ESSAYS ON LABOUR MARKET FRICTIONS IN DEVELOPING ECONOMIES

SIMON FRANKLIN

Oriel College
Oxford University

A thesis submitted for the degree of
Doctor of Philosophy in Economics
Hilary 2015

Essays on Labour Market Frictions in Developing Economies

Simon Franklin, Oriel College
DPhil in Economics
Hilary 2015

Abstract

This thesis is about imperfections in urban labour markets of three developing countries. I study how physical living conditions place constraints on labour force participation, and increase risks associated with unemployment.

In Chapter One I test for the impact of high search costs on labour market outcomes of job seekers. I use a randomized trial of transport subsidies among youth living far away from the centre of the city in Addis Ababa, Ethiopia. Lowering transport costs increases the intensity of job search and leads to better employment outcomes. Weekly phone call data shows that treatment works to stop job search activity from declining over time. I show that the results are consistent with a dynamic model of job search with cash constraints and monetary search costs. Income from temporary work is used to smooth consumption and pay for the costs of search. I find that subsidies reduce participation in temporary work.

Chapter Two looks at the links between poor housing conditions in slums and market labour supply. I test for the effect of free government housing in South Africa on households, using four waves of panel data and a natural experiment due to the allocation of new housing according to proximity from housing projects. I then use planned but cancelled projects to control for non-random selection of housing project sites. I find that government housing leads to large increases in household incomes from wage work, and increases in the labour supply of female household members. I argue that these results are due to reduced burdens of work in the home of improved housing, especially for women.

In Chapter Three we look at how labour markets respond to large but temporary economic shocks caused by typhoons in the Philippines. We use quarterly aggregate, repeated-cross sectional and panel data to demonstrate robust evidence of downward wage flexibility. Lay-offs do not occur when storms hits, but hours per worker fall. We explain these results with a model of implicit contracts under which risk is shared between workers and firms through wage cuts, but workers are insured against lay-offs so that adjustments in labour demand occur through reductions in hours per worker. Our results are particularly strong for workers in long term contractual relationships in the private sector.

Word count

This thesis consists of approximately 85,094 words. This is calculated using the number of words on page 2 (314) multiplied by the number of pages (271).

Declaration of authorship

The third chapter of this thesis was prepared jointly with Julien Labonne. We collaborated on the conception of the project, the estimation strategies and the writing of the paper. I prepared the geographic and climate data, performed the detailed regression analysis, and developed the theoretical model.

ACKNOWLEDGEMENTS

Firstly, this thesis would not have been possible without the help of my supervisor, Marcel Fafchamps. His advice was always challenging and constructive. In his way, he taught me what it is to do good research and how to ask the right questions. Simon Quinn played a crucial role as a second supervisor. I thank him for his enthusiasm for my ideas, and for encouraging me to visit Ethiopia for the first time in 2011.

I acknowledge funding for fieldwork for this thesis from Magdalen College, Oriel College, OxCarre, and the World Bank. I benefited from presenting the work for this thesis at various conferences with funding from the George Webb Medley Fund. Support from Rose Page and Gail Wilkins at the Centre for the Study of African Economies was absolutely invaluable.

The Ethiopian Development Research Institute provided me support and assistance throughout my time in Ethiopia. There are too many skilled, competent and enthusiastic researchers, enumerators and miracle-workers to list. Girum Abebe is their incredible leader.

Coursemates, colleagues and friends provided encouragement, ideas and support. Joe St Clair, in Ethiopia, and Julien Labonne, in Oxford, deserve particular mention. I could not have got through it, nor got through it so happily, without Lucie Moore.

To my family, my mother, father and sister. Maybe now that it's finished I will tell you about what I've been working on. I'm so grateful for your love and support.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This thesis is about frictions in urban labour markets of developing countries.[1] Frictions are disruptive to the way in which labour markets facilitate matches, determine wages and respond to volatility (Rogerson and Shimer, 2011; Diamond, 1982). While a long literature has looked at market failures in developing country labour markets, this literature has focused on failures arising from risk in agrarian economies and from segmentation between the formal and informal sectors (Rosenzweig, 1988; Behrman, 1999; Fields, 1975). In fact the prevailing assumption seems to be that markets for wage labour in developed and developing urban areas operate in more or less the same ways.[2] Unemployment has received very little attention as a distinct theoretical problem and is even thought to be rare in the formal sense, since there is such an abundance of agricultural and informal sector income opportunities for those without good jobs (Lewis, 1954; Gollin, 2014).[3]

However, unemployment rates in the cities of the developing world are high, especially among the youth (World Bank, 2012). Wage paying employment in non-agricultural jobs is becoming an increasingly important source of liveli-

1 Here I define frictions as restrictions on the matching of unemployed individuals and firms who would otherwise like to hire them. This includes: factors which constrain labour supply for those who would like to work, constraints that make job search costly, and informational asymmetries which make it particularly hard for firms to hire workers.

2 To quote: "Few distinct analytical models specifically targeted in any meaningful way to problems of low-income country urban labor markets have emerged in the literature" (Rosenzweig, 1988). "In my judgement a decade later this conclusion still holds." (Behrman, 1999).

3 Note that considerable attention has been paid to the problem of under- or disguised-unemployment (World Bank, 2009; Lewis, 1954). It is assumed that individuals cannot afford to remain unemployed for any length of time. While my first paper suggests that this is true, it is also appropriate to consider someone unemployed while they find irregular income from temporary work, for as long as they continue to look for alternative work in the formal sector.

hoods as economies and cities continue to change and grow. Yet the risk of unemployment may be particularly salient in these environments.

Cities have great potential to be places of productivity where the concentration of economic activity reduces frictions (Moretti, 2014). Yet rapid urban growth is also placing considerable pressures on labour markets in cities (Harris and Todaro, 1970). Cities may become sprawling and under-serviced such that the benefits of urban agglomeration are mitigated by congestion costs and poor living conditions in slums (Marx et al., 2013). Investments in urban roads, transport systems, denser neighborhoods and well-located housing for the poor are important to ensure that labour markets function efficiently and equitably.

There are many constraints on participants in urban labour markets that are particularly relevant to developing countries. Subsistence constraints are more salient (Lewis, 1954), since households face particularly severe consequences when they experience income shocks. They find ways to mitigate against risk through the use of adjustments in labour supply (Banerjee and Duflo, 2007; Rose, 2001) and migration (Imbert and Papp, 2015). The unemployed have no access to institutions that would help them to find employment, or to welfare or unemployment insurance while they are without work. There is a growing evidence that credit constraints limit the mobility of labour (Bryan et al., 2014; Ardington et al., 2009). And information flows about employment opportunities may be subject to additional frictions in developing country settings (Jensen, 2012). Labour relationships are more often characterised by temporary or casual contracts, labour regulation is often weakly enforced, and women are more likely to be confined to work in the home (Duflo, 2012).

This thesis focuses on how risk, subsistence constraints and urban living conditions impact labour markets in three distinct settings.

The first paper tests whether young job seekers who live far from jobs are constrained in their ability to search for jobs by their place of living and high search costs. I ran a randomized controlled trial of transport subsidies in Addis Ababa, Ethiopia. In addition to face-to-face baseline and endline surveys, I collected a high-frequency phone survey. I find that individuals who were given the subsidies search more intensively and find better jobs.

I use a dynamic job search model to illuminate the mechanisms through which cash constraints and high search costs can make search a risky activity.[4] Job seekers give up searching when their savings run low, even when they have enough money to pay the costs of search at least once. The use of the high-frequency phone data allows me to analyze the trajectory of treatment effects over time and compare these to the predictions of the model. I show that this model of job search successfully replicates the main empirical results.

The paper makes three main contributions. Firstly, it adds to a literature on the impacts of cash constraints on labour market decisions. While a growing literature looks at constraints on rural-urban migration (Bryan et al., 2014; Ardington et al., 2009), studies on the impacts of cash constraints on job search and employment are predominately from developed countries (Chetty, 2008; Card et al., 2007).[5] I show that search costs create frictions in urban labour markets of a large developing country.

Secondly the paper contributes to a literature that links place of living to labour outcomes. It becomes more costly to search for work the further one lives from the city centre. In this way the results have important implications for how new opportunities for jobs are shared across space in cities. Certain individuals may be locked out of opportunity by virtue of where they live. This is the first experimental evidence, to my knowledge, on the spatial mismatch hypothesis in a developing country context (Kain, 1992; Kling et al., 2007).

4 I calibrate and simulate the model in Matlab.

5 My setting differs fundamentally from these studies because monetary search costs cause job seekers to remain unemployed for longer, rather than exiting unemployment too quickly.

Finally the paper contributes to a literature which examines the role of informal and temporary work in allowing prospective employees to "queue" for jobs (Lewis, 1954; Harris and Todaro, 1970). Job seekers in my setting cannot afford to remain unemployed for long, and must take other forms of less desirable work in order to search. These jobs are not predominately in the informal (self-employed) sector, but include casual or daily labour, or other temporary wage paying jobs. I contribute to this literature by providing evidence for how temporary income is used to smooth consumption and pay the costs of search. My results show that alleviating cash constraints can reduce participation in these forms of temporary work in the short run.

Chapter Two looks at the impact of poor housing conditions on employment outcomes of slum dwellers in South Africa. The government has provided free housing to over 3 million households in the last 20 years. This is an important large scale policy which has transformed the urban landscape of South Africa, and warrants rigorous evaluation. I causally identify the impacts of the programme by exploiting a natural experiment in the way in which units in new housing projects were allocated to households.

I develop a unique instrumental variables estimator based on household proximity from government housing projects to predict who gets a house, using GIS data that I collected from the municipal government. I use these instruments, in conjunction with 4 waves of panel data on households from informal settlements in Cape Town to estimate the impact of government housing on incomes and labour supply. In addition I use projects that were planned, but never built, to control for non-random selection of project locations. I find that improved housing leads to increases in household earnings from wage employment, and increases employment rates among females.

I contribute to a literature that shows that poor housing conditions can negatively impact a number of long term outcomes. I add to the existing evidence by showing that constraints on labour supply are an additional channel through which housing can act as a poverty trap (Marx et al., 2013). Policies that try to deal with the conditions of informal housing should take into account the effects that these policies have on the labour outcomes of recipients.

Secondly, I look for mechanisms for these results and argue that they are driven by labour saving technology that removes burdens of work in the home on female household members (Dinkelman, 2011). Housing also reduces feelings of insecurity at home. This could allow households to substitute time away from protecting their homes from crime, toward work in labour market (Field, 2007).

The third paper uses typhoons to identify large shocks to labour markets in the Philippines.[6] We look at how labour markets adjust to shocks, and find little impact on employment. Instead we find considerable evidence of nominal wage flexibility. Downward wage adjustments prevent lay-offs and allow the labour market to clear. We improve on the existing literature by using aggregate, repeated-cross sectional and individual panel data to argue that the results are driven by a reduction in wages for individual workers at the same jobs. We find no evidence that the results are driven by employment composition or labour supply changes. We contribute to the literature on adjustments to shocks in developing countries by presenting results that apply to both rural and urban labour markets.[7]

Typhoons are a source of considerable risk to both firms and employees in the Philippines, which impact on the functioning of labour markets.[8] We argue that implicit contracts in long term labour relationships allow risk sharing between workers and firms, in an outcome whereby workers are insured against layoffs

6 This is joint work with Julien Labonne. We use our own geo-coded data on storm tracks to develop a dataset of municipal level wind-speeds in each quarter.

7 See Kaur (2014) for a recent study on labour market adjustment in a rural setting, as well as an overview of the existing literature.

8 In light of evidence that such disasters are likely to become more common in other developing countries as a result of changes in the global climate, it is crucial to understand how labour markets react to shocks of this size.

(work-sharing among employees) but endure wage cuts during shocks (risk sharing with firms). Surprisingly, our empirical results are particularly strong for workers with long term contracts, contrasting with the literature that has used implicit contracts as an explanation for wage *rigidities* (Hall and Milgrom, 2008; Holmstrom, 1983). We adapt theoretical models of Azariadis (1975) and Rosen (1985) to explain our findings.

While these long term arrangements exhibit remarkable flexibility that prevents disruption to employment levels, they also imply a cost to productive efficiency as a second best insurance outcome (Holmstrom, 1983). Mechanisms that would allow firms to better insure against storm damage, and unemployment insurance to protect workers would lead to more efficient outcomes.

In summary, the three papers in this thesis focus on how aspects of daily life in developing countries have an effect on the functioning of labour markets. I show that search costs, location and conditions of housing, and weather shocks, all impact labour outcomes. Risk plays an important role in preventing agents from making optimal labour market decisions. In each paper, these constraints are underscored by failures of credit and insurance markets, absence of state welfare or unemployment insurance, or inadequate provision of public infrastructure in urban areas.

BIBLIOGRAPHY

Ardington, C., Case, A., and Hosegood, V. (2009). Labor supply responses to large social transfers: Longitudinal evidence from South Africa. *American economic journal. Applied economics*, 1(1):22–48.

Azariadis, C. (1975). Implicit contracts and underemployment equilibria. *The Journal of Political Economy*, pages 1183–1202.

Banerjee, A. and Duflo, E. (2007). The economic lives of the poor. *The Journal of Economic Perspectives*, 21(1):141–167.

Behrman, J. R. (1999). Labor markets in developing countries. *Handbook of labor economics*, 3:2859–2939.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5):1671–1748.

Card, D., Chetty, R., and Weber, A. (2007). Cash-on-hand and competing models of intertemporal behavior: New evidence from the labor market. *The Quarterly Journal of Economics*.

Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment ensurance. *Journal of Political Economy*, 116:1197–1197.

Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *The Journal of Political Economy*, pages 881–894.

Dinkelman, T. (2011). The effects of rural electrification on employment: new evidence from South Africa. *American Economic Review*, 101(7):3078–3108.

Duflo, E. (2012). Women's empowerment and economic eevelopment. *Journal of Economic Literature*, 50(4):1051–1079.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in Peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Fields, G. S. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. *Journal of development economics*, 2(2):165–187.

Gollin, D. (2014). The Lewis Model: A 60-year retrospective. *Journal of Economic Perspectives*, 28(3):71–88.

Hall, R. E. and Milgrom, P. R. (2008). The limited influence of unemployment on the wage bargain. *American Economic Review*, 98:1653–1674.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American Economic Review*.

Holmstrom, B. (1983). Equilibrium long-term labour contracts. *Quarterly Journal of Economics*, 98(1983):23–54.

Imbert, C. and Papp, J. (2015). Labor market effects of social programs: Evidence from India's employment guarantee. *American Economic Journal: Applied Economics*, (Forthcoming).

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics*, 127(2):753–792.

Kain, J. F. (1992). The spatial mismatch hypothesis: three decades later. *Housing policy debate*, 3(2):371–460.

Kaur, S. (2014). Nominal wage rigidity in village labor markets. *Working Paper, Columbia University*, (October):1–66.

Kling, J. R., Liebman, J., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75:83–119.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2):139–191.

Marx, B., Stoker, T., and Suri, T. (2013). The economics of slums in the developing world. *Journal of Economic Perspectives*, 27(4):187–210.

Moretti, E. (2014). Cities and Growth. *International Growth Centre*.

Rogerson, R. and Shimer, R. (2011). Search in macroeconomic models of the labor market. *Handbook of Labor Economics*, 4:619–700.

Rose, E. (2001). Ex ante and ex post labor supply response to risk in a low-income area. *Journal of Development Economics*, 64:371–388.

Rosen, S. (1985). Implicit contracts: A survey. *Journal of Economic Literature*, 23(3):1144–1175.

Rosenzweig, M. R. (1988). Labor markets in low income countries. In Chenery, H. B. and Srinivasan, T. N., editors, *Handbook of Development Economics*, volume 1, pages 713–762, Amsterdam, The Netherlands. North Holland.

World Bank (2009). African development indicators 2008/09: Youth and employment in Africa: The potential, the problem, the promise. Technical report, Washington, DC.

World Bank (2012). *World Development Report 2013: Jobs*. World Bank, Washington, D.C.

1

LOCATION, SEARCH COSTS AND YOUTH UNEMPLOYMENT: A RANDOMIZED TRIAL OF TRANSPORT SUBSIDIES IN ETHIOPIA

I acknowledge funding from the World Bank and the OxCarre. Thank you to Forhad Shilpi and Marcel Fafchamps for input, oversight and guidance of the project. Helpful input to the design and implementation of this experiment was provided by Girum Abebe, Stefano Caria, Simon Quinn and Paolo Falco, to whom I am grateful. I received vital institutional support from the Ethiopian Development Research Institute in Addis Ababa. Thank you to helpful seminar participants and discussants at the CSAE Conference and Workshop, PacDev, EDePo and NEUDC, in particular James Fenske, Kate Vyborny, Orazio Attanasio, Julien Labonne, Simon Quinn, Kate Orkin, Robert Garlick, Erika Deserranno, Clement Imbert, and Joe St Clair.

Abstract:

Do high costs of search affect the labour market outcomes of young job seekers living far away from the centre of cities? I randomly assign temporary and non-fungible transport subsidies to unemployed youth living in spatially dislocated areas of Addis Ababa, Ethiopia. Lowering transport costs increases the intensity of job search (during and after treatment), and increases the likelihood of finding permanent employment by 6 percentage points in the short run. Analysis of weekly phone call data show that search activity declines over time but the subsidies prevent this from happening in the treatment group. The subsidies reduce participation in temporary and informal work, suggesting an important role for alternative sources of income to support job seekers. I explain these results with a dynamic model of job search with savings, cash constraints and monetary search costs. The predictions of the model are quantitatively consistent with the estimated impacts on increased job search activity, and in turn the rates at which jobs are found. These results suggest that the cost of transport in large cities can lead to frictions in the matching of firms and workers and reinforce spatial inequality.

1.1 INTRODUCTION

Young people in African cities often spend long periods of time in unemployment while trying to find satisfying livelihoods. This has long been a concern of researchers and policy-makers (Harris and Todaro, 1970; World Bank, 2012). Cash constraints create frictions in labour markets if job-seekers cannot optimally invest in finding employment (Card et al., 2007; Bryan et al., 2014). These constraints are compounded when transport costs make job search particularly expensive for those living far away from jobs in sprawling cities.

As African economies expand, the growth of cities is expected to create thicker labour markets (Marshall, 1890) and allow workers and firms to make better matches faster, and reduce the time spent in unemployment (Moretti, 2014; Puga, 2010; Wheeler, 2006). However the rapid growth of cities has often not come with corresponding investments in infrastructure and well located housing stock, leaving an increasing number of people living on the outskirts of cities, far away from access to employment opportunities.[1] The resulting increase in the costs of search could inhibit the creation of new and higher quality jobs (Pissarides, 2000; Acemoglu and Shimer, 1999) and leave particular individuals locked out of employment opportunities.

This paper studies whether young job seekers who live far from the centre of Addis Ababa are constrained in their ability to find jobs by their place of living. A randomized controlled trial of transport subsidies tests directly for an impact of search costs and location on job search outcomes. The subsidies were non-fungible (they could only be used for transport to the city centre) and provided for two days of travel for a pre-announced period of time (12 weeks). Subsidies covered entirely the costs of transport to the city centre but no more. In this way the subsidies exogenously reduced the costs of search *without* introducing

1 This is especially true in cities where policy has made it particularly hard for the poor to live in the centre.

wealth effects or reducing the costs of commuting to work in the long run. Since the treatment acts only to bring individuals living far away to a more a equal footing with those living closer in, the results should be interpreted as evidence both for the existence of search frictions and for inequality in access to jobs due to living location.

I find that individuals spend more time searching for work, and are more likely to have applied for jobs, and to have visited the central job boards during all weeks of the study. They are also less likely to be engaged in forms of temporary or casual work in the weeks when they are receiving treatment, suggesting that they substitute away from taking informal work to concentrate on finding good jobs. By the end of the treatment period they are significantly more likely to being working in good jobs.

The subsidy programme is evaluated using a detailed baseline and two endline surveys (4 and 10 months after baseline). A phone call survey was also conducted, whereby respondents were phoned every week for 3 months, and asked a series of short questions about labour market outcomes. I track the trajectories of search intensity over all weeks using this high-frequency data, and look at the timing of changes in employment outcomes.[2]

I used a two sample approach to compare and validate the impact of these subsidies on two different, but both policy-relevant sub-populations. The first (referred to as the *city* sample) surveyed unemployed people found at home in randomly selected slum areas of Addis Ababa. Most of the respondents in this sample were unskilled, worked in the informal sector, and were unqualified for white-collar jobs. The second sample (referred to as the *board* sample) was taken from individuals found at the main vacancies boards in Ethiopia. They were relatively well-educated, active job seekers, who aspired to highly sought-after

2 There is very little evidence from high-frequency data on employment outcomes. For recent evidence from the US see Krueger and Mueller (2011).

professional jobs. I argue that the latter group are those more likely to self-select into participating in youth employment programmes.

Four months after baseline, shortly after the transport treatment ended, treated individuals in the *board* sample are 7 percentage points (relative to a control mean of 19%) more likely to have permanent work. They were not more likely to be working overall, but treated individuals had jobs of higher quality, in higher skilled sectors, and that were more likely to be located in central parts of the city.

By contrast, treated individuals in the *city* sample were not more likely to have found permanent jobs, but the treatment improved their labour market outcomes in other ways. They were more likely to be working (by about 8 percentage points relative to a control mean of 46%) at the endline, they had jobs of higher average quality, and were less likely to be working as day labourers. Furthermore these treatment effects are at least partly persistent. While there is some catch up, the control group among the *board* sample are still less likely, by 3.5 percentage points, to have permanent work 10 months after the baseline.

To explain these results I develop a dynamic discrete-time model of risk averse and cash-constrained job seekers, who must choose whether or not to search for work in each period, and, conditional on that decision, how much of their remaining savings to consume. Job search is costly in monetary terms, it enters the utility function directly such that the marginal utility-cost of search is higher for someone with lower savings. In this model search is risky: the future value of remaining unemployed, after failing to find a job, falls dramatically when someone has very low savings and has to spend a significant portion on searching for work. The model implies a single crossing point in the value of searching versus not searching so that workers with savings above this amount find it optimal to search. Individuals starting with wealth above this crossing point search actively and run down their savings until they no longer find it optimal to search in every period.[3]

I estimate this model numerically, for a wide range of parameter values. I am able to solve for (1) the critical value of savings required to search, (2) the number of weeks it takes a (relatively) wealthy individual to run down savings to that critical point, and (3) the proportion of workers who search for work in steady-state. I find that these predicted outcomes are largely consistent with their empirical moments estimated for the control group in my sample. I then simulate the effect of reduced transport costs on search and consumption activities, and so estimate the effect of treatment on each of the three model predictions mentioned above. The model predicts that lowering transport costs reduces the critical value at which one gives up search, reduces the time taken for job seekers to run down their savings, and results in a higher proportion of individuals searching for work in the steady-state.

The predictions of this cash constraints model are consistent with the estimated impacts of search costs on job search. The high-frequency data shows that the treatment effect on search exhibits a particular trajectory over time. In the control group the percentage of respondents searching for jobs declined over the course of the study as respondents gave up job search and fell into unemployment, or took up temporary jobs to make ends meet. Three key facts emerge. Firstly, the subsidies seem to have prevented respondents from *giving up* job search over time as they ran down their savings, or from skipping weeks of search once they no longer had the savings to search in every week. Secondly, treated individuals are still more likely to be searching *after* the treatment has ended, confirming a prediction of the model that the subsidies allowed them to hold on to savings and thus to search for longer. Finally I show that the treat-

3 Below that level of savings job seekers save up for job search.

ment effects of the subsidies are particularly strong among the relatively poor and cash constrained.

By contrast I find no impact of the treatment on reservation wages, perceptions of the job market, or aspirations. While I cannot rule out that learning effects or other psychological mechanisms might play a role in explaining the results, the model of monetary search costs and cash constraints better explains the trajectory of impacts over time. Qualitative interviews with a number of young job seekers confirm that the treatment did not introduce the job seekers to a new search technology: almost everyone had used the job boards at least once before. In addition, I test for the impact of the phone calls on job search behaviour or endline job outcomes, by using a "pure control" group who received neither phone calls nor transport subsidies. I find no impact of the calls on job search behaviour or labour outcomes, which suggests that results are not driven by Hawthorne effects.[4]

These findings contribute to the literature in a number of ways. A large literature investigates the relationship between cash constraints and job search (Danforth, 1979; Chetty, 2008; Card et al., 2007; Acemoglu and Shimer, 1999). In these models unemployment insurance plays a role in allowing job seekers to smooth consumption while in unemployment.[5] I contribute to this literature with a model in which search costs are monetary and can be subsidized directly. Existing work generally assumes utility functions that are separable in consumption and the (dis)utility of search.[6] In my model, individuals who are particularly poor or running out of savings cannot simply search more. Rather than increas-

4 The main results are robust to using the two control groups pooled together, or using each of the two control groups separately.

5 Indeed the youth's inability to smooth consumption during long spells of unemployment is one of the concerns motivating this paper. Without adequate family support, young people can suffer extended periods of long periods of poverty in unemployment, and can have their life aspirations postponed if not severely blunted (Mains, 2013).

6 See Danforth (1979) for a discussion of this issue.

ing the rate of exit from unemployment, binding cash constraints make it less likely that a job seeker will find a permanent job.

Monetary search costs are a likely to a fitting assumption for many developing countries where the poor spend high proportions of their earnings on transport. This setting, and my results, are most similar to literature looking at migration decisions in developing countries. Migration, like job search, may entail costs that are too high or risky to pay, especially for cash constrained individuals. In this context, providing subsidies to migration (Bryan et al., 2014) or increased liquidity (Ardington et al., 2009) allow young people to migrate to cities, which greatly increases employment and household earnings.[7]

The finding that search costs are creating frictions in labour markets suggests that reducing the costs of transport and improving access to information about jobs could have impacts on access to employment. Other studies have found relatively weak impacts of active labour market policies (Betcherman et al., 2004; Groh et al., 2012; Ibarraran et al., 2012). However, this experiment was conducted on a small scale. These effects could be partly displaced in general equilibrium if reducing search costs improves labour market outcomes at the expense of other job seekers (Crépon et al., 2013).

Still the costs of search fall are likely to fall particularly hard on individuals who live further away from the city since the costs of acquiring information about jobs increases with distance (Ihlanfeldt, 1997). Individuals who suffer most acutely from a lack of access to cash to cover the costs of search are also likely to be particularly constrained. In this sense my results provide evidence for the spatial mismatch hypothesis (Kain, 1992; Holzer, 1991; Zenou, 2009), which

7 Jensen (2012) and Beam (2014) provide evidence on how information about jobs, in a rural context, can increase employment rates in urban jobs. However the literature has not addressed how rural-urban migrants, and other poor individuals living on urban peripheries, search for employment *within* cities, and the constraints that they face in doing so. There is some evidence from developing countries which suggests that freeing up certain physical constraints in relation to housing (Field, 2007; Franklin, 2012) and access to electricity (Dinkelman, 2011) allows the poor to increase labour supply. One of the mechanisms suggested for these effects is that individuals are constrained in their ability to search for jobs.

has been applied and confirmed empirically in the context of large cities in the United States (Kling et al., 2007; Phillips, 2014) but, to my knowledge, never rigorously tested in developing country context until now (Banerjee et al., 2007; Verick, 2011). My results and theoretical model of cash constraints contribute to this literature by explaining the mechanisms through which search costs leave certain individuals at a distinct disadvantage in the labour market.

Finally my results shed light on the role of the informal sector in providing livelihoods for urban youth. The segmentation of African labour markets between formal and informal sectors has been well studied by economists, since at least Lewis (1954) and later Harris and Todaro (1970) and Fields (1975). Debates have raged about whether the informal sector provides a temporary livelihood for those queueing for better work (Serneels, 2007) or if it is a vibrant entrepreneurial sector in its own right (Maloney, 2004). I find evidence that young people use temporary work as means to finding better work, with "planned separations" from these temporary jobs (Browning et al., 2007). I show that improving access to formal sector employment opportunities (by reducing search costs) reduces labour supply to forms of temporary and casual work.[8] My theoretical model shows that reducing cash constraints by providing job seekers with regular income would diminish the risks of unemployment and allow them to search more intensively, while also mitigating the need to take up undesirable forms of temporary work.[9]

The paper proceeds as follows: Section 1.2 discusses the setting and design of the experiment, while Section 1.3 discusses the data and randomization. The results turn first to the impact of the subsidies on employment outcomes at

8 An extensive literature on labour in Ethiopia (Mains, 2013; Serneels, 2007; Haile, 2005), the descriptive data collected for this paper, as well as anecdotal evidence from field work, all suggest that most young Ethiopians still aspire to wage employment in the formal sector. Public sector work is sought after, but not to the extent that it once was. Formal employment, particularly permanent jobs, provide surety, security and social prestige.

9 Evidence on high-turnover for factory jobs in Ethiopia (Blattman and Dercon, 2015) lends support to this idea, in a very similar context: workers seem to have taken low skill jobs due to cash needs rather than aspiring to keep them long term.

endline only in Section 1.4. Section 1.5 looks at the trajectory of the impacts on job search over time using the phone survey. The theoretical Section 1.6 explains these impacts on job search with a model, which is used to generate further predictions, which are tested in turn in Section 1.7, which also investigates other potential mechanisms. Section 1.8 concludes.

1.2 SETTING AND EXPERIMENT

Addis Ababa is an ideal setting to study the effects of urban growth and transport costs on youth labour markets. The city's population has been doubling nearly every decade for last 40 years, and is now estimated to be 4.5 million, and is estimated to grow to 12 million by 2024 (UN-Habitat, 2005). Addis Ababa has been the major destination city for rural-urban migration in Ethiopia, and is one of the most rapidly growing cities in Africa as a result. Many of the new migrants can not access well positioned land in the centre of the city, and long term residents are also being forced out of the inner city slums to make way for new development.[10]

The labour market appears to suffer from many frictions, many of which are induced by distance. Rates of youth unemployment are high in Addis Ababa, at 28% in estimates of Broussard and Teklesellasie (2012). Yet firms still often express frustration at their inability to find the right skilled candidates. The secondary and tertiary education system has expanded enormously in recent years, increasing the supply of labour of formal sector work, but making it harder for firms to distinguish among applicants.

For job seekers the costs of gathering information about jobs are high, and transports costs often comprise a high proportion of their weekly expenditures.

10 Compensation is usually poor and many of those displaced are suffering from having to move to dislocated areas where they no longer have access to their social networks and business links in the center of the city. Existing research documents the loss of income, and transport related problems of those living in worse locations within the city (Yntiso, 2008).

The majority of white-collar jobs are found on job boards, located in the center of the Addis Ababa. For job seekers living far away from the city center, finding a job for the first time or after a spell of unemployment is difficult, especially for those without savings or financial support. Many college graduates spend years looking for their first job (Serneels, 2007).

Young people search for work with limited budgets and savings (and no welfare support or unemployment insurance), and often eed to take up forms of temporary work to make ends meet, while looking for better, permanent work. These are jobs like casual labour or self-employment and work for the family, usually in and around their local areas. The better jobs are usually located in the centre of the city and require formal applications.

The costs of searching for a good job are not evenly distributed: they are considerably higher for those living far from the city centre, those that are particularly poor or find it hard to find other sources of income to support themselves, and new migrants who do not have access to the social networks that might otherwise allow to access employment opportunities. Thus some young people are at a disadvantage in their ability to share in the continued growth of Addis Ababa's economy.

1.2.1 *Experimental design*

I test for the relationship between search costs and labour outcomes by providing a cash transfer that could only be used for travel. Individuals were given money to cover the costs of the transport if they arrived to collect it at designated spot in the center of the city. The amount given out was enough to cover the costs of return trip from their place of living to the centre. We budgeted enough money to make the trip by mini-bus which is the preferred mode of transport because of its flexibility and regularity.[11] The average one way trip to the centre took 33 minutes by mini-bus.

The subsidy amount was tailored to the distance an individual travelled; using the transport costs for a trip from each respondents place of living, using the current fares in Addis Ababa, as surveyed by the enumeration team. The modal amount given was 15 birr for a return trip, or just less than $1 per day.[12] No one received less than 12 birr or more than 20 birr. There was little or no spare cash left over after covering the costs of transports. The money was collected from a makeshift kiosk near the main bus terminal and transport hub in the centre of Addis Ababa, which is close to some of the main job boards in the city.

The sample was assigned to treatment and control groups randomly, with the sample split into three groups: the treatment group who received the subsidies and were also called in every week, and two treatment groups; one that received the same weekly phone calls, and one that was not called at all.[13] This allows me to test both for an impact of the subsidies on trajectories over time with the phone data, and test for an impact of the phone calls at endline. Furthermore, treated respondents were randomly divided into those receiving the subsidies for 8 weeks and those receiving them for 11 weeks.

Immediately after the completion of the baseline data collection the sample was assigned to treatment and control groups for the purposes of the experiment. Randomization was done by stratifying the sample by a number of different baseline covariates, including gender and education. I followed the standard blocking procedure as suggested by Bruhn and McKenzie (2009). Within these strata, 30% of individuals were assigned to both the transport and the calls only

11 58% of respondents said their main mode of travel was in a mini-bus, 38% said that they used the large yellow government buses. These buses are cheaper but unreliable and uncomfortable. A negligible number used other modes of transport such as walking or getting lifts with acquaintances with cars. These mini-buses are similar to those used in many African countries, an overview of the industry can be found in Kumar and Barrett (2008).

12 The US dollar - Birr Exchange stood at around 18 Birr to $1 at the beginning of the experiment.

13 No one was assigned to just the transport treatment without getting the phone calls.

groups. The remaining 40% were designated as pure controls. A more detailed discussion of the variables used to block the randomization is given in the data description section. Figure A.2 in the appendix gives an overview of the randomization design and timeline. 551 respondents received the phone calls, of them a further 255 were offered the transport treatment. 326 were not contacted again until the endline.

The treatment group began to receive the subsidies one week after the end of the baseline, which will be referred to as *Week 1* throughout the rest of the paper. Phone call surveys, which are described in more detail in section 1.3.2 were also begun in that week. The treatment group were clearly informed of the nature of the subsidies, that they had been randomly selected, and for how long they would be to collect them. They could collect the subsidies no more than twice a week and only before midday on any week-day, by showing an identification document and signing for the money.[14]

In the week before they were last allowed to collect the money, they were phoned and reminded that they would no longer be receiving the transport in the next week. The last respondents received their money in week 11 of the study. The last phone calls of the study were completed in week 12. In total respondents could collect the money *up to* 22 times.

1.2.2 *Transport Costs in Perspective*

The cost of transport to and from the centre of the city seems small, at less than \$1 a day, but it is a substantial cost to pay for unemployed youth in Addis Ababa. Reported expenditure from the baseline survey for this experiment confirms that search costs are high. A *single* trip by mini-buses costing 9.50 Birr represents 12% of median weekly expenditure for individuals in the sample, which was just 80

14 This was designed to limit the use of the transport subsidy for recreational use, to make it useful for job seekers who would come in early to see the new job postings, and apply for jobs in the afternoon, but not individuals who wanted to take advantage of evening entertainment in the city center.

Birr ($4). Weekly transport costs were on average 25 birr, or 20% of median total expenditure.

Job seekers found money to pay for these costs and other expenditure items from difference sources. Some took casual labour jobs which pay about 200 Birr a week on average. Others (about 50% of the sample) received money from their family; those that did received about 150 Birr per week on average. So the cash provided by the intervention, of up to 30 Birr per week, constitutes a significant transfer for many of the respondents.

This subsidies do not necessarily cover all of the costs of job search. Respondents reported that there are many other costs associated with search, such as paying for printing and photo-copying, buying clothes for interviews, renting or buying newspapers with information about vacancies, and sometimes paying firms to make applications.

Individuals who were offered the full 11 weeks of the program had the option to collect up to 330 Birr ($17) over the course of the study. If a respondent collected the subsidies at every opportunity, the monthly subsidy was worth about 10% of the monthly salary of a good permanent job, or 15% – 20% of some of the lower paid, informal jobs.

1.3 DATA

Here I describe the data collection and the randomization procedure used in the experiment. I discuss the sampling strategy used for surveying, report on rates of attrition from the survey, and show tests of balance on baseline covariates after randomization.

Figure A.2 in the Appendix provide an overview of the timeline of the project. The experiment unfolded as follows: the baseline was conducted, finishing in week 0 of the study (10 April 2013), immediately after which individuals were assigned to treatment and control groups. Both the transport subsidies and the weekly phone calls began in the week after after baseline, week 1, and continued until week 12 of the study. Three weeks after the end of the phone call and transport treatment the endline survey was conducted (week 16). Respondents were interviewed in a random sequence at the endline survey, in order to avoid correlation between individual characteristics and time effects.

1.3.1 *Sampling Strategy*

This study uses a two sample approach: I study the effects of transport subsidies in two representative but distinct populations in Addis Ababa. This was done to verify the validity of the results across two different population relevant sub-populations, and to compare the nature of the impacts across heterogeneous individuals. This allows me comment on which individuals would most benefit from the subsidy.

From here on, I refer to the one sample as the *city* sample and the other as the *board* sample. The two samples, taken together without dividing respondents in this way, will be referred to as the *pooled* sample. The distinction between the two samples is central to the interpretation of the results because they differed in terms of the margins at which their employment outcomes could be improved.

Screening: Both samples comprise of men and women aged 18-30 who able and available to start a new job in Addis Ababa in the next 2 weeks. Individuals who had some kind of temporary work were included, but those who had no interest in taking a job other the one they already had were not. All individuals were screened on their place of living: only individuals living in neighbourhoods least 5km away from the center of Addis Ababa were included. See the map in Figure A.1 in the Appendix for an idea of the layout of Addis and the radius outside of which the sample was drawn. Individuals in the sample live, on average, 6.8km as the crow flies (sometimes considerably further by road) from the city center where the transport money was collected.

The individuals making up the two samples, both screened for eligibility, were found in the following ways:

City sample: This sample was randomly drawn by going door-to-door in small enumeration areas around the city. Addis Ababa is made of 10 subcities. Four of these, contained within a 5km radius from the centre of the city, were not included in the sample. The chosen sampling areas were stratified to be representative of the remaining 6 subcities located on the periphery of the city. led by selecting two Kebeles from that subcity.[15] Two enumerator teams then moved outward in different directions from the center of the chosen Kebeles, surveying about 60 individuals per Kebele. The survey sites are marked in Figure A.1. Respondents were interviewed at home, enumerators returned for interviews if they were not around at the time of the first visit.

Board sample: The board sample was drawn by randomly approaching individuals who were gathered in the areas around the job boards in the center of Addis Ababa. Although they were all interviewed in the center of the city, these respondents were screened on their place of living, all of them lived in same subcities used in the sampling for the city sample, ensuring that they lived at least 5km away from the center. Since they were all at the job boards at the time of screening, they were almost all, by definition, job seekers, and fit the screening criteria outlined above.

I find impacts of the subsides on both samples, but the two samples respond at very different employment margins: this difference reflects the different backgrounds of the two samples, and the employment opportunities that are available to them. While the characteristics of the *city* sample were similar on average to young people in representative survey data, the *board* sample is clearly a selective group. Notably, they were far more likely to be highly educated, they were more likely to be recent migrants from smaller towns or rural areas. Represen-

15 With a population of over half a million, this subcity is more populous than the next biggest subcity by more than 50%.

tative data from Addis Ababa in 2012 suggests that 22% of the same age cohort had some kind of post-secondary education, about 10% had university degrees. The corresponding figures for the board sample is 72% and 43% respectively. At endline they were far more likely to have permanent jobs, which the *city* sample were very unlikely to find.

I consider my two samples to be representative of two sizeable and policy relevant sub-populations: The *city* sample are those who are withdrawn from the formal labour market and could be induced to (re-)enter by lowered, while the *board* sample represents the group of individuals who are most likely to be part of the growing formal and high-productivity sector in Ethiopia, but may give up search too soon, or become discouraged by their inability to find a good match in the market. The *board* sample started out searching actively, but stood a high chance of becoming discouraged or running out of the funds required to keep pursuing employment. They are also a group most likely to be aided by youth program and active labour market policies initiated by governments or donors, since they would rely on youth to self-select into these programs.

1.3.2 *Phone Survey*

I used a weekly phone call survey to measure the trajectories of job search, employment, and treatment effects over time. These trajectories are used to identify mechanisms through which the treatment effects impact job search, by fully accounting for job seekers activities during the time of treatment. As described in the experimental design, I also randomly assigned some individuals in the control group to not receive the phone calls. In all 524 individuals were assigned to the phone call survey, and 4,510 interviews were conducted over 11 weeks, an average of just over 400 individuals contacted each week, and each respondent contacted on average 10.4 times.

The phone calls were short: they would take between 2 and 4 minutes to complete. The questionnaire was short, giving only a handful of measures that can be used in the analysis of the phone call data. While this restricts the detail of the investigation that can be conducted, it has the advantaged of pre-committing me to testing the significance of just a few major outcomes. These are the outcomes that I analyse in detail throughout the paper, using more detailed endline surveys to investigate further where necessary. Most importantly, the only measure of job quality used in the survey was whether the job was a permanent job- this is the main outcome analysed at endline.

Since this outcome receives special attention throughout the analysis, it is useful to note that it was not hand-picked from a range of different quality measures in the phone survey.[16]

1.3.3 *Tests of Balance*

Table A.4 in the Appendix presents test for balance on variety of job market outcomes, focusing on the main employment outcomes that will be used to test for impact, and that were measured in the phone survey, and then respondent characteristics and other labour market outcomes.

Blocking and randomization was done for each sample separately, so that treatment is necessarily balanced across the two samples. I make sure to test for balance in the pooled sample, and in each sub-sample separately.

There is balance across a wide range of measures, in the pooled sample, and the two samples separately. Very few measures, and none of the variables used for the blocking or employment outcome variables are statistically different across groups. For the *board* sample individuals in the control group are more likely to be recent graduates (individuals who finished school, university or vo-

16 The nature of permanent work is discussed in more detail in Section A.1.

cational training in the last 15 months). This is a group that may not have been searching for work for quite as long.

1.3.4 *Attrition*

There is no difference in rates of attrition between individuals who were given the transport subsidies and those who were given the calls but not the subsidies. In addition, almost no baseline covariates are significant predictors of attrition at endline. Table 1.1 provides an overview of rates of attrition at various points of survey. Attrition is high in this sample.[17] I argue that much of the attrition is due to respondent who changed phone numbers, or provided poor contact information at baseline, before the treatment began.

We could not find 14% of the total sample at all after the baseline, and about 25% were not interviewed at the endline survey. However, among the phone call survey respondents, a large proportion (just less than half) of the total attrition took place between the first phone call surveys.[18] This sort of attrition is unlikely to be correlated with the transport treatment, since treatment did not start until *after* the first phone calls.

The phone calls do seem to have reduced attrition, but only by about 4% (the effect is not significant). The main results of this paper are robust to using either control group (the "no calls" and the "no transport" group), so this is not of great concern.

The group who received the subsidies had almost identical rates of attrition (at all stages of data collection) to those receiving calls but not treatment. Table A.6 shows the main determinants of whether a respondent was found at follow

17 This was because of the limited resources of the survey team (which required us to use phone calls as the main means of response during a time when the Ethiopian cellular network was very unreliable). This was also a very mobile study population: young people often left town, changed phone numbers, or did not want to respond, with short notice.

18 524 individuals were enrolled in the phone survey, 8% were never reached by phone. Out of the 465 that were reached by phone at least once, on average 400 were reached each week, an average rate of attrition on the phone survey of about 15%.

Table 1.1: Attrition by treatment status

		Calls		
	Control	No Transport	Transport	Total
Never found	81 24.85%	22 7.43%	22 8.63%	125 14.25%
Contacted by phone, not Endline	0 0%	35 11.82%	31 12.16%	66 7.53%
Refused at Endline	9 2.76%	12 4.05%	7 2.75%	28 3.19%
Interviewed at Endline	236 72.39%	227 76.69%	195 76.47%	658 75.03%
Total	326 100%	296 100%	255 100%	877 100%

up (in week 16). The first two columns show that the transport subsidies had no effect on attrition after controlling for receiving the calls. Columns (3)-(6) show that very little else impacted the probability of being found at follow up. I look at predictors of attrition for the *board* and *city* separately and find that attrition is particularly hard to predict for the *board* sample.

Furthermore, I show that the sample is balanced on baseline variables between treatment and control *after* attrition. Table A.5 Panel B shows balance after attrition to the endline survey, while Panel C shows balance among those ever reached for the phone call surveys. This shows that the actual samples used for estimating treatment effects are broadly balanced on covariances. All variables that are relatively strong predictors of attrition are used as covariates in estimating regressions as robustness checks. I conclude that the results in the paper are not driven by attrition.[19]

19 There is one notable exception to this. The *city* sample were particularly hard to track down at the second endline, 10 months later. The attrition in that sample is a problem. However, the results in this paper focus on the first endline. I discuss this issue further in Section 1.4.1.

I estimate the impact the transport subsidies on the labour market outcomes of job seekers at the main endline survey, 16 weeks (4 months) after the baseline, and about one month after the end of the transport subsidy program itself. I then present results from the second endline survey (conducted by phone) 40 weeks after the baseline to look at the persistence of these effects long after the experiment ended.

I leave aside the impacts on job search for now. Section 1.5 will use high-frequency phone call data, and argue that the increased job search intensity over the weeks of the study is the main driving force behind the large job outcome impacts documented in this Section. In Section 1.7 I test additional predictions of the theoretical model (outlined in Section 1.6) and investigate mechanisms other than increased job search that may be driving the results.

All results are intent-to-treat (ITT) estimates on binary labour market outcomes, using difference and difference-in-difference OLS estimators. Standard errors are clustered at the Woreda level (the lowest urban administrative unit in Ethiopia), of which there are 70 in my data, suggesting that I do not have problems with too few clusters (Cameron et al., 2008).[20]

For robustness, I employ a series of different estimators, and present the results from all of these specifications here, to show that results are not sensitive to specification. I focus on the limited set of binary outcomes used in the phone survey. In this sense I tie my hands to only look at a small group of outcomes that were chosen before the study began. Later I turn to explore other employment outcomes from the endline survey in more detail.

20 The Woreda system recent replaced the communist-era Kebele system in Addis Ababa. Woreda's were formed by the combination of 2, 3 or 4 former Kebele's into a large consolidated administrative units.

Regressions on endline outcomes take the form:

$$y_i = \alpha + T_i\lambda + \epsilon_i \tag{1.1}$$

$$y_i = \alpha + T_i\lambda + X_{i0}\beta + \epsilon_i \tag{1.2}$$

$$y_{is} = \alpha_s + \sum_s T_i S_{si}\lambda_s + X_{i0}\beta + \epsilon_i \tag{1.3}$$

T_i is the treatment variable dummy. Equation (1.1) estimates the basic difference in means between the treatment and control group (this specification is labelled BAS in the tables). Equation 1.2 includes a set of individual covariate controls (COV), based on baseline outcomes, which also include basic individual characteristics, and especially those that exhibit any minor imbalance between treatment and control at baseline. X_{i0} could easily be replaced with a set of blocking dummy's on which assignment to treatment was based (see 1.3), this specification is labelled in tables as (BLK).

Equation 1.3 provides the basic form for estimating different treatment effects for different groups (or heterogeneous treatment effects) by baseline group S_i. Usually this is used to estimate treatment effects for the two samples, the *board* and *city* samples, but will be employed to estimate treatment effects by different education outcomes, or poverty levels at baseline.[21]

Further I estimate difference-in-difference style estimators, by looking at the impact of treatment in the change in labour market outcomes between baseline and endline, as in equation 1.4 below labeled (FD) throughout. The ANCOVA estimator (labelled ANC), in equation 1.5, is similar but looks at endline out-

21 These coefficients measure the size of the treatment effect for each category separately. A simple t-test can be used to test the difference in the size of the coefficients

comes and includes a lagged dependent variable to account for differences the dependent variable at baseline in a flexible way.[22]

$$y_{i16} - y_{i0} = \alpha + T_i\lambda + X_{i0}\beta + \epsilon_i (t = 16) \tag{1.4}$$

$$y_{i16} = \alpha + y_{i0}\rho + T_i\lambda + X_{i0}\beta + \epsilon_i \tag{1.5}$$

1.4.1 *Jobs and Permanent Jobs*

I have discussed the labour market for young urban Ethiopians briefly. I provide more detail on the types of jobs available and who finds them in the Appendix Section A.1. Briefly, job seekers look for good jobs. Among highly educated individuals, this means looking for permanent jobs.[23] For less skilled job seekers, this means looking for any formal sector or white collar work. Other forms of casual or temporary work are readily available. These are used for subsistence and are considered inferior.[24]

Therefore, if transport subsidies improve access to good job opportunities for treated individuals we would expect the following for treated respondents: 1) better quality jobs and more permanent jobs for those can get them, 2) an increase in the probability of having any work which may be larger or smaller than the effect on the probability of having a permanent job. It is possible that the impacts on having any work at all are zero, if the treatment only induces displacement of temporary work for permanent jobs.

In Table 1.2 shows the impact of subsidies on finding permanent work at endline: treatment increases the probability of finding a permanent job, among the board sample, from 19% among the control group, by about 7 percentage points, indicating about a 30% increase in the probability of having a permanent

22 The ANCOVA estimator is more efficient that either difference in difference estimator or the standard POST estimator which ignores baseline outcomes (Frison and Pocock, 1992; McKenzie, 2012).

23 Permanent usually come with a written contract and an understanding that the job will be available to the employer indefinitely or for a set, but relatively long, period of time.

24 I discuss these mechanisms in more detail in the model of Section 1.6.

job. This is the central result of this paper. It shows that subsidies allowed active job seekers to find the jobs that they were looking for.

Table 1.2: Effects of transport subsidies on having permanent employment at endline

	Control Mean	(1) BAS	(2) LOG	(3) COV	(4) ANC	(5) BLK	(6) FD
Panel A: Average Treatment Effects At Follow Up (Pooled Sample)							
All	0.13	0.028	0.027	0.042	0.043	0.032	0.044*
		(0.027)	(0.024)	(0.026)	(0.026)	(0.026)	(0.026)
Observations	657	657	657	657	657	657	657
R^2		0.001		0.088	0.098	0.151	0.097
Panel B: Treatment Effects At Follow Up by Sample							
Board	0.19	0.068*	0.046*	0.078**	0.078**	0.073*	0.078**
		(0.038)	(0.028)	(0.037)	(0.037)	(0.040)	(0.037)
City	0.06	-0.019	-0.036	-0.004	-0.002	-0.020	0.001
		(0.032)	(0.054)	(0.034)	(0.032)	(0.026)	(0.033)
Observations	657	657	657	657	657	657	657
R^2		0.186		0.221	0.230	0.276	0.218

[1] The dependent variable is a dummy variable equal to one if the individual reported having a permanent job, measured at endline (week 16). Results are from OLS regressions on endline outcomes.
[2] Panel A gives average ITT effect for the two samples together. Panel B shows results two different samples- "board" and "city"
[3] Standard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

The *city* sample are not more likely to find permanent work, but this is because it is difficult for this group to find permanent work at all.[25]

I then look at the impact of the treatment on having any job at all. If job seekers had access to better employment opportunities we should see more individuals in work rather than unemployed. However, if the effects on "good jobs" were concentrated among individuals who would overwise have been working at temporary jobs, the effect on employment could be dampened.

I find that there is about a 6 percentage point increase in the probability of having employment at endline in the pooled sample, over a control mean of 53%. These results are concentrated among the *city* sample, for whom the effect

25 In fact the results are concentrated among individuals with universities degrees, to whom these permanent jobs are actually available. Results on heterogenous treatment effects in Section 1.7 discusses this in more detail. The effect is only present for individuals with degrees.

is large (at around 8 percentage points). The effect is smaller and not statistically significant for the *board* sample.

Table 1.3: Effects of transport subsidies on having employment at endline

	Control Mean	(1) BAS	(2) LOG	(3) COV	(4) ANC	(5) BLK	(6) FD
Panel A: Average Treatment Effects At Follow Up (Pooled Sample)							
All	0.53	0.058*	0.059*	0.062*	0.064*	0.057*	0.081*
		(0.034)	(0.035)	(0.035)	(0.034)	(0.034)	(0.043)
Observations	657	657	657	657	657	657	657
R^2		0.003		0.066	0.078	0.159	0.062
Panel B: Treatment Effects At Follow Up by Sample							
Board	0.58	0.044	0.046	0.043	0.046	0.049	0.067
		(0.051)	(0.052)	(0.052)	(0.051)	(0.051)	(0.062)
City	0.46	0.076	0.075*	0.086*	0.088**	0.068	0.099*
		(0.046)	(0.044)	(0.044)	(0.041)	(0.041)	(0.057)
Observations	657	657	657	657	657	657	657
R^2		0.553		0.066	0.079	0.159	0.062

[1] The dependent variable is a dummy variable equal to one if the individual reported having done work in the last 7 days, measured at endline (week 16). Results are from OLS regressions on endline outcomes.
[2] Panel A gives average ITT effect for the two samples together. Panel B shows results two different samples- "board" and "city".
[3] Standard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level.

Persistence and Dissipation

Are these impacts on job outcomes persistent a further 6 months after the first endline survey? Treated individuals were more likely to have jobs after 16 weeks, but the control group may have caught up over the proceeding 6 months by continuing to catch up. In this case the treatment effects will have dissipated after the control group have had long enough to search as much the treatment group did during the subsidy period. In this case the treatment would have had an impact on unemployment durations, an important finding, but nothing more than that. Alternatively, if job search intensity dropped off after a few months of search because job seekers became discouraged or ran out of the money required to search (for which I present considerable evidence in the next section) the control group might not catch up to the treatment group at all.

The treatment effects seem to be at least partly persistent. In Table 1.4 I show that when surveyed 6 months later (at week 40 of the project) those who were treated among the *board* sample are now 3.7 percentage points (roughly 10%) more likely to have permanent work. This is displayed along side the impacts for week 16, showing that the coefficient has roughly halved over time. The coefficient at 40 weeks is not statistically significant but is reasonably large.

Table 1.4: Impacts on having permanent work at both endlines (weeks 16 & 40)

Estimator	CM		Basic		Controls		First Diff	
			(1)	(2)	(3)	(4)	(5)	(6)
Week	16	40	16	40	16	40	16	40
Panel A: Average Treatment Effects At Follow Up (Pooled Sample)								
All	0.130	0.210	0.028	0.018	0.042	0.018	0.044*	0.017
			(0.027)	(0.038)	(0.026)	(0.033)	(0.026)	(0.034)
Obs			657	605	657	605	657	605
R^2			0.001	0.000	0.088	0.133	0.097	0.143
Panel B: Treatment Effects At Follow Up by Sample								
Board	0.190	0.310	0.068*	0.035	0.078**	0.033	0.078**	0.032
			(0.038)	(0.052)	(0.037)	(0.051)	(0.037)	(0.051)
City	0.065	0.080	-0.019	0.007	-0.004	-0.001	0.001	-0.002
			(0.032)	(0.037)	(0.034)	(0.038)	(0.033)	(0.042)
Obs			657	605	657	605	657	605
R^2			0.186	0.285	0.091	0.133	0.100	0.143

[1] The dependent variable is a dummy variable equal to one if the individual reported having a permanent job, measured at endline (week 16). Results are from OLS regressions on endline outcomes.
[2] Panel A gives average ITT effect for the two samples together. Panel B shows results two different samples- "board" and "city".
[3] Standard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level.

I look to see whether the effects on having any work are persistent in the *city* sample. Before doing so, it is important to note that there were problems with attrition for the *city* sample at 6 months, that were not present at the first endline or in the *board* sample in either survey. Attrition rates were high and resulted in a lack of co-variate balance at baseline for the sample that were found 6 months later.

In the basic difference specification, treated individuals in the the *city* sample were 17% more likely to be working at endline. However, this result should be read with caution: it coefficient is not stable when controlling for baseline covariates as illustrated in column (6). This is indicative of the fact that among the sample found at endline the treatment group were more likely to have had work at baseline, due to the attrition. Therefore the first difference (FD) estimates are the most trustworthy, suggesting that among the *city* sample the probability of working was increased by 10 percentage points (over a mean of 41%), although even this is not quite significant (these results are in Table A.8 in the Appendix).

The persistence of these effects suggests that reducing the duration of unemployment (or as it may be, transition into a first job) can have long lasting effects. Unemployment could lead to scarring if jobs seekers' skills worsen with time in unemployment, or if psychological factors make it hard for job seekers to resume looking for work after becoming discouraged or running out of savings.[26]

Job Quality

I now turn to look at the impacts on other measures of job quality and type at the *first* endline survey at 16 weeks into the study. Here I look at a number of measures that were not specified for inclusion into the phone survey. I test for whether the increase in employment for the *city* sample were in jobs of better quality, and whether the permanent jobs found by *boards* respondents are qualitatively different in other ways.

In Figure A.3 in the Appendix, I classify the jobs of all respondents working at follow up into occupational groups, rank those groups by average weekly salary earned at follow up, and plot the cumulative distribution among these

26 I discuss other issues related to the persistence of impacts on job search in Section 1.7 but otherwise these issues are beyond the scope of this paper.

occupations by treatment and control groups. The results clearly show a positive shift in the quality of jobs among the treated group.

Many job quality outcomes were significantly improved by the transport subsidies, as shown in Table 1.5. I look at a series of dummy variables indicating that a respondent has a job with a certain quality, all of which are in some ways proxies for the permanence, formality or desirability of work. The key variables are described in the notes to Table 1.5. For instance, treated respondents are 14 percentage points more likely to be working in office, as opposed to at another kind of other work site, and 4.3 percentage points more likely to have found the job through formal means (an application and a proper interview).

The results indicate that the *city* respondents are more likely to find jobs of better quality. They are not simply getting work faster by accepting inferior jobs. If I restrict the sample to just individuals who had work, and run the same regression, I confirm that, conditional on having a job, *city* respondents are more likely to have better jobs.

The *board* respondents, who are already likely to have jobs in office, or be paid by the month, do not see significant treatment effects on these variables. However, they are more likely to have found jobs that require at least a degree as a qualification. During focus group discussions run during the endline survey, many respondents attached special importance to the goal of finding work for which they were specifically trained.

I find no significant difference between the average wages of treated individuals and the control group. Although the coefficient is large for the *board* sample (9%) it is not estimated precisely. I am unable to reject the Kolmogorov-Smirnov test of equality in distribution of wages (in levels or logs) between treatment and control groups. This is not too surprising. The description of jobs presented in Appendix A.1 shows that permanent jobs are desirable for their security. They do not actually pay more than other jobs for entry level positions.

Table 1.5: Effects of treatment on job quality and type at endline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
			log			in	pay			
	work	casual	ln wage	hours	degree	office	monthly	satisfied	formally	in city
Panel A: Impacts on work outcomes at week 16										
TE Pooled	0.062*	-0.022	0.051	3.74**	0.047**	0.070*	0.069*	0.061**	0.054*	0.059*
	(0.035)	(0.024)	(0.088)	(1.71)	(0.018)	(0.037)	(0.037)	(0.028)	(0.029)	(0.032)
Observations	658	596	356	656	596	596	596	596	596	596
R-squared	0.067	0.077	0.115	0.079	0.228	0.059	0.107	0.058	0.114	0.051
Panel B: Heterogeneous impacts on work at week 16 by Sample										
TE board	0.043	0.0026	0.091	2.53	0.075**	0.020	0.032	0.015	0.064	0.097**
	(0.051)	(0.025)	(0.11)	(2.34)	(0.033)	(0.052)	(0.053)	(0.045)	(0.049)	(0.042)
TE city	0.087*	-0.050	-0.0090	5.27**	0.014	0.13**	0.11**	0.11***	0.042*	0.015
	(0.044)	(0.042)	(0.15)	(2.34)	(0.011)	(0.050)	(0.049)	(0.029)	(0.023)	(0.046)
Observations	658	596	356	656	596	596	596	596	596	596
R-squared	0.067	0.079	0.116	0.080	0.230	0.063	0.108	0.062	0.114	0.053

Results are from Difference OLS regressions on endline outcomes at week 16, using the simple (BAS) specification without covariates. Panel A shows results for the two samples pooled together.S tandard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level
Unusual Dependent variables: (5) *Degree*: Respondent has a job that required a degree as minimum qualification (6) *In Office:* Job is performed in an office, or formal business house- proxy for "white collar" work (7) *Pay Monthly:* Respondent is paid every month, usually according to set a contract (9) *Formally:* The job was acquired through an official application and interview process (this excludes referral from a friend or family, or jobs given after just a conversation with the employer

It could be the case that treatment brought individuals who otherwise would not have been employed into work: they might have average expected earnings, which would bring down the average wage in the treatment group and downward bias the impacts on wages. I check for this using a Heckman (1979) selection model. Using both the two-step inverse Mills ratio and log-likelihood approaches to control for selection, I find similar coefficients on the impact on log incomes which are still not significant, although they are slightly larger.

Discussion: Better Jobs or More Jobs?

Subsidized transport improved job market outcomes for job seekers. The *board* sample were more likely to be working at permanent jobs, some instead of being unemployed, others instead of having temporary jobs. Thus the increase in employment among the *board* sample is smaller than the increase in the probability in finding permanent work. This would be expected if permanent work was the main outcome on which these individuals search, and there was no impact on the probability of having temporary work.

It is interesting to note that at the second endline the incidence of permanent jobs has increased in both the treatment and control groups. The incidence of temporary jobs has stayed constant among those who have not found permanent work. Thus the incidence of temporary has actually declined to 28% from 37% at week 16, in the control group. This reinforces the idea that job seekers in the *board* sample are taking temporary jobs when they have to, while looking for permanent jobs.

By contrast, the *city* sample seems to have found a better set of temporary work opportunities, ones that are still temporary but look more like formal employment. In this sample treatment seems to have an induced individuals to find better jobs when they otherwise would have remained without work, which leads to a significant increase in the employment rate for this sample. Coefficients for the impacts on job quality are usually slightly larger than the impact on having any work at all, although not much larger. As with the *board* sample, this suggests that the effect on employment is partly driven by individuals finding high quality jobs who otherwise would have been working at worse jobs.

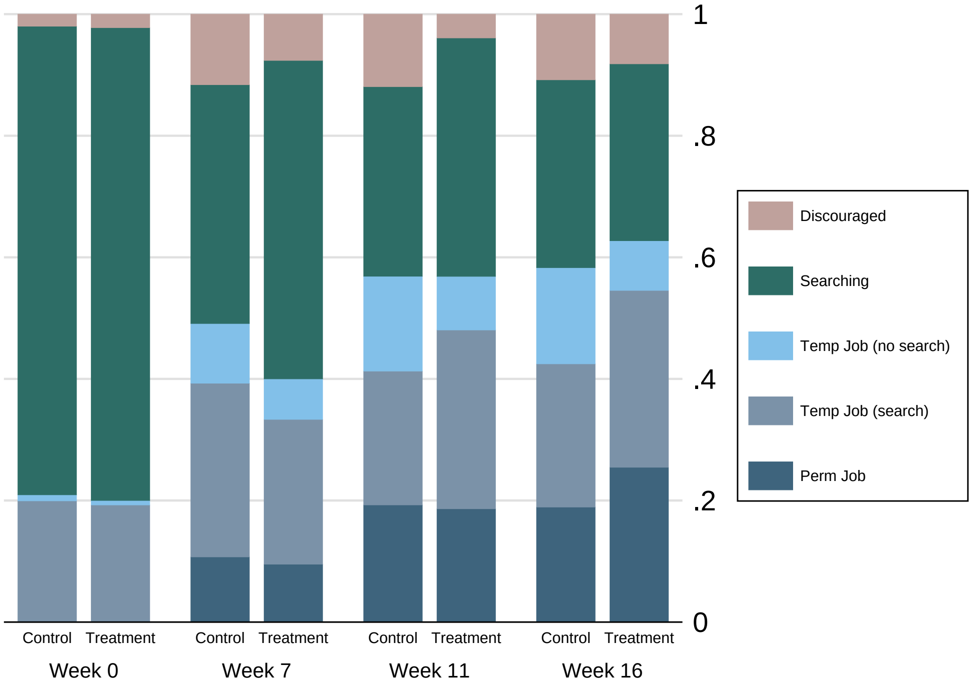
Firstly I turn to the next pressing question: what is the more proximate cause of the improved labour market outcomes documented in this section? I look at how job search responded to the transport subsidies.

1.5 MAIN RESULTS: JOB SEARCH

This Section looks at the impacts of the transport subsidies on job search at both the intensive and extensive margins, as well as on methods of job search. These impacts on search activity appear to be driving the impacts on employment outcomes. In Section 1.6 I discuss the link between the magnitude of the impact in job search and the magnitude of the impact on employment, to argue that job search could explain the main results on permanent employment.

I focus on the trajectory of the impacts of subsidies over time. This illuminates the role of cash constraints in the job search process. I find the impacts are not constant over time, but increase as job seekers run down their savings and become discouraged. Section 1.6 fully outlines the theoretical mechanisms by which cash constraints could produce this trajectory.

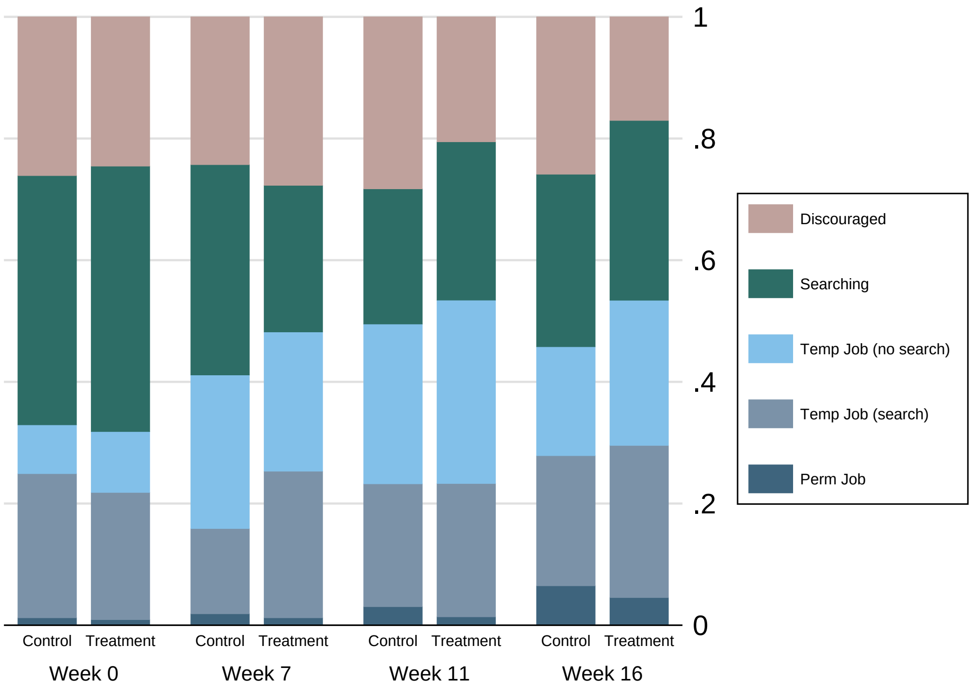
Figure 1.1: Composition of sample by week and treatment status: Board sample



1.5.1 *Overview: Composition Effects*

Before turning to the impacts on specific job search outcomes, it is useful to look at how the treatment effected the composition of labour activities in the sample over time. I do this separately for the *board* (Figure 1.1) and *city* (Figure 1.2) samples. For each week, I show the percentage of individuals engaged in 5 distinct states of employment, for treatment and control individuals. For the sake of clarity, I present only a select number of weeks to provide an overview changing compositions, the Appendix provides similar graphs for all 16 weeks.

Figure 1.2: Composition of sample by week and treatment status: City sample



The 5 states plotted here are (from the top of the bar chart): Discouragement (not working or searching); Searching (but no work) Temporary work (not searching) Temporary work (but also searching)[27], and Permanent work. In this ordering, the top of the blue bars in the graphs indicate the employment rate for the relevant group.

Week 0 shows the strong balance between treatment control in terms of the activities they were engaged in at baseline: most individuals were searching at this stage anyway, especially in the *board* sample. As the study goes on, more of the unemployed are likely to become discouraged (stop searching for work), while more of those with work are likely to give up looking for better jobs.

27 The distinction between Searching or not searching among temporary workers is important, as on the job search is extremely important, especially for individuals how do not consider their work to satisfactory or long term. If much temporary work is used as a means to short run subsistence, and perhaps to make money to search for other work, it is as interesting to look at job search in this group as those without work.

Table 1.6: Ordered logistic regression

	(1) All Weeks	(2) After Week 7	(3) Week 16 Only
Panel A: Effects across samples			
Effect for *boards*	0.20	0.42**	0.53***
	(0.14)	(0.18)	(0.19)
Effect for *city*	0.21	0.32*	0.30*
	(0.17)	(0.17)	(0.16)
Panel B: Effects in pooled Sample			
Pooled Effect	0.20*	0.37***	0.43***
	(0.11)	(0.12)	(0.13)
Obs (both panels)	5,011	2,202	658

Dep Var is a categorical variable: 1- Discouraged; 2- Temp work (no searching) 3- Searching (no work) 4 - Temp work (and searching); 5 - Permanent work. Log-odds coefficients are reported. All regressions include a full set of control variables. Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

These trends hold for both samples. The rate at which people give up job search is similar among those with, and those without temporary jobs.

The transport treatment, at each margin, pushes respondents away from discouragement, towards work, more permanent work, and increased job search intensity (regardless of employment status). This is true for most of the later periods of the study, after which the treatment had been running for a while and had time to take effect. A key focus of these results is how the subsidies had an effect sooner among the *city* sample than among the *board* sample.

Importantly the *board* sample seem to be *less* likely to be working in temporary jobs for the middle weeks of the study. As I discuss in Section 1.6, this is consistent with a theory of respondents substituting low quality temporary work in favour of more intensive search for jobs that they are actually interested in. If some job seekers are giving up looking for good permanent jobs by taking their outside options in temporary, or informal work, the treatment seems to have prevented this happening, at least temporarily. This effect goes away after the transport subsidies are removed.

I confirm that these trends are quantitatively significant. I employ an ordered logistic regression specified in equation 1.6 using the 5 job market outcomes are ranked ordinally in the order that they are presented.

$$P(j|X_{0i}, T_i) = P(\eta_j \leqslant y_i^\star \leqslant \eta_{j+1}) \tag{1.6}$$

where

$$y_i^\star = T_i\lambda + X_{0i}\beta + \epsilon_{it}$$

The results in Table 1.6 clearly show a statistically significant impact of the treatment on the ordered categorical variable, in the positive direction: away from discouragement to more job search and better jobs, for both samples. The effect is more pronounced for the *board* respondents.

1.5.2 *Job Search Trajectories*

How did the treatment impact the job search activity of recipients during the weeks that they were receiving it, and how did these impacts change over the course of the study? Did search intensity change change over time, and how and when did treated individuals diverge from the control group?

$$y_{it} = \alpha_t + T_i\lambda + X_{i0}\beta + \epsilon_{it} \qquad \forall t \neq 0 \tag{1.7}$$

$$y_{it} = \alpha_t + \sum_t T_i W_{it}\lambda_t + X_i\beta + \epsilon_{it} \tag{1.8}$$

I begin by presenting estimates of the treatment effect on the propensity to search for work over time, looking at the 12 post-baseline surveys: 11 phone call surveys (denoted by week 1-11), and the final face-to-face survey (week 16). I estimate the average impact on the probability of searching for a job across all 12 weeks combined (equation 1.7). Using equation 1.8, I then estimate the treatment effect in each week separately.[28] I estimate the trend over time, estimating an intercept term, linear, quadratic and cubic trend terms, as in equation 1.9 and 1.10, below.

$$y_{it} = \alpha_t + T_i\lambda_0 + T_i w\lambda_1 + X_i\beta + \epsilon_{it} \tag{1.9}$$

$$y_{it} = \alpha_t + T_i\lambda_0 + T_i w\lambda_1 + T_i w^2\lambda_2 + X_i\beta + \epsilon_{it} \tag{1.10}$$

Figure 1.3 summarizes all of these results for both the *board* sample (Column 1) and the *city* sample (Column 2). In Panel A, non-parametric estimates of the probability of searching for employment as a function of time are presented. They show that search behaviour declined over time, as individuals either found employment or became discouraged and stopped searching for work. However, for both samples, the treated group clearly shows a different trajectory.

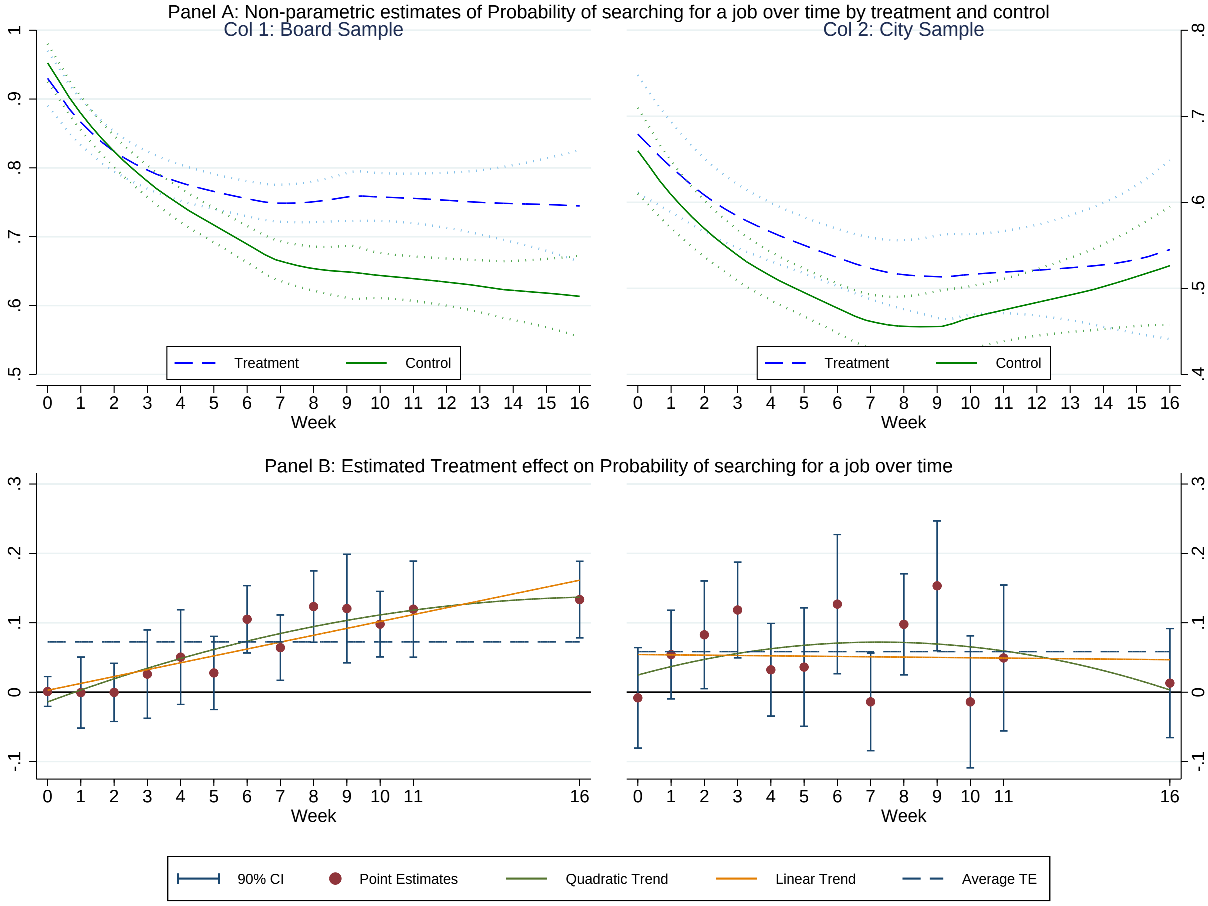
Table A.9 in the Appendix estimates (parametrically) these treatment effects week by week. I estimate one treatment effect coefficient for each week of phone data using equation 1.8. The control mean (CM) shows how the proportion of individuals searching for a job declined over time, but by considerably less for the treatment group, who were as much as 10% more likely to be searching in particular weeks of the study. I show results for the two samples pooled together, and for each separately.[29]

These weekly estimates of the impact of the treatment in each week are plotted in the Panel B of Figure 1.3, showing, for both samples separately, a clear upward trend in the treatment effect over time. For the *board* sample, these effects seem

28 In all specifications, "treatment" is defined as having received the transport subsidies as any point in the past, the treatment switches on in week 1, and does not "switch off". In later analysis, I exploit variation in when the subsidy treatment was ended for different individuals, and the fact that the treatments ended by at least week 11 for everyone (5 weeks before the follow up paper survey) to estimate the persistence of the treatment effects.

29 Power is low for weekly-sample specific treatment effect estimates, so the pooled estimates more often statistically significant, but hide some heterogeneity between the two groups.

Figure 1.3: Impact on job search: Non-parametric trends & treatment effects over time



to increasing linearly with time, whereas for the *city* sample (in Column 2), these effects seem to have a more immediate effect at the beginning of the study, remaining constant with a decline towards the end (the effect is negligible in week 16).[30] Panel B also overlays the linear and quadratic estimates of the trend in the treatment effect over time.

The Appendix contains Tables that show the parameter values used in these plots. They show a significant linear trend for the *board* sample and a mostly constant effect *city* sample. The quadratic term for the *city* sample is negative, reflecting the decline in the effect in the final period, but is not statistically significant.

30 I show, shortly, that this decline in search activity may be driven by these individuals finding better work.

The results suggest unambiguously that individuals in both samples were more likely to search for jobs while receiving the transport subsidies, but the trajectory of these impacts differs slightly between the samples.[31] For the boards sample, who were initially more likely to be searching for employment, the impacts took some time to kick in, doing so only as individuals become discouraged. For the *city* respondents, the effect seems to have been more immediate, but less persistent through to the later weeks.

Search Methods

The nature of the transport subsidy, which required job seekers arrive at the centre of the city, also had an impact on the method of job search employed by job seekers. I find that the treatment had an impact on the probability of respondents searching for employment at the vacancy boards (see Figure A.6 and Table A.11 in the appendix). Again I find some heterogeneity by sample: the effects not as strong for the *city* sample, perhaps because the boards were never likely to be their preferred method of search.[32]

Search at the intensive or extensive margin

The evidence suggests that treatment did not induce job seekers to search more intensively during a given week.[33] The results presented in the Appendix Table A.14 show a positive impact of treatment on the average number of days spent searching for work (and days visiting the boards), but these effects are driven entirely by the increase in proportion of individuals searching for work. Indeed,

31 The impact on job search at week 16 shows that treated respondents were more likely to search *after* the subsidies ended. I return to this point in Section 1.7.

32 Although there are individual weeks, early in the study, where the effect is significant, suggesting that the intervention "nudged" or at least encouraged respondents to try to check the boards, possibly with little tangible reward

33 However, it is hard to be certain of this because the treatment induced an increase in search at the extensive margin. This may have induced job seekers who were likely to be relatively inactive job seekers to search, bringing down the average search intensity in the treatment group.

estimates not presented here, show that there was no significant impact in the number of days searched, conditional on an individual search at all.[34]

This suggests that the treatment had the effect of increasing search only at the extensive margin. This conforms with responses in qualitative interviews, that job search was only worthwhile if done for a few days a week. Checking the boards regularly, following up with applications, and responding to interviews required two or three days a week. Searching more was not adequate, doing more was subject to decreasing returns. Instead the treatment increased the probability that individuals decided to make that effort to undertake search for the week at all. This result is reflected in the assumptions of model developed in Section 1.6.

Effects on temporary employment

I find evidence that for the *board* sample, in intermediate weeks of the treatment period, there is a statistically significant drop in the proportion of individuals taking work (Figure A.7 and Table 1.7). There was no impact on the rates of finding permanent employment in these early weeks: the effect is driven completely by jobs in temporary casual or other informal sector work. These are jobs that do not last long, and the effect is only present for a few weeks, but provides strong evidence that active job seekers choose to reduce their supply to labour to temporary work because of the subsidies.[35]

These results suggest that temporary work acts as a source of cash while for job seekers looking for permanent work. They may do so to pay the costs of search directly, to smooth income while unemployed, or to generate buffer savings. Transport subsidies lower the financial burden of searching, and so reduce

34 Although, again, this could be because the individuals who were motivated to begin searching for work where ones that were not naturally inclined to search, and thus searched less when they were searching.

35 For the *city* sample, there also seems to be evidence of an initial fall in employment rates, with a strong upward linear trend. However the negative impact on work for only one period may be an outlier and should not be interpreted with too much confidence.

the need for job seekers to take temporary work. These issues are discussed in more detail in Section 1.6.4.

Impacts on search activities by Months

The small samples used to estimate the impact on job search in each week individually have limited power to detect significant treatment effects, despite the large coefficients estimated. In order to increase the precision of my estimates, I follow McKenzie (2012) and pool weekly observations together and estimate average effects for all weeks at once. I pool observations into sets of four consecutive weeks, or 3 successive *months*, which allows me to confirm the trajectories of the treatment effects, with considerably more power. Monthly results are presented in Table 1.7 using Specification 1.11:

$$y_{imt} = \alpha_t + \sum_m T_i M_{mi} \lambda_m + X_{i0}\beta + \epsilon_{it}$$

$$y_{imt} = \alpha_{st} + \sum_s \sum_m T_i M_{mi} S_{is} \lambda_{sm} + X_{i0}\beta + \epsilon_{it} \quad (1.11)$$

The results emphasise the trajectories illustrated in the figures above. The treatment effects on search activity take some time to take effect, and grow over time for the *board* sample. The impacts are seen as early as the first month for the *city* sample, but seem to have diminished by the third month. For the final month, both samples are significantly less likely to be discouraged. For the *board* this is driven largely by increased search activity among the unemployed, for the *city* sample is a combination of increased search, and increased employment rates.

To allay concerns that these months were chosen strategically to boost significance, I present complementary results where I *restrict* the sample to groups of four weeks, starting with the first four weeks of the study, and then iteratively move this window forward by one month. I re-estimate the treatment effects separately for each iteration. This provides a type of moving average monthly

Table 1.7: Monthly impacts of treatment on main job market outcomes

	(1) work	(2) work perm	(3) searchnow	(4) searchboards	(5) discouraged	(6) days search
Panel A: Average Impacts By Month						
month 1	-0.018	-0.017	0.040*	0.033	-0.012	0.19*
	(0.024)	(0.012)	(0.023)	(0.023)	(0.018)	(0.097)
month 2	-0.024	-0.009	0.070***	0.084***	-0.034*	0.110
	(0.024)	(0.012)	(0.024)	(0.023)	(0.018)	(0.098)
month 3	0.038	0.012	0.083***	0.081***	-0.064***	0.28***
	(0.023)	(0.012)	(0.023)	(0.023)	(0.018)	(0.096)
R-squared	0.487	0.141	0.648	0.545	0.231	0.471
Panel B: Average Impacts By Month and Sample						
board month 1	0.018	-0.009	0.011	0.033	0.007	0.170
	(0.033)	(0.016)	(0.032)	(0.031)	(0.025)	(0.13)
board month 2	-0.067**	-0.013	0.074**	0.090***	-0.024	0.130
	(0.033)	(0.016)	(0.032)	(0.031)	(0.025)	(0.13)
board month 3	0.021	0.028*	0.11***	0.11***	-0.054**	0.47***
	(0.031)	(0.016)	(0.031)	(0.029)	(0.024)	(0.13)
city month 1	-0.048	-0.021	0.069**	0.009	-0.039	0.200
	(0.036)	(0.018)	(0.035)	(0.033)	(0.027)	(0.14)
city month 2	0.028	-0.004	0.058*	0.070**	-0.043	0.055
	(0.036)	(0.018)	(0.035)	(0.034)	(0.027)	(0.14)
city month 3	0.058*	-0.014	0.043	0.033	-0.072***	0.036
	(0.035)	(0.017)	(0.034)	(0.033)	(0.027)	(0.14)
R-squared	0.492	0.159	0.651	0.572	0.235	0.478
Observations	5,011	5,010	5,011	5,011	5,011	4,949

[1] Dependent Variables are listed at the top of each column. Results are from POST-OLS regressions on endline outcomes,
[2] Analysis excludes the follow up survey, just restricting analysis to the sample contacted in the phone surveys, with Month 1 defined as weeks 1-4, Month 2 as weeks 5-8 and Month 3 as weeks 9-12.
[3] Panel A gives average ITT effects across the full sample. Panel B estimates different coefficients for the two subsamples.
[4] Standard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

treatment effect, and clearly shows the trajectory of treatment effects. These estimates are shown for the pooled sample in Appendix Table A.16. These results confirm that the treatment starts to work slowly, and is strong and significant in the later weeks.

1.6 THEORY

I develop a model which considers the consumption and job search decisions of an unemployed cash- and credit-constrained job seeker. I use this model to explain the job search behaviour and the trajectory of treatment effects observed in the data. The model demonstrates why job seekers don't search for work when the returns to search are very high and they have enough cash on hand to pay the costs of search (at least once). I based the key assumptions of the model on observations from the data as well as qualitative insights from focus group discussions conducted during the baseline surveys.

The models describes the experience of an individual who is cash constrained, and who receives income from taking spells of temporary work in order to smooth consumption and earn money to pay search costs. In each week, he or she pursues a permanent job by choosing to pay a fixed monetary search cost.[36] The model incorporates search-on-the-job. Agents can search for work and have temporary employment at the same time.

The key intuition of the model is that a risk-averse job seeker with low levels of cash on hand finds it too risky to look for a job because the returns to search are low in expectation. The value of a permanent job is high, but probability of finding one is small: only 19% of my control group had permanent jobs after 15 weeks.

After describing the main features of the model, I proceed as follows: Firstly, I provide stylized solutions to the stationary version of the model. The key prediction of the baseline model is that job seekers find it optimal to not search below a critical value of savings, and that this critical value can be many orders of magnitude larger than the cost of search itself.

36 For the sake of simplicity, the model is framed in terms of the decisions of a job seeker from the *board* sample, in search of permanent work. The model could easily be applied to the *city* sample, in search of a higher quality job. In the model the two samples differ according to their initial savings levels and search behaviour. I return to this distinction later in this section.

Secondly, I simulate time series predictions of savings, consumption and search paths using the solved model. This allows me to predict stylized dynamics of job search behaviour over time. I show that these predicted dynamics are consistent with those found in the data.

As well as providing an intuition for the descriptive trends observed in the data, the model is used to predict the effect of reducing the costs of search in the way that the experimental treatment did. I show that these predictions are consistent with the estimated treatment effects, in both the static and dynamic cases.

Thirdly, having described the key predictions that the model is able to produce, I summarize these predictions for a wide range of combinations of parameter choices. This shows that for a great number of plausible parameter values, the predictions of the model are consistent with the empirical results found in the paper. Finally, I generate new predictions using the model, which I test in detail in Section 1.7.

1.6.1 *The Model*

The model is presented in discrete time. As is the case in the surveys and empirical analysis, the time period under consideration is one week. The job seeker begins each week with personal savings x, which could be cash on hand or formal savings in the bank. S/he starts the week by deciding whether she will search for a permanent job and pay the corresponding cost of search p. The decision to search is binary, so that the cost of search is constant and corresponds to the costs of transport, photocopying, making applications and buying newspapers.[37]

Searching for work at the start of the period leads to a permanent job with probability σ. If a permanent job is found, unemployment ends, and the newly

37 The assumption of a binary search cost is motivated by the fact that, conditional on searching, days of search per week are relatively constant. Most job seekers make two trips to the centre to look for work, corresponding to key days when new information about vacancies is published. The majority of variation in search activity is between rather than within weeks.

employed person receives income Y with certainty in the current period and in all future periods. Permanent employment is an absorbing state: these jobs cannot be lost. The permanently employed worker solves a kind of cake eating problem, augmented with a permanent income stream:

$$V(x) = \max_{0 \leqslant c \leqslant x+Y} u(c) + \beta V(x - c + Y) \tag{1.12}$$

Someone who hasn't paid the cost of search cannot find a permanent job in that period. Individuals remain unemployed for the rest of the period if they have searched but failed to find work, or have not searched at all. Unemployed individuals can earn: with probability θ they earn a known income W, which is paid at the end of the period.[38] W could be interpreted as cash transfers from other family members but is more realistically thought of as income from temporary employment, earned for work done as a casual/temp labourer, for a family member, or in self-employment in the informal sector. Search is not required to get these jobs.[39]

Let $U(x)$ be the value of being unemployed at the start of the period with savings x. An individual who chooses not to search for a job solves:

$$F(x) = \max_{0 \leqslant c \leqslant x} u(c) + \beta(\theta U(x - c + W) + (1 - \theta)U(x - c)) \tag{1.13}$$

So $F(x)$ is the value of the decision not to search. A job seeker who has failed to find work faces the same problem, but has already spent p on search. Therefore $F(x - p)$ gives the value of having searched but failed to find work.

38 This income is only realised in the next period, earned income increases the value of unemployment in the future, but cannot alleviate credit constraints in the current period if savings are already close to the zero lower-bound.

39 I take the arrival of these forms of temporary work to be random occurrences: the job seeker has no choice about whether to take the job (there is no leisure cost to taking temporary work) and the probability of getting this work is independent from the decision to search.

The optimal consumption decision of course differs between searchers and non-searchers.[40] Consumption when searching (c^s) is inevitably lower than consumption after not choosing to search c^{ns} because the searching individual has lower savings after paying the cost of search, and does not risk running down savings any further.

Using expressions for the value of permanent work and the value of failing to find permanent work, I write the value of searching for work at the beginning of the period:

$$S(x) = \sigma V(x-p) + (1-\sigma)F(x-p) \tag{1.14}$$

Here, σ is the probability of finding permanent employment. I can now write an expression for the value of unemployment, which is given by the envelope of the value of searching and not searching, reflecting the job-seeker's decision to search at the beginning of each period.

$$U(x) = max\{S(x), F(x)\} \tag{1.15}$$

I am interested in when individuals choose to search for work. That is when $S(x) \geqslant F(x)$. The key insight of the model comes from Expression 1.14. $F(x)$ is monotonically increasing and concave, and $S(x)$ is a convex combination of $F(x-p)$ and $V(x)$. Therefore, as is the standard result for models of this kind (Bryan et al., 2014; Vereshchagina and Hopenhayn, 2009; Buera, 2009) $S(x)$ crosses $F(x)$ once from below.

I define x^* as the critical value for which $S(x^*) = F(x^*)$. This is the level of savings below which an individual gives up looking for work. At these levels search becomes prohibitively risky, job seekers trade off the benefits of search

40 In each case the model is solved using an optimal control decision which solves the stochastic euler equation equating the marginal utility of consumption in the current period with expected margin utility of consumption in the future period.

against the costs of reducing their buffer savings (Deaton, 1991). The result is further illuminated by rewriting the expression for this critical value as:

$$\sigma(V(x^* - p) - F(x^*)) = (1 - \sigma)(F(x^*) - F(x^* - p)) \tag{1.16}$$

The right hand side of this expression represents the expected benefit of searching and the left hand side the expected loss, relative to not searching. At low levels of savings $F(x) - F(x - p)$ becomes large and searching becomes very risky because marginal utility of future consumption is very high. The left Panel of Figure 1.4 illustrates this intuition. $S(x)$ falls quickly as savings get lower.

The model implies a steady state level of savings at x^* for most calibrations of the model. Below this level individuals save up for search. Above it they run down their savings by paying the costs of search.

The model predicts that lowering the cost of search reduces the critical value x^*. Lowering search costs reduces the risk of searching. Define x^{*t} as the new critical value when the costs of search have been reduced permanently. Lowering search costs also has implications for the dynamics of savings over time: job seekers run down their savings less slowly when costs are lower. These dynamic implications are discussed in detail in Section 1.6.3.

This model cannot be solved analytically, because job seekers' optimal policy depends on wealth levels, which are in turn a function of their optimal decisions. I now turn to describe the numerical solutions of the model.

Note that by design the model can generate the 4 states of activities observed among individuals without permanent work in the data (see Figure 1.1): searching without working, searching and working, discouragement (not searching and not working), or working without searching.

1.6.2 *Solution to the Stationary Model*

In this section, I solve the model numerically.[41] This happens in two steps. First I estimate the value of of permanent employment, which is given simply by Expression 1.12. Secondly I use that value of permanent work to estimate the value of searching, not searching, and unemployment, and solve for these and the corresponding consumption levels in each state.

The left panel of Figure 1.4 shows the value for searching and not searching for different levels of savings, and the critical value above which it pays to search. I verify that the model implies that $S(x)$ crosses $F(x)$ once from below, and I solve for the critical value x^* at which an unemployed individual is indifferent between searching and not searching. In this case it is $x = 200$. Figure A.10 in the Appendix shows the optimal consumption paths for an individual who is forced to search, and an individual who cannot search in the current period.[42]

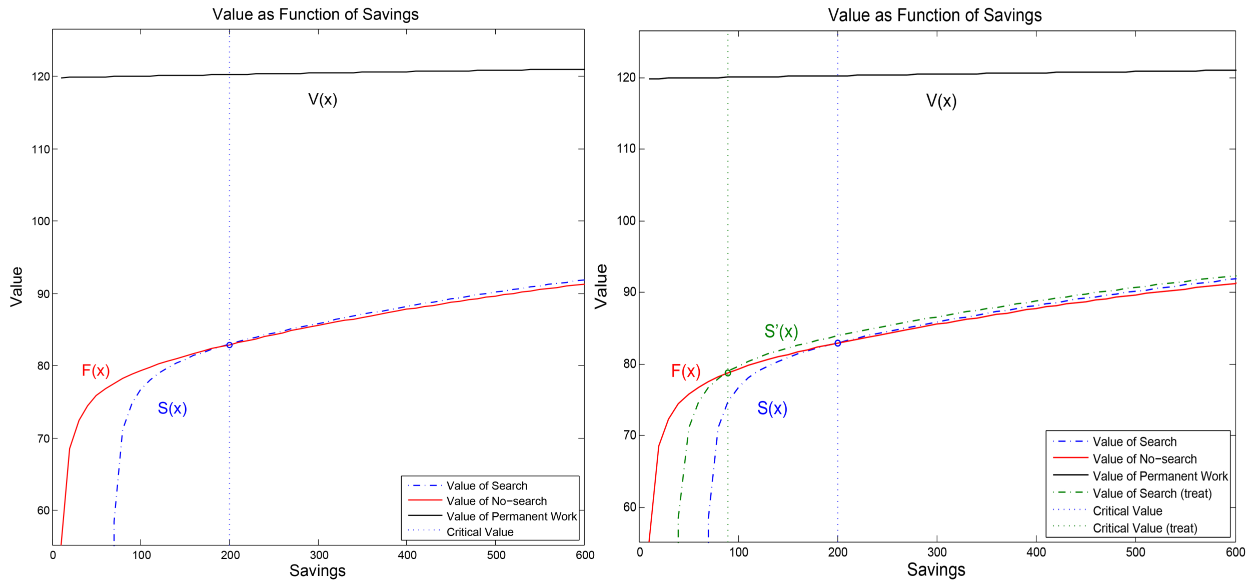
I estimate the impact of *permanently* reducing the cost of search (usually by a half or a third) on x^*. I call this x^{*t}. I then also estimate x^{*s}, the critical value for the case when search costs are reduced for just one period, before it revests back to the original search costs, so that the continuation value of unemployment remains unchanged.[43]

I then estimate the proportion of individuals that would be induced to search by the change in the critical value, given the distribution of savings in the data. That is, the proportion of individuals with savings $x^{*s} < x < x^*$ at baseline. This gives a rough idea of the static impact of reduced searched costs due to

41 The model is solved by iterating over the value function, using a grid for the state space (savings) and re-estimate the value function until it converges. This is done in Matlab.

42 In this model consumption does not follow the 45 degree line at low levels of savings: because income is always uncertain, cash constrained individuals need to always be keep savings in hand in case no income is received in the following period.

43 Unsurprisingly, x^{*s} is always (weakly) less than x^{*t}: people are more likely to search when the costs are search are set to rise again in the next period. However, the difference between the two solutions is usually relatively small. The treatment induced by this experiment reduced the cost of search temporarily, but for 12 weeks rather than one. Thus we might expect the effect in each week of this experiment to be somewhere between x^{*t} and x^{*s}.



Left panel: baseline model. Right panel: Effect of lowered search costs (treatment)

Figure 1.4: Single Crossing Point of the Value of Searching and not Searching

the experimental treatment. See the right hand Panel of Figure 1.4: reducing the search costs for one period shifts the value of search to left, such that (in this example) everyone with savings $90 < x < 200$ are induced by treatment to start searching in that period.

I provide an illustrative example of how the model's predictions can be tested against the empirical findings presented in the paper before looking at the calibration results across a wide range of parameter choices.

Illustrative Example

Can the model's predictions of the impact on x^* be reconciled with the impact of the randomized subsidy on the probability of someone searching for work? For this example and all calibrations, I use the estimates from the data to parameterize wages, the cost of search and the average probability of finding permanent work. Table 1.8 show the chosen values for these parameters. Other parameters of the model are allowed to vary across calibrations. For this example, I set probability of finding informal work while unemployed to $\theta = 0.3$, and use utility given by a simple log-utility function, $u(x) = log(x)$, and a discount rate $\beta = 0.9$.

For these parameters a job seeker gives up searching when his/her savings falls below the critical value $x^* = 600$. This is 10 times the cost of search, and 50% more than a week's salary in a permanent job.

I simulate the effect of reducing the cost of search from 60 to 40 birr per week (a one third reduction). This changes the critical value to $x^{t*} = 360$. Reducing the cost of search for one period has the same effect $x^{s*} = 360$ in this simulation, although this is not the case in all calibrations.

In the baseline data savings are 874 at the average, with a standard deviation of about 1300, and a median of 400. Here 12.2% of the sample have savings between 360 and 600, and would be induced to search for one period according to the predictions of the model. This is not out of line with my experimental estimated treatment effect of 8% on the probability of searching for work.

1.6.3 *Dynamic Simulations*

The model can be used to predict, for a given starting level of savings, search, saving and consumption behaviour over time. I use these predictions to confirm the dynamics observed in the data. Most notably, the data shows that job seekers give up search over time after they start out searching, and that many job seekers oscillate between searching and not searching every few weeks. The proportion of individuals searching in the data stabilizes over time. In section A.2 in the appendix, I discuss these patterns in greater detail.[44]

To show that the model replicates these patterns, I simulate job search behaviour and a series of random shocks for an individual starting with savings well above x^* and estimate the expected number of weeks it takes for that individual to run savings to below x^*. This is the number of weeks before an agent

44 I do this using data from a new and enlarged phone call survey currently underway in Ethiopia.

will switch from searching to not searching. I denote this as w^*. I then reestimate this value with search subsidy in place, and call this w^{*t}.

Secondly, I show that the model implies a steady state probability of searching for work in each period. I simulate an initial savings level for 1000 individuals, in such a way that it resembles the savings distribution in my sample at baseline. Further I simulate a series of random temporary employment shocks for each individual in each time period. Using the optimal consumption and search decisions from the solution to the calibrated model, I look at the evolution of job search and savings over time. Job seekers' savings converge to the critical x^*, above this level they search and run down their savings, below this they save up for search. This generates oscillating search behaviour from week to week. I find that the proportion of individuals searching stabilises around steady state given by s^*. s^{*t} gives the corresponding steady state proportion searching for work when the search costs are reduced.

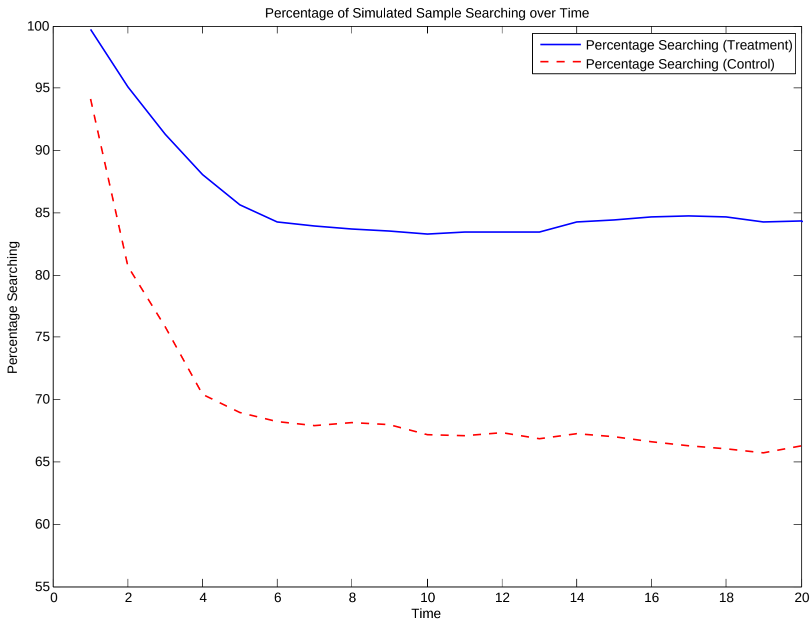
Treatment reduces the speed with which individuals run down savings, and increases the steady state proportion of individuals searching for work. Figure 1.5 shows a calibrated example of the proportion of the sample that are searching for work over time. These impacts on the dynamics of job search are often larger than those predicted by the static analysis.

Illustrative Example

For this example I use a representative agent with a log-utility function, $\theta = 0.3$, $\sigma = 0.03$, $\beta = 0.95$, and $p = 60$ (see Table 1.8 for a summary of the parameters). The critical value at which someone gives up search is $x^* = 600$ for this calibration. I then simulate initial savings for 1000 job seekers. Savings follow a logNormal distribution with mean 800 and variance 1000, which is closely resembles the distribution of savings at baseline in the data.

I generate time-series for each individual by simulate a series of income shocks and following the optimal consumption and search paths ascribed by the model. Figure 1.5 shows proportion of individuals in such a simulation searching for work over time, along with the corresponding trajectory for a treatment group for whom the costs of search are reduced by a third but initial savings are identical. Time series simulations of the model are consistent with the patterns observed in the data. The trajectories predicted in Figure 1.5 are similar to those estimated empirically in Figure 1.3 (Col 1). The calibration starts from Week 1, to correspond to the week in which subsidies started in the experiment.

Figure 1.5: Percentage of the sample searching for a job over time (simulation)



Almost all individuals start with savings above x^*, so that nearly 95% are searching for work in week 1. As a result the predicted *static* effect of reduced search costs in the first period is small. This mirrors the result from the experiment. However, job seekers who start out searching give up search over time. Someone starting with savings of 800 searches in every period to begin with. On average it takes such an individual 11 weeks to down savings below $x^* = 600$.

When search costs are reduced by a third, this time taken to run down savings below x^* is extended to 19 weeks on average.

The model implies a steady state proportion of individuals searching for the treatment group that is significantly higher than the control group. In this calibration someone the control group searches about 68% of the time in steady-state. The treatment group is roughly 20% more likely to search in each period. This result is much larger than the effect implied by the static analysis.

Note that these effects on the steady state proportion of searchers are driven by two factors. Firstly there is the impact on the critical value of x^* which I have described in detail. Secondly the lower search costs lead to savings being run down slower. Even when an individual has savings close to the critical value this second factor is at play: an individual who waits for a positive income shock to increase savings above x^* can search for longer before savings dip below x^* again.

1.6.4 *Calibrations*

Here I repeat the analysis in the examples above for a range of parameter values. For each paramterization of the model I estimate the x^*, s^* and w^*.[45] I look at the effect of reducing search costs from p to p_t (60 to 40 Birr) on each of these outcomes. I show that the predictions of the model are broadly consistent with the data for a wide range of parameter values. The calibrations also illustrate interesting additional predictions of the model.

Table 1.8 shows the calibration choices. In the top panel I show the values that are constant across calibrations, in the lower the panel the range over which other parameters vary in the calibrations.

45 These are the the critical value for savings, the steady-state proportion of job seekers, and the number of weeks that it takes for a job seeker to run down savings to x^* when starting with savings higher than x^*, respectively.

Table 1.8: Key parameters values for model calibrations

parameter	description	value
Y	Weekly wage for permanent wages	400 Birr
W	Weekly wage in temporary work	320 Birr
p	Weekly cost of searching actively work	60 Birr
p_t	Subsidies cost of searching for work	40 Birr
σ	Probability of finding permanent work if searching	0.03
θ	Probability of offer of temporary work if unemployment	(0.1-0.5)
β	Probability of offer of temporary work if unemployed	(0.8-0.99)
$u(c)$	Utility function: log/power utility	
δ	CRRA for power utility function	(1.2-2.8)
Solves for:		
x^*	Critical value of savings below which agent does not search	
w^*	Number of weeks it takes someone starting with x=1600 to run down savings to x^*	
s^*	Steady state value probability of searching in each week	

Table 1.10 presents the solution with different parameter values for β and θ for an agent with log-utility. The critical value decreases with risk aversion: unemployment individuals are more likely to search when the value future consumption more and thus are willing to sacrifice consumption in the short run for the chance to get a permanent job. The impact on x^* is slightly bigger when search costs are reduced for only one period, rather than permanently reduced ($x^{*t} \leqslant x^{*s}$), but this difference is usually small. As was the case in the illustrative examples, I calculate the % induced to search (Col 4) in the static case by looking at the proportion of my sample who had savings between x^{*t} and x^*.

In Table 1.10 I look at the effect of risk aversion on the results. I use power-utility, with different values for the constant rate of risk aversion δ. As expected, the critical value x^* is monotonically increasing in the risk aversion, job seekers are less likely to take the risk of searching. Similarly the steady state proportion of searchers falls with risk aversion.

Table 1.9: Patience & job search: Calibration of the main outcomes & simulated treatment effects
(Agent with log-utility. $p = 60$, $p_t = 40$, $Y = 400$, $W = 320$, $\sigma = 0.03$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Statics				Steady State			Discouragement		
	x^*	x^{*t}	x^{*s}	induced	s^*	s^{*t}	s^{*t} -s^*	$\overline{w}$	$\overline{w}^t$	$\overline{w}^t - \overline{w}$
β					$\theta = 0.1$					
0.8	1140	620	620	12.2%	0.0%	1.5%	1.5%	4.27	6.66	2.39
0.9	700	420	400	10.0%	2.2%	14.8%	12.6%	8.03	12.97	4.94
0.95	480	320	300	7.6%	15.0%	33.7%	18.6%	15.05	21.66	6.61
0.99	280	220	180	9.6%	40.3%	61.8%	21.5%	30.09	43.94	13.85
β					$\theta = 0.3$					
0.8	1200	600	600	12.5%	0.0%	11.3%	11.3%	3.22	6.55	3.33
0.9	600	360	360	12.2%	19.4%	52.8%	33.5%	8.85	13.91	5.06
0.95	400	260	240	5.2%	53.1%	76.5%	23.3%	17.68	25.8	8.12
0.99	240	180	180	8.6%	84.7%	93.8%	9.1%	40.45	43.4	2.95
β					$\theta = 0.4$					
0.8	1280	600	600	13.3%	0.0%	14.9%	14.9%	3.41	6.59	3.18
0.9	620	360	340	14.1%	25.8%	61.2%	35.4%	8.64	13.13	4.49
0.95	380	260	240	5.0%	66.1%	83.3%	17.2%	18.84	30.21	11.37
0.99	240	180	160	8.8%	91.7%	96.3%	4.6%	41.26	45.1	3.84

High rates of risk aversion are not needed to rationalize my results. In fact, calibrations with relatively low risk aversion parameters are more consistent with the job search behaviour and treatment effects estimated in the data. The parameters that most closely matches the empirical results are $\theta = 0.4$ and $\delta = 1.2$. The calibrations of the model are consistent with the estimated treatment effects in the data.

In these calibrations changes in the critical value x^* is highly sensitive the cost of search. For almost all parameter values the results are qualitatively similar to the results in the paper. Even for very low rates of risk aversion and impatient consumers, a significant proportion of individuals can be induced to search more, and for longer, by lower search costs.

In fact, the predicted treatment effects in the model are considerably larger than the treatment effect found in the paper, for a large number of calibrations. Of course, the model seeks to explain the behaviour of a marginal consumer for

Table 1.10: Risk aversion & job search: Calibration of the main outcomes & simulated treatment effects

(Agent with power utility function. $\beta = 0.95$, $p = 60$, $p_t = 40$, $Y = 400$, $W = 320$, $\sigma = 0.03$)

	x^*	x^{*t}	x^{*s}	induced	s^*	s^{*t}	s^{*t} -s^*	$\overline{w}$	$\overline{w}^t$	$\overline{w}^t - \overline{w}$
δ					$\theta = 0.1$					
1.2	640	440	400	9.0%	10.2%	23.8%	13.6%	12.9	20.89	7.99
1.4	840	560	540	8.0%	5.2%	19.0%	13.9%	10.94	18.29	7.35
1.8	1280	860	820	8.2%	1.5%	10.6%	9.1%	5.68	12.75	7.07
2	1500	1020	980	8.0%	0.5%	7.7%	7.3%	3.51	11.33	7.82
2.4	1820	1340	1320	2.2%	0.3%	2.8%	2.5%	1	5.85	4.85
2.8	1940	1660	1620	0.8%	0.3%	1.1%	0.8%	1	1	0
δ					$\theta = 0.3$					
1.2	520	340	320	11.6%	44.4%	71.0%	26.6%	15.75	25	9.25
1.4	640	440	400	9.0%	38.2%	64.3%	26.2%	13.82	23.74	9.92
1.8	920	640	580	8.0%	23.9%	54.2%	30.3%	10.61	20.26	9.65
2	1060	740	680	10.4%	19.3%	48.5%	29.2%	8.26	19.3	11.04
2.4	1360	960	900	7.6%	10.9%	37.0%	26.2%	5.99	15.06	9.07
2.8	1640	1180	1120	3.4%	6.7%	29.1%	22.3%	1	10.07	9.07
δ					$\theta = 0.4$					
1.2	480	320	300	7.6%	59.1%	80.6%	21.5%	18.07	27.31	9.24
1.4	600	400	380	12.0%	49.8%	76.7%	26.9%	14.42	24.62	10.2
1.8	820	580	540	7.6%	37.9%	66.3%	28.4%	12.36	22.39	10.03
2	960	680	620	6.6%	30.1%	60.8%	30.7%	10.47	19.91	9.44
2.4	1220	860	800	8.4%	20.5%	53.1%	32.6%	6.19	18.04	11.85
2.8	1480	1060	1000	4.0%	13.2%	44.2%	31.0%	4.77	14.26	9.49

whom search costs are salient. The subsidies were not likely to be useful to all job seekers. In fact, on an average week, between 40% and 50% of the subsidies were collected. This could explain why the model predicts treatment effects that are larger than the estimated effects, at least for high levels of risk aversion.[46] In this sense the model can be said to rationalize the large treatment effects induced by a small change in the cost of search.

Temporary work and job search

An important result from the calibrations deserves special mention: increasing the probability of getting temporary work (θ) while unemployed actually increases the probability that an unemployed person will search for work. This is

46 In addition, it may be the case that I have over- or under-estimated the extent to which transport influence the cost of search. In calibrations performed here the treatment reduces the cost of search from 60 to 40 birr (the average subsidy amount was between 20 and 30 birr per week).

a counter-intuitive finding, since the value of unemployment has been increased. In most job search models, this would be predicted to reduce incentives to find permanent work more quickly (Chetty, 2008). However, in this model temporary work provides income that allows for job search. Forward looking unemployed agents are more likely to pay the cost of job search for a given level of savings because search is less risky. For lower values of θ, search is made risky by the lower probability of getting income after that search decision has been made.

Table 1.11: Risk and job search: Solution for x^*
(Power utility with $p = 60$, $\beta = 0.98$, $\sigma = 0.03$)

	δ (Risk Aversion)					
θ	1.2	1.4	1.8	2	2.4	2.8
0.1	350	450	650	900	1500	2050
0.2	300	400	550	800	1400	1850
0.4	300	350	500	750	1300	1700
0.6	300	350	450	800	1250	1600
0.8	300	350	450	750	1200	1550
1	250	300	450	750	1150	1600

In Table 1.11 I look at the case with power utility, and look the predictions of the model for a range of values for θ, and for δ. The critical value decreases with the value of θ in these calibrations, except for when $\theta = 1$ and very high risk aversion.[47]

This result has important implications. It suggests that job seekers rely on temporary work for money to search for work, and as a method to deal with risk while unemployed. The need to take temporary work seems symptomatic of the cash constraints that job seekers face.

The model assumes that job search and temporary work are non-rival activities in terms of the time of the unemployed. However, in reality taking temporary work could seriously interfere with a job seekers ability to search for work. This

47 Only in cases where θ is very close to 1 do job seekers become relatively indifferent between unemployment and permanent work, and reduce search effort. Here permanent work is still valuable when $\theta = 1$ because permanent work pays more than temporary work. When $Y = W$, $\theta = 1$, of course, no one searches for work for any calibration since permanent work is not preferred in any way.

would drive the result in the paper that the subsidies allow job seekers to take *less* temporary to concentrate on search.[48] In this setting an appropriate policy response, if subsidizing search directly were not feasible would be provide money to search for work directly.

I simulate the effect of pure unemployment insurance. That is I set $\theta = 1$, but income while unemployed (now interpreted as indemnity payments rather than wages from temporary work) low at $W = 60 = p$. The model predicts a huge increase in the proportion of individuals induced to search. This suggests that job search could be improved by providing a small income stream to remove the risks associated with unemployment.

1.6.5 *The probability of finding permanent work*

In this section I argue that the increases in permanent employment induced by the transport subsidies could plausibly have been induced by the increased rates of job search predicted by the model and estimated from the experiment. The model assumes that paying for job search leads to a constant probability of finding work. In reality the mechanism I have in mind is that increased job search leads to more information about good jobs, more applications to good jobs, and increased probability of receiving and accepting a job offer.

On average, for the *board* sample, the probability of finding a permanent job for *each week* of search is about 2.5 percent (leading to a 19% probability of permanent employment at endline) for the control group. In the model I use calibrations with $\sigma = 0.03$. In each period, treated individuals are more likely to be searching for work by about 8% for every week, over 16 weeks to endline, relative to a control mean of 65%. This implies that the treatment increases the

48 The Ethiopian government has spent a lot of money on employment programs that have engaged the unemployed youth in forms of unskilled casual work, such as acting as parking attendants or building cobble-stoned streets. The results in this paper argue that for some responds searching for temporary work, these forms of work may provide some support, but that the funds may be better spent allowing job seekers to search, or by subsidizing search directly.

average number of weeks of search by about 1.5 weeks over a control mean of about 9 weeks of search. It also leads to about 1.7 additional weeks of visiting the job boards. If returns to search are linearly and constantly related to the number of search activities, these averages would explain about a 2.9 percentage point increase in the probability of finding a job. The actual effect is estimated to be about 7 percentage points. This seems to inadequately capture the size of the treatment effect found here.

However, I would argue that the returns to search exhibit non-linearities in the time spent searching. I find that the combined effect of the treatment on search across all weeks is that treated respondents are far more likely to have searched during *all* weeks by a very large margin. Among the board sample, treated respondents are about 50% more likely to have search in all (or all but one) weeks of the study. The median number of trips to the boards among treated individuals is 18 compared to 12 in the control group, for the full sample.

If noticing that one job vacancy that others missed makes all the difference, or an application is most likely to be successful if one has already made a few applications in recent weeks, then the marginal returns to search are highest when search is sustained and persistent over weeks. Then shifting the distribution of search persistence at the right tail, as the treatment did, is likely to have a larger impact than the impact on the mean would suggest, which would explain the 7% treatment effect on permanent work estimated in the paper.

1.6.6 *Summary*

The model generates the following results:

RISK AVERSION: Job seekers give up search when savings fall below a critical value x^*, which is usually many multiples larger than the cost of search itself.

Discouragement: Job seekers who start with enough income to search ($x > x^*$) give up search quickly when the costs of search are high. Reduced search costs lead to savings being run down more slowly.

Persistence: Because savings are run down more slowly the model predicts that the treatment effects on search should be persistent for individuals who have not run down their savings to the steady state by the end of the experiment period. The keep buffer savings that allow them to search for longer.

Volatility: Individuals who have savings close to x^* alternative between searching and not searching, a pattern confirmed in the data. They save up to search and then run down their savings again when they search. In the model, subsidies lead to more weeks searching when individuals are at this steady state.

Temporary Work: Rather than discourage job search by making unemployment more attractive, improving the probability of temporary work facilitates job search and makes it more likely that an agent will search for a given level of savings.

Heterogeneity: The treatment effects should be larger for individuals who are more cash constrained at the baseline. Individuals who are very wealthy to start with, or receive income from their family, should not be constrained by the costs of search.

1.7 Mechanisms & Persistence

In this section, I provide further experimental evidence to corroborate the intuition and predictions of the theoretical model. I also investigate some of the possible alternative explanations for the results, such as effects on aspirations, motivation and learning about job search. I find no convincing evidence for any of these alternative explanations.

Estimation of heterogenous treatment effects show that individuals who came from poorer households, or had lower levels of savings at the start of the subsidy period, had significantly larger responses to treatment than wealthier individuals. In addition I confirm that the treatment effects were persistent, in the short run, after the subsidies had ended: confirming the prediction of the model that the treatment could allow job seekers to maintain their savings and thus search for longer even after the treatment ended.

The transport treatment, and the initial increase in job search among respondents, could have had a range of additional effects other than the price and income effects described here. The treatment may have changed the information set of respondents, by alerting them to the existence of the job boards. I argue that this unlikely, and find no statistical evidence for changes in respondents perceptions or information about search. The regular phone calls could have induced additional search by acting as reminders to search. I use a control group who received no phone calls to show that there is no evidence of Hawthorne effects of this kind. Similarly, I find no impact on reservation wages, attitudes or aspirations. I argue that while some of these alternative mechanisms may have been at play, the evidence points to the cash constraints story as the dominant explanation, for which there is the strongest evidence.

1.7.1 *Heterogeneous Treatment Effects*

I estimate differential treatment effects for individuals above and below median household wealth, expenditure and savings, and as in equation 1.3, and test for equality of these coefficients. The model would predict that wealthier individuals should not be effected by the subsidies. They receive enough income from their families to always have enough wealth to keep searching. In other words they'd be above the critical level x^* implied by the model. Poorer individuals for whom the costs of search are more salient, should benefit more.

For the *board* sample, who were more likely to respond to the treatment, I find clear evidence that poorer individuals benefited more from the treatment than wealthier ones. Individuals from poorer households were more likely to find permanent jobs, or any work at all, after receiving subsidies, while individuals with low savings at baseline saw a disproportionately large impact on the probability of being discouraged at endline. The effects on poorer households are large and significant, although I am not always able to reject the test of equality between the two groups in such small sub-samples.

The results for the *city* sample are less clear. I look baseline expenditure as a proxy for cash constraints, I find that those with low expenditure and savings seem to have benefited more from the treatment, they are particular less likely to be discouraged. The results by household wealth are not as strong.[49]

According to the model, poorer individuals should be more effected by the price of search. In Table 1.13 I show that the average impact on search, across all weeks, is higher for poorer individuals and the F-stat on the different between coefficients shows that this difference is significant at the 10% level. The impact on richer households is not significant on its own, suggesting that the constraints of job search do not bind for these individuals.

Effects by unemployment duration

The model presented in this paper is stationary- the length of time that a person has been employed or searching for work does not have any effect on their future labour decisions, for a given level of savings. However, there are many ways that unemployment history could matter. If individuals with strong labour market attachment, who have been out of work for a shorter period of time, have more job contacts or think that they may be rehired by their industry, they may have less need for the job boards and official applications as a means of search. Therefore

49 However for this sample the household wealth measure may not be the best for measuring cash constraints: individuals who were living at home with their parents were more likely to appear wealthy, whereas individuals who were living alone may actually have been more cash constrained. To the extent that the ability to find a good job might be correlated with wealth and background, the analysis here could be confounded.

Table 1.12: Heterogeneous effects on endline outcomes by respondent wealth

	Board Sample			City Sample		
	(1) work perm	(2) work	(3) discouraged	(4) work perm	(5) work	(6) discouraged
Heterogeneous Treatment Effects by Household Wealth Index (Above/Below Median)						
poor hh	0.13**	0.12*	-0.002	-0.044	-0.027	0.005
	(0.060)	(0.062)	(0.043)	(0.035)	(0.064)	(0.063)
not poor hh	-0.008	-0.110	-0.074	0.017	0.19**	-0.20**
	(0.078)	(0.10)	(0.051)	(0.054)	(0.067)	(0.069)
R^2	0.085	0.088	0.038	0.076	0.086	0.108
Equality F-stat	1.320	3.460	1.150	1.040	5.300	5.160
Equality p-val	0.260	0.068	0.290	0.330	0.038	0.041
Heterogeneous Treatment Effects by Savings at Baseline (Above/Below Median)						
low savings	0.092**	0.027	-0.027	-0.005	0.074	-0.13**
	(0.044)	(0.064)	(0.038)	(0.032)	(0.064)	(0.057)
not low savings	0.029	0.091	-0.009	-0.030	0.098	0.006
	(0.092)	(0.12)	(0.063)	(0.075)	(0.13)	(0.11)
R^2	0.084	0.075	0.039	0.064	0.074	0.097
Equality F-stat	0.340	0.220	0.074	0.120	0.021	1.160
Equality p-val	0.560	0.640	0.790	0.730	0.890	0.300
Heterogeneous Treatment Effects by Expenditure at Baseline (Above/Below Median)						
low exp	0.074	0.064	-0.100***	-0.028	0.110	-0.16**
	(0.057)	(0.081)	(0.035)	(0.040)	(0.068)	(0.070)
not low exp	0.078	-0.004	0.047	0.019	0.026	0.014
	(0.067)	(0.078)	(0.059)	(0.044)	(0.097)	(0.074)
R^2	0.082	0.081	0.055	0.063	0.097	0.103
Equality F-stat	0.002	0.310	4.150	0.860	0.390	2.480
Equality p-val	0.970	0.580	0.046	0.370	0.540	0.140
Observations	368	369	369	289	289	289

[1] Results are from OLS regressions on endline outcomes. Standard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

an individual with weaker ties to the labour market would benefit more from the treatment. On the other hand, an individual who has been discouraged from search by months of unemployed might be less motivated and thus respond less to treatment.

I categorize individuals as having been employed for a "long duration" if they have spent more than 4 months without work.[50] I find that those who had not worked in more than 4 months benefited far from the treatment than those who had spent less time out of the labour force. These results are in Appendix

50 If they have not worked before, this means they haven't worked since graduating

Table 1.13: Heterogeneous effects on job search trajectories by respondent household wealth (Board sample)

	(1) searchnow	(2) searchnow	(3) searchnow
Treat*Poor HH	0.075***	-0.0097	-0.040
	(0.021)	(0.035)	(0.045)
Treat*Poