninety-Sixth Annual Meeting of the American Economic Association.

Alba, Richard and Victor Nee, “Rethinking Assimilation Theory for a New Era of Immigration,” *International Migration Review*, Winter 1997, 31 (3), 826–874.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” *Journal of Political Economy*, 2005, 113 (1).

Andersson, Jens A., “Reinterpreting the Rural-Urban Connection: Migration Practices and Socio-Cultural Dispositions of Buhera Workers in Harare,” *Africa: Journal of the International African Institute*, 2001, 71 (1), 82–112.

- Angrist, Joshua D. and Jörn-Steffen Pischke**, *Mostly Harmless Econometrics: An Empirical Companion*, Princeton University Press, 2009.
- Attanasio, Orazio and Katja Kaufmann**, “Educational Choices, Subjective Expectations, and Credit Constraints,” June 2009. Working Paper 15087, NBER Working Paper Series.
- Azam, Jean-Paul and Jean-Yves Leuseur**, “Efficiency Wage and Supervision: Theory and Application to the Ivorian Manufacturing Sector,” *Journal of African Economies*, 1997, (3).
- Baily, Martin Neil and Eric Zitzewitz**, *Service Sector Productivity Comparisons: Lesson for Measurement*, University of Chicago Press, 2001. pp. 419-464, Chapter 11 in Charles R. Hulten, Edwin R. Dean and Michael J. Harpe (editors), *New Development in Productivity Analysis*, National Bureau of Economic Research (NBER) Book Series Studies in Income and Wealth.
- **and Robert M. Solow**, “International Productivity Comparisons Built from the Firm Level,” *Journal of Economic Perspectives*, Summer 2001, 15 (3). 151-172.
- Banerjee, Abhijit V. and Andrew F. Newman**, “Occupational Choice and the Process of Development,” *Journal of Political Economy*, April 1993, 101 (2), 274–298.
- Banerjee, Biswajit**, “Rural-urban Migration and Family Ties: An Analysis of Family Considerations in Migration Behaviour in India,” *Oxford Bulletin of Economics and Statistics*, 1981, 43 (4), 321–355.
- , “Social Networks in the Migration Process: Empirical Evidence on Chain Migration in India,” *Journal of Development Areas*, 1983, 17, 185–196.
- , “The Probability, Size, and Uses of Remittances from Urban to Rural Areas in India,” *Journal of Development Economics*, 1984, 16, 293–311.
- **and S. M. Kanbur**, “On the Specification and Estimation of Macro Rural-urban Migration Functions: With an Application to Indian Data,” *Oxford Bulletin of Economics and Statistics*, 1981, 43 (1), 7–29.
- Baptist, Simon and Francis Teal**, “Why do South Korean Firms Produce So Much More Output Per Worker than Ghanaian One?,” February 2008, 10. Centre for the Study of African Economies (CSAE), Working Paper Series.

- Barham, Bradford and Stephen Boucher**, “Migration, Remittances, and Inequality: Estimating the Net Effects of Migration on Income Distribution,” *Journal of Development Economics*, 1998, *55*, 307–331.
- Barnum, H. N. and R. H. Sabot**, “Education, Employment Probabilities and Rural-Urban Migration in Tanzania,” *Oxford Bulletin of Economics and Statistics*, 1977, *39* (2), 109–126.
- Barr, Abigail**, “Social Capital and Technical Information Flows in the Ghanaian Manufacturing Sector,” *Oxford Economic Papers*, 2000, *52*, 539–559.
- **and Abena Oduro**, “Ethnic Fractionalization in an African Labour Market,” *Journal of Development Economics*, 2002, *68*, 355–379.
- Barrett, C.B., T. Reardon, and P. Webb**, “Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications,” *Food Policy*, 2001, *26*, 315–331.
- Bartel, Ann P.**, “The Migration Decision: What Role Does Job Mobility Play?,” *American Economic Review*, December 1979, *69* (5), 775–786.
- Bartelsman, Eric J. and Mark Doms**, “Understanding Productivity: Lessons from Longitudinal Microdata,” *Journal of Economic Literature*, September 2000, *XXXVIII*, 569–594.
- Basu, Susanto and John G. Fernald**, “Are Apparent Productive Spillovers a Figment of Specification Error?,” *Journal of Monetary Economics*, 1995, *36*, 165–188.
- **and —**, “Returns to Scale in U.S. Production: Estimates and Implications,” *Journal of Political Economy*, 1997, *105* (2), 249–283.
- Battese, George E. and Tim J. Coelli**, “Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data,” *Journal of Econometrics*, 1988, *38*, 387–399.
- Becker, Charles M. and Christopher D. Grewe**, “Cohort-Specific Rural-Urban Migration in Africa,” *Journal of African Economics*, 1996, *5* (2), 228–270.
- Beegle, Kathleen**, “Labor Effects of Adult Mortality in Tanzanian Households,” *Economic Development and Cultural Change*, April 2005, *53*, 655–683.

- , **Joachim Deweerdt**, and **Stefan Dercon**, “Adult Mortality and Consumption Growth in the Age of HIV/AIDS,” *Economic Development and Cultural Change*, January 2008, 56, 299–326.
- , —, and —, “Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey,” *Policy Research Working Paper*, December 2008, 4798. World Bank, Development Research Group, Poverty Team.
- Benjamin, Dwayne**, “Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models,” *Econometrica*, March 1992, 60 (2), 287–322.
- Berndt, Ernst R. and Laurits R. Christensen**, “The Internal Structure of Functional Relationships: Separability, Substitution, and Aggregation,” *Review of Economic Studies*, July 1973, 40 (3), 403–410.
- and —, “The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-1968,” *Journal of Econometrics*, 1973, 1, 81–114.
- Bhagwati, Jagdish N. and T. N. Srinivasan**, “On Reanalyzing the Harris-Todaro Model: Policy Rankings in the Case of Sector-Specific Sticky Wage,” *American Economic Review*, June 1974, 64 (3), 502–508.
- Bhattacharya, Prabir C.**, “Rural-urban Migration in Economic Development,” *Journal of Economic Surveys*, 1993, 7 (3), 243–281.
- Biggs, Tyler and Manju Kedia Shah**, “African Small and Medium Enterprise Networks, and Manufacturing Performance,” February 2006. World Bank Policy Research, Working Paper 3855, World Bank.
- , **Manju Shah**, and **Pradeep Srivastava**, “Technological Capabilities and Learning in African Enterprises,” December 1995. World Bank Technical Paper Number 288, Africa Technical Department Series.
- Bigsten, Arne and Måns Söderbom**, “What Have We Learnt from a Decade of Manufacturing Enterprise in Africa?,” February 2006. mimeo.
- , **Anders Isaksson**, **Måns Söderbom**, **Albert Zeufack**, **Stephen Dercon**, **Marcel Fafchamps**, **Jan Willem Gunning**, **Francis Teal**, **Bernard Gauthier**, **Abena Oduro**,

- Remeco Oostendorp, and Cathy Pattillo**, “Rates of Return on Physical and Human Capital in Africa’s Manufacturing Sector,” *Economic Development and Cultural Change*, July 2000, 48, 801–827.
- , **Paul Collier, Stefan Dercon, Bernand Gauthier, Jan Willem Gunning, Anders Isaksson, Abena Oduro, Remco Oostendorp, Cathy Pattilo, Måns Söderbom, Francis Teal, and Albert Zeufack**, “Investment in Africa’s Manufacturing Sector: A Four Country Panel Data Analysis,” *Oxford Bulletin of Economics and Statistics*, 1999, 61 (4), 489–511.
- , —, —, —, **Marcel Fafchamps, Bernand Gauthier, Jan Willem Gunning, Abena Oduro, Remco Oostendorp, Catherine Pattillo, Måns Söderbom, Francis Teal, and Albert Zeufack**, “Do African Manufacturing Firms Learn from Exporting?,” *Journal of Development Studies*, February 2004, 40 (3), 115–141.
- , —, —, —, —, **Bernard Gauthier, Jan Willem Gunning, Jean Habarurema, Abena Oduro, Remco Oostendorp, Catherine Pattillo, Måns Söderbom, Francis Teal, and Albert Zeufack**, “Exports and Firm-Level Efficiency in African Manufacturing,” December 1999. mimeo.
- Black, Richard, Claudia Natali, and Jessica Skinner**, “Migration and Inequality,” *World Development Report*, 2006. Background Papers.
- Blalock, Garrick and Paul Gertler**, “Learning from Exporting Revisited in a Less Developed Setting,” *Journal of Development Economics*, 2004, 75, 397–416.
- Blanchflower, David G., Andrew J. Oswald, and Peter Sanfey**, “Wages, Profits, and Rent-sharing,” *Quarterly Journal of Economics*, February 1996, 111 (1), 227–251.
- Blundell, Richard and Ian Preston**, “Consumption Inequality and Income Uncertainty,” *Quarterly Journal of Economics*, May 1998, 113 (2), 603–640.
- **and Stephen Bond**, “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics*, 1998, 87, 115–143.
- **and —**, “GMM Estimation with Persistent Data: An Application to Production Functions,” *Econometric Reviews*, 2000, 19 (3), 321–340.

- , **Lorraine Dearden**, and **Barbara Sianesi**, “Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey,” *Journal of the Royal Statistical Society. Series A*, July 2005, *168* (3), 473–512.
- Boissiere, M., J. B. Knight**, and **R. H. Sabot**, “Earnings, Schooling, Ability and Cognitive Skills,” *American Economic Review*, December 1985, *75* (5), 1016–1030.
- Bond, Stephen**, “Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice,” 2002. cemmap working paper CWP09/02, The Institute for Fiscal Studies Department of Economics, UCL.
- Borjas, George J.**, “Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants,” *Journal of Labor Economics*, October 1985, *3* (4), 463–489.
- , “Self-selection and the Earnings of Immigrants,” *American Economic Review*, September 1987, *77* (4), 531–553.
- , “The Economics of Immigration,” *Journal of Economic Literature*, December 1994, *32* (4), 1667–1717.
- , “Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?,” *Journal of Labor Economics*, April 1995, *13* (2), 201–245.
- , “The Earnings of Mexican Immigrants in the United States,” *Journal of Development Economics*, 1996, *51*, 69–98.
- , **Stephen G. Bronars**, and **Stephen J. Trejo**, “Assimilation and the Earnings of Young Internal Migrants,” *Review of Economics and Statistics*, February 1992, *74* (1), 170–175.
- Bowles, Samuel**, “Migration as Investment: Empirical Tests of the Human Investment Approach to Geographical Mobility,” *Review of Economics and Statistics*, November 1970, *52* (4), 356–362.
- Bowsher, Clive G.**, “On Testing Overidentifying Restrictions in Dynamic Panel Data Models,” *Economics Letters*, October 2002, *77*, 211–220. Issue 2.
- Byerlee, Derek**, “Rural-Urban Migration in Africa: Theory, Policy and Research Implications,” *International Migration Review*, Winter 1974, *8* (4), 543–566.

- Canagarajah, S., C. Newman, and R. Bhattamishra**, “Non-farm Income, Gender, and Inequality: Evidence from Rural Ghana and Uganda,” *Food Policy*, 2001, 26, 405–420.
- Card, David**, “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems,” *Econometrica*, September 2001, 69 (5), 1127–1160.
- Carrington, William J., Enrica Detragiache, and Tara Vishwanath**, “Migration with Endogenous Moving Costs,” *American Economic Review*, September 1996, 86 (4), 909–930.
- Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert**, “The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity,” *Econometrica*, November 1982, 50 (6), 1393–1414.
- , —, and —, “Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers,” *Economic Journal*, March 1982, 92, 73–86.
- Charnes, A., W. W. Cooper, and E. Rhodes**, “Measuring the Efficiency of Decision Making Units,” *European Journal of Operational Research*, 1978, 2, 429–444.
- Chen, Kong-Pin, Shin-Hwan Chiang, and Siu Fai Leung**, “Migration, Family, and Risk Diversification,” *Journal of Labor Economics*, 2003, 21 (2), 353–378.
- Chiquiar, Daniel and Gordon H. Hanson**, “International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States,” *Journal of Political Economy*, 2005, 113 (2), 239–280.
- Chiswick, Barry R.**, “The Effect of Americanization on the Earnings of Foreign-born Men,” *Journal of Political Economy*, October 1978, 86 (5), 897–921.
- Christensen, Laurits R. and William H. Greene**, “Economies of Scale in U.S. Electric Power Generation,” *Journal of Political Economy*, August 1976, 84 (4, Part 1), 655–676.
- , **Dale W. Jorgenson, and Lawrence J. Lau**, “Conjugate Duality and the Transcendental Logarithmic Function,” *Econometrica*, July 1971, 39 (4), 255–256. In *Economic Theory of Organization II* (pp. 251-266).
- , —, and —, “Transcendental Logarithmic Production Frontiers,” *Review of Economics and Statistics*, February 1973, 55 (1), 28–45.

- Christofides, Louis N. and Andrew J. Oswald**, “Real Wage Determination and Rent-Sharing in Collective Bargaining Agreements,” *Quarterly Journal of Economics*, August 1992, *107* (3), 985–1002.
- Clerides, Sofronis K., Saul Lach, and James R. Tybout**, “Is ”Learning-by-Exporting” Important? Micro-Dynamic Evidence from Colombia, Mexico and Morocco,” 1996. NBER Working Paper Series, Working Paper 5715, National Bureau of Economic Research.
- , —, and —, “Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, Morocco,” *Quarterly Journal of Economics*, August 1998, *113* (3), 903–947.
- Cole, William E. and Richard D. Sanders**, “Internal Migration and Urban Employment in the Third World,” *American Economic Review*, June 1985, *75* (3), 481–494.
- Daveri, Francesco and Riccardo Faini**, “Where do Migrants GO?,” *Oxford Economic Papers*, 1999, *51*, 595–622.
- Delgado, Miguel A., Jose C. Fariñas, and Sonia Ruano**, “Firm Productivity and Export Markets: A Non-parametric Approach,” *Journal of International Economics*, 2002, *57*, 397–422.
- Denny, Michael and Melvyn Fuss**, “The Use of Approximation Analysis to Test for Separability and the Existence of Consistent Aggregates,” *American Economic Review*, June 1977, *67* (3), 404–418.
- Dercon, Stefan**, “Growth and Shocks: Evidence from Rural Ethiopia,” *Journal of Development Economics*, 2004, *74*, 309–329.
- , “Economic Reform, Growth and the Poor: Evidence from Rural Ethiopia,” *Journal of Development Economics*, 2006, *81*, 1–24.
- and **Pramila Krishnan**, “Income Portfolios in Rural Ethiopia and Tanzania: Choices and Constraints,” *Journal of Development Studies*, August 1996, *32* (6), 850–875.
- Eckstein, Zvi and Yoram Weiss**, “On the Wage Growth of Immigrants: Israel, 1999-2000,” *Journal of the European Economic Association*, June 2004, *2* (4), 665–695.
- Ericson, Richard and Ariel Pakes**, “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, January 1995, *62* (1), 53–82.

- Fafchamps, Marcel and Agnes R. Quisumbing**, “Human Capital, Productivity, and Labor Allocation in Rural Pakistan,” *Journal of Human Resources*, Spring 1999, *34* (2), 369–406.
- **and Måns Söderbom**, “Wages and Labor Management in African Manufacturing,” *Journal of Human Resources*, 2006, *XLI* (2), 346–379.
- **and Said El Hamine**, “Firm Productivity, Wages, and Agglomeration Externalities,” June 2004, *32*. Centre for the Study of African Economies (CSAE), Working Paper Series.
- **, Jan Willem Gunning, and Remco Oostendorp**, “Inventories and Risk in African Manufacturing,” *Economic Journal*, 2000, *110*, 861–893.
- Farrell, M. J.**, “The Measurement of Productive Efficiency,” *Journal of the Royal Statistical Society. Series A (General)*, 1957, *120* (3), 253–290.
- Fields, Gary S.**, “Rural-urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs,” *Journal of Development Economics*, 1975, *2*, 165–187.
- Frazer, Garth**, “Which Firm Die? A Look at Manufacturing Firm Exit in Ghana,” *Economic Development and Cultural Change*, April 2005, *53*, 585–617.
- Friedberg, Rachel M.**, “The Labor Market Assimilation of Immigrants in the United States: The Role of Age at Arrival,” December 1992. mimeo.
- **, “You Can’t Take it with You? Immigrant Assimilation and the Portability of Human Capital,”** *Journal of Labor Economics*, 2000, *18* (2), 221–251.
- Friedman, Milton**, *A Theory of Consumption Function*, Princeton University Press, 1957.
- Garen, John**, “The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable,” *Econometrica*, September 1984, *52* (5), 1199–1218.
- Ghatak, Maitreesh and Neville Nien-Huei Jiang**, “A Simple Model of Inequality, Occupational Choice, and Development,” *Journal of Development Economics*, 2002, *69*, 205–226.
- Ghatak, Subrata, Paul Levine, and Stephen Wheatley Price**, “Migration Theories and Evidence: An Assessment,” *Journal of Economic Surveys*, 1996, *10* (2), 159–198.
- Glaeser, Edward L.**, “Learning in Cities,” *Journal of Urban Economics*, 1999, *46*, 254–277.

— and David C. Maré, “Cities and Skills,” *Journal of Labor Economics*, 2001, 19 (2), 316–341.

Gong, Xiaodong, Arthur Van Soest, and Elizabeth Villagomez, “Mobility in the Urban Labor Market: A Panel Data Analysis for Mexico,” *Economic Development and Cultural Change*, October 2004, 53, 1–34.

Grant, E. Kenneth and John Vanderkamp, “The Effects of Migration on Income: A Micro Study with Canadian Data 1965-71,” *Canadian Journal of Economics*, Augst 1980, 13 (3), 381–406.

Greene, William H., *Frontier Production Functions*, Blackwell Publishers Ltd., 1999. pp. 81-166, Chapter 3 in M. Hashem Pesaran and Peter Schmidt (editors), *Handbook of Applied Econometrics Volume II: Microeconomics* (paperback).

— , *Econometric Analysis* 2003. fifth edition, Prentice Hall, Pearson Education International.

Greenwood, Michael J., “A Regression Analysis of Migration to Urban Areas of a Less-Developed Country: The Case of India,” *Journal of Regional Science*, 253-262 1971, 11 (2).

Griliches, Zvi and Jacques Mairesse, *Production Functions: The Search for Identification*, Cambridge University Press, 1998. pp. 169-203, Chapter 6 in Steinar Strøm (editor), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*.

Hagen-Zanker, Jessica and Melissa Siegel, *The Determinants of Remittances: A Review of the Literature* June 2007. Maastricht University, Maastricht Graduate School of Governance, Working Paper MGSOG/2007/WP003.

Hansen, Lars Peter, “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, July 1982, 50 (4), 1029–1054.

Harris, John R. and Michael P. Todaro, “Migration, Unemployment and Development: A Two-Sector Analysis,” *American Economic Review*, 1970, 60 (1), 126–142.

Hartog, Joop and Rainer Winkelmann, “Comparing Migrants to Non-migrants: The Case of Dutch Migration to New Zealand,” *Journal of Population Economics*, 2003, 16, 683–705.

Heckman, James J., “Sample Selection Bias as a Specification Error,” *Econometrica*, January 1979, 47 (1), 153–161.

- Holmlund, Bertil and Johnny Zetterberg**, “Inside Effects in Wage Determination: Evidence from Five Countries,” *European Economic Review*, 1991, 35, 1009–1034.
- House, William J. and Henry Rempel**, “The Determinants of Interregional Migration in Kenya,” *World Development*, 1980, 8, 25–35.
- Jacoby, Hanan G.**, “Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra,” *Review of Economic Studies*, October 1993, 60 (4), 903–921.
- Janvry, Alain De, Marcel Fafchamps, and Elisabeth Sadoulet**, “Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained,” *Economic Journal*, November 1991, 101 (409), 1400–1417.
- Johnson, George E.**, “The Structure of Rural-urban Migration Models,” *Eastern Africa Economic Review*, 1971, 3 (1), 21–28.
- Jondrow, James, C. A. Knox Lovell, Ivan S. Materov, and Peter Schmidt**, “On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model,” *Journal of Econometrics*, 1982, 19, 233–238.
- Jones, Patricia and Abigail Barr**, “Learning by Doing in Sub-Saharan Africa: Evidence from Ghana,” *Journal of International Economics*, 1996, 8 (3), 445–466.
- Jovanovic, Boyan**, “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, October 1979, 87 (5), 972–990.
- , “Selection and the Evolution of Industry,” *Econometrica*, May 1982, 50 (3), 649–670.
- Katayama, Hajime, Shihua Lu, and James R. Tybout**, “Firm-level Productivity Studies: Illusion and a Solution,” *International Journal of Industrial Organization*, 2009, 27, 403–413.
- Katz, Eliakim and Oded Stark**, “Labor Migration and Risk Aversion in Less Developed Countries,” *Journal of Labor Economics*, January 1986, 4 (1), 134–149.
- Kessy, Flora**, “Rural Income Dynamics in Kagera Region, Tanzania.” A Report Prepared for the World Bank, Economic and Social Research Foundation (ESRF).

- Kochar, Anjini**, "Smoothing Consumption by Smoothing Income: Hours-of-Work Response to Idiosyncratic Agricultural Shocks in Rural India," *Review of Economics and Statistics*, February 1999, 81 (1), 50–61.
- Konings, Jozef and Patrick P. Walsh**, "Evidence of Efficiency Wage Payments in UK Firm Level Panel Data," *Economic Journal*, May 1994, 104 (424), 542–555.
- Laber, Gene and Richard X. Chase**, "Interprovincial Migration in Canada as a Human Capital Decision," *Journal of Political Economy*, July - August 1971, 79 (4), 795–804.
- LaLonde, Robert J.**, "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, September 1986, 76 (4), 604–620.
- **and Robert H. Topel**, "Immigrants in the American Labor Market: Quality, Assimilation, and Distributional Effects," *American Economic Review*, May 1991, 81 (2), 297–302. Papers and Proceedings of the Hundred and Third Annual Meeting of the American Economic Association.
- **and Robert H. Topel**, *Economic Impact of International Migration and the Economic Performance of Migrants*, Elsevier Science B.V., 1997. pp. 799–850, Chapter 14 in Mark R. Rosenzweig and Oded Stark (editors), *Handbook of Population and Family Economics Volume: 1B*.
- Lanzona, Leonardo A.**, "Migration, Self-selection and Earnings in Philippine Rural Communities," *Journal of Development Economics*, 1998, 56, 27–50.
- Leontaridi, Marianthi Rannia**, "Segmented Labour Markets: Theory and Evidence," *Journal of Economic Surveys*, 1998, 12 (1), 63–101.
- Levin, Carol E., Marie T. Ruel, Saul S. Morris, Daniel G. Maxwell, and Margaret Armar-Klemesu**, "Working Women in an Urban Setting: Traders, Vendors and Food Security in Accra," *World Development*, 1999, 27 (11), 1997–1991.
- Levinsohn, James and Amil Petrin**, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, 70 (2), 317–341.
- Lipton, Michael**, "Migration from Rural Areas of Poor Countries: The Impact on Rural Productivity and Income Distribution," *World Development*, 1980, 8, 1–24.

- Little, I. M. D.**, “Small Manufacturing Enterprises in Developing Countries,” *World Bank Economic Review*, 1987, 1 (2), 203–235.
- Lucas, Jr., Robert E.**, “Life Earnings and Rural-Urban Migration,” *Journal of Political Economy*, 2004, 112 (1), 29–59.
- Lucas, Robert E. B. and Oded Stark**, “Motivations to Remit: Evidence from Botswana,” *Journal of Political Economy*, 1985, 93 (5), 901–918.
- Luke, Nancy and Kaivan Munshi**, “New Roles for Marriage in Urban Africa: Kinship Networks and the Labor Market in Kenya,” *Review of Economics and Statistics*, May 2006, 88 (2), 264–282.
- Lundvall, Karl and George E. Battese**, “Firm Size, Age and Efficiency: Evidence from Kenyan Manufacturing Firms,” *Journal of Development Studies*, 2000, 36 (3), 146–163.
- Maloney, William F.**, “Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transition in Mexico,” *World Bank Economic Review*, 1999, 13 (2), 275–302.
- Mansuri, Ghazala**, “Migration, School Attainment and Child Labor: Evidence from Rural Pakistan,” *World Bank Policy Research Working Paper*, June 2006, 3945.
- McKenzie, David and Hillel Rapoport**, “Network Effects and Dynamics of Migration and Inequality: Theory and Evidence from Mexico,” *Journal of Development Economics*, 2007, 84, 1–24.
- **and —**, “Can Migration Reduce Educational Attainment?: Evidence from Mexico,” *Journal of Population Economics*, 2010. Published online 06 April 2010, <http://www.springerlink.com/content/95ht846m8m306517/fulltext.pdf>.
- **and Marcin J. Sasin**, “Migration, Remittances, Poverty, and Human Capital: Conceptual and Empirical Challenges,” *World Bank Policy Research Working Paper*, July 2007, 4272.
- **, John Gibson, and Steven Stillman**, “How Important is Selection?: Experimental vs Non-experimental Measures of the Income Gains from Migration,” March 2006. University of Waikato, Department of Economics, Working Paper in Economics 3/06.

- McKinsey Global Institute**, “Service Sector Productivity,” October 1992. Washington, D.C., McKinsey Global Institute with assistance from Martin Baily, Francis Bator, Ted Hall, Bob Solow, Eisenhower Center (Columbia University), and McKinsey & Company, Inc.
- McPherson, Michael A.**, “The Hazards of Small Firms in Southern Africa,” *Journal of Development Studies*, October 1995, *32* (1), 31–54.
- , “Growth of Micro and Small Enterprises in Southern Africa,” *Journal of Development Economics*, 1996, *48*, 253–277.
- Mead, Donald C.**, “The Contribution of Small Enterprises to Employment Growth in Southern and Eastern Africa,” *World Development*, 1994, *22* (12), 1881–1894.
- **and Carl Liedholm**, “The Dynamics of Micro and Small Enterprises in Developing Countries,” *World Development*, 1998, *26* (1), 61–74.
- Meeusen, Wim and Julien Van Den Broeck**, “Efficiency Estimation From Cobb-Douglas Production Functions with Composed Error,” *International Economic Review*, June 1977, *18* (2), 435–444.
- Mengistae, Taye and Catherine Pattillo**, “Export Orientation and Productivity in Sub-Saharan Africa,” 2004, *51* (2). IMF Staff Papers.
- Miller, Robert A.**, “Job Matching and Occupational Choice,” *Journal of Political Economy*, December 1984, *92* (6), 1086–1120.
- Mincer, Jacob**, “Family Migration Decisions,” *Journal of Political Economy*, October 1978, *86* (5), 749–773.
- Mitti, Respichius D. and Thadeus G. Rweyemamu**, “A Socio-Economic Profile of Nyakatoke,” January 2001. mimeo.
- Monk, Courtney, Justin Sandefur, and Francis Teal**, “Does Doing an Apprenticeship Pay Off?: Evidence from Ghana,” 2008. The Centre for the Study of African Economies Working Paper Series, paper 288.
- Mortensen, Dale T.**, *Wage Dispersion: Why are Similar Workers Paid Differently?* Zeuthen Lecture Book Series, paperback ed., MIT Press, 2005.

- Moschini, Giancarlo**, “A Non-Nested Test of Separability for Flexible Functional Forms,” *Review of Economics and Statistics*, May 1992, 74 (2), 365–369.
- Munshi, Kaivan**, “Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market,” *Quarterly Journal of Economics*, May 2003, pp. 549–599.
- Nayga, Jr., Rodolfo M.**, “Wife’s Labor Force Participation and Family Expenditures for Prepared Food, Food Prepared at Home, and Food Away from Home,” *Agricultural and Resource Economics Review*, October 1996, 25 (2), 179–186.
- Nickell, Stephen**, “Biases in Dyanamic Models with Fixed Effects,” *Econometrica*, November 1981, 49 (6), 1417–1426.
- and **Sushil Wadhvani**, “Insider Forces and Wage Determination,” *Economic Journal*, June 1990, 100 (401), 496–509.
- Olley, G. Steven and Ariel Pakes**, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, November 1996, 64 (6), 1263–1297.
- Pack, Howard**, “The Substitution of Labour for Capital in Kenyan Manufacturing,” *Economic Journal*, March 1976, 86 (341), 45–58.
- , “Aggregate Implications of Factor Substitution in Industrial Processes,” *Journal of Development Economics*, February 1982, 11, 1–37.
- , “Productivity and Industrial Development in Sub-Saharan Africa,” *World Development*, 1993, 21 (1), 1–16.
- Page, Jr, John M.**, “Technical Efficiency and Economic Performance: Some Evidence from Ghana,” *Oxford Economic Papers*, July 1980, 32 (2), 319–339.
- Pakes, Ariel and Richard Ericson**, “Empirical Implication of Alternative Models of Frm Dynamics,” *Journal of Economic Theory*, 1998, 79, 1–45.
- Petrin, Amil, Brian P. Poi, and James Levinsohn**, “Production Function Estimating in Stata Using Inputs to Control for Unobservables,” *Stata Journal*, 2004, 4 (2), 113–123.

- Quinn, Simon and Francis Teal**, “Private Sector Development and Income Dynamics: A Panel Study of the Tanzanian Labour Market,” February 2008. The Centre for the Study of African Economies (CSAE), Working Paper Series/2008-09.
- Ramachandran, Vijaya and Manju Kedia Shah**, “Minority Enterprises and Firm Performance in Sub-Saharan Africa,” *Journal of Development Studies*, December 1999, 36 (2), 71–87.
- Rankin, Neil, Justin Sandefur, and Francis Teal**, “Learning & Earning in Africa: Evidence from Ghana,” May 2006. mimeo.
- , **Måns Söderbom, and Francis Teal**, “Exporting from Manufacturing Firms in Sub-Saharan Africa,” December 2005. GPRG-WPS-036, Global Poverty Research Group.
- Reardon, Thomas, Christopher Delgado, and Peter Matlon**, “Determinants and Effects of Income Diversification Amongst Farm Households in Burkina Faso,” *Journal of Development Studies*, January 1992, 28 (2), 264–296.
- , **J. Edward Taylor, Kostas Stamoulis, Peter Lanjouw, and Arsenio Balisacan**, “Effects of Non-farm Employment on Rural Income Inequality in Developing Countries: An Investment Perspective,” *Journal of Agricultural Economics*, May 2000, 51 (2), 266–288.
- Roberts, Mark J. and James R. Tybout**, “The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs,” *American Economic Review*, September 1997, 87 (4), 545–564.
- Robinson, Chris and Nigel Tomes**, “Self-selection and Interprovincial Migration in Canada,” *Canadian Journal of Economics*, 1982, 15 (3), 474–502.
- Roodman, David**, “How to Do xtabond2: An Introduction to ”Difference” and ”System” GMM in Stata,” December 2006. Working Paper 103, Center for Global Development.
- Rosenzweig, Mark R.**, “Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries,” *Economic Journal*, December 1988, 98, 1148–1170.
- **and Oded Stark**, “Consumption Smoothing, Migration, and Marriage: Evidence from Rural India,” *Journal of Political Economy*, August 1989, 97 (4), 905–926.
- Salop, Steven C.**, “A Model of the Natural Rate of Unemployment,” *American Economic Review*, March 1979, 69 (1), 117–125.

- Sandefur, Justin**, “Explaining the Trend toward Informal Employment in Africa: Evidence from Ghanaian Manufacturing,” March 2006. mimeo.
- , “On the Evolution of the Firm Size Distribution in an Africa Economy,” 2008, pp. 18–73. DPhil Dissertation, Chapter 2, Department of Economics, University of Oxford.
- Sargan, J. D.**, “The Estimation of Economic Relationships Using Instrumental Variables,” *Econometrica*, July 1958, *26* (3), 393–415.
- Schultz, Theodore W.**, “Investment in Human Capital,” *American Economic Review*, March 1961, *51* (1), 1–17.
- Seiford, Lawrence M. and Robert M. Thrall**, “Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis,” *Journal of Econometrics*, 1990, *46*, 7–38.
- Shapiro, Carl and Joseph E. Stiglitz**, “Equilibrium Unemployment as a Worker Discipline Device,” *American Economic Review*, 1984, *74* (3), 433–444.
- Shiferaw, Admasu and Arjun Bedi**, “The Dynamics of Job Creation and Job Destruction: Is Sub-Saharan Africa Different?,” December 2009. The Institute for the Study of Labor (IZA), Discussion Paper No. 4623.
- Singh, Ram D.**, “Underinvestment, Low Economic Returns to Education, and the Schooling of Rural Children: Some Evidence from Brazil,” *Economic Development and Cultural Change*, April 1992, *40* (3), 645–664.
- Sjaastad, Larry A.**, “The Costs and Returns of Human Migration,” *Journal of Political Economy*, October 1962, *70* (5), 80–93. Part 2: Investment in Human Beings.
- Skoufias, Emmanuel**, “Labor Market Opportunities and Intrafamily Time Allocation in Rural Households in South Asia,” *Journal of Development Economics*, 1993, *40*, 277–310.
- Sleuwaegen, Leo and Micheline Goedhuys**, “Growth of Firms in Developing Countries, Evidence from Côte d’Ivoire,” *Journal of Development Economics*, 2002, *68*, 117–135.
- Söderbom, Måns and Francis Teal**, “Are Manufacturing Exports the Key to Economic Success in Africa?,” *Journal of African Economies*, 2003, *12* (1), 1–29.

- **and** — , “Size and Efficiency in African Manufacturing Firms: Evidence from Firm-level Panel Data,” *Journal of Development Economics*, 2004, *73*, 369–394.
- , — , **and Alan Harding**, “The Determinants of Survival Among African Manufacturing Firms,” *Economic Development and Cultural Change*, April 2006, *54*, 533–555.
- , — , **and Anthony Wambugu**, “Unobserved Heterogeneity and the Relation between Earnings and Firm Size: Evidence from Two Developing Countries,” *Economics Letters*, 2005, *87*, 153–159.
- Solow, Robert M.**, “Another Possible Source of Wage Stickiness,” *Journal of Macroeconomics*, 1979, *1* (1), 79–82. Winter.
- Soubbotina, Tatyana P.**, *Beyond Economic Growth: An Introduction to Sustainable Development* WBI Learning Resources Series, second ed., 1818 H Street, N.W., Washington, D.C. 20433, U.S.A.: The International Bank for Reconstruction and Development / The World Bank, 2004.
- Sparks, Roger**, “A Model of Involuntary Unemployment and Wage Rigidity: Worker Incentives and the Threat of Dismissal,” *Journal of Labor Economics*, October 1986, *4* (4), 560–581.
- Stark, Oded**, “On the Optimal Choice of Capital Intensity in LDCs with Migration,” *Journal of Development Economics*, 1981, *9*, 31–41.
- , “Research on Rural-to-urban Migration in LDCs: The Confusion Frontier and why We Should Pause to Rethink Afresh,” *World Development*, 1982, *10* (1), 63–70.
- **and David Levhari**, “On Migration and Risk in LDCs,” *Economic Development and Cultural Change*, October 1982, *31* (1), 191–196.
- **and J. Edward Taylor**, “Migration Incentives, Migration Types: The Role of Relative Deprivation,” *Economic Journal*, September 1991, *101* (408), 1163–1178.
- **and Robert E. B. Lucas**, “Migration, Remittances, and the Family,” *Economic Development and Cultural Change*, April 1988, *36* (3), 465–481.
- , **J. Edward Taylor**, **and Shlomo Yitzhaki**, “Remittances and Inequality,” *Economic Journal*, September 1986, *96*, 722–740.

- Stiglitz, Joseph E.**, “Rural-Urban Migration, Surplus Labour, and the Relationship between Urban and Rural Wages,” *Eastern Africa Economic Review*, 1969, pp. 1–27. Cowles Foundation Paper 324.
- Teal, Francis**, “The Size and Sources of Economic Rents in a Developing Country Manufacturing Labour Market,” *Economic Journal*, July 1996, *106* (437), 963–976.
- Todaro, Michael P.**, “A Model of Labor Migration and Urban Unemployment in Less Developed Countries,” *American Economic Review*, 1969, *59* (1), 138–148.
- , “Urban Job Expansion, Induced Migration and Rising Unemployment: A formulation and Simplified Empirical Test for LDCs,” *Journal of Development Economics*, 1976, *3*, 211–225.
- Tybout, James R.**, “Making Noisy Data Sing: Estimating Production Technologies in Developing Countries,” *Journal of Econometrics*, 1992, pp. 25–44.
- , “Manufacturing Firms in Developing Countries: How Well do They Do, and Why?,” *Journal of Economic Literature*, March 2000, *XXXVIII*, 11–44.
- UNEP**, “Ministers to Tackle Increasing Air Pollution in Sub-Saharan African Cities,” 2006. unofficial document for Regional Conference on Better Air Quality in African Cities, Division of Policy Development and Law.
- United Nations**, “World Urbanization Prospects - The 2007 Revision,” February 2008. Population Division of the Department of Economic and Social Affairs, ESA/P/WP/205.
- Van Biesebroeck, Johannes**, “Exporting Raises Productivity in Sub-Saharan African Manufacturing Firms,” *Journal of International Economics*, 2005, *67*, 373–391.
- , “Firm Size Matters: Growth and Productivity Growth in African Manufacturing,” *Economic Development and Cultural Change*, April 2005, *53*, 545–583.
- , “The Sensitivity of Productivity Estimates: Revisiting Three Important Debates,” *Journal of Business and Economic Statistics*, July 2008, *26* (3), 311–328.
- Vanzo, Julie De**, “Repeat Migration in the United States: Who Moves Back and Who Moves on?,” *Review of Economics and Statistics*, November 1983, *65* (4), 552–559.

- Velenchik, Ann D.**, “Government Intervention, Efficiency Wages, and the Employer Size Wage Effect in Zimbabwe,” *Journal of Development Economics*, 1997, 53.
- Vijverberg, Wim P. M.**, “Measuring Income from Family Enterprises with Household Surveys,” *Small Business Economics*, 1992, 4, 287–305.
- **and Jonathan Houghton**, “Household Enterprises in Vietnam,” *Policy Research Working Paper*, February 2002, 2773. World Bank, Development Research Group, Macroeconomics and Growth.
- Wadhvani, Sushil B. and Martin Wall**, “A Direct Test of the Efficiency Wage Model Using UK Micro-Data,” *Oxford Economic Papers*, October 1991, 43 (4), 529–548.
- Weerdt, Joachim De**, “Defying Destiny and Moving out of Poverty: Evidence from a 10-year Panel with Linked Qualitative Data from Kagera, Tanzania,” December 2008. mimeo.
- Weiss, Andrew**, “Job Queues and Layoffs in Labor Markets with Flexible Wages,” *Journal of Political Economy*, 1980, 88 (3), 526–538.
- Weiss, Yoram, Robert M. Sauer, and Menachem Gotlibovski**, “Immigration, Search, and Loss of Skill,” *Journal of Labor Economics*, 2003, 21 (3), 557–591.
- Wertheimer, Richard**, *The Monetary Rewards of Migration within the U.S.*, The Urban Institute, 1970. Washington, D. C.
- Windmeijer, Frank**, “A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators,” *Journal of Econometrics*, 2005, 126, 25–51.
- Wooldridge, Jeffrey M.**, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002. Cambridge, Massachusetts.
- World Bank**, “Reshaping Economic Geography,” 2009. World Development Report 2009.
- Yamauchi, Futoshi**, “Are Experience and Schooling Complementary?: Evidence from Migrants’ Assimilation in the Bangkok Labor Market,” *Journal of Development Economics*, 2004, 74, 489–513.
- **and Sakiko Tanabe**, “Nonmarket Networks among Migrants: Evidence from Metropolitan Bangkok, Thailand,” *Journal of Population Economics*, 2008, 21, 649–664.

- Yang, Dean**, “Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants’ Responses to Exchange Rate Shocks,” *Review of Economic Statistics*, November 2006, 88 (4), 715–735.
- Yellen, Janet L.**, “Efficiency Wage Models of Unemployment,” *American Economic Review*, May 1984, 74 (2), 200–205. Papers and Proceedings of the Ninety-Sixth Annual Meeting of the American Economic Association.
- Yezer, Anthony M. J. and Lawrence Thurston**, “Migration Patterns and Income Change: Implications for the Human Capital Approach to Migration,” *Southern Economic Journal*, April 1976, 42 (4), 693–702.
- Zellner, Arnold**, “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,” *Journal of the American Statistical Association*, June 1962, 57 (298), 348–368.
- Zhou, Min**, “Segmented Assimilation: Issues, Controversies, and Recent Research on the New Second Generation,” *International Migration Review*, Winter 1997, 31 (4), 975–1008.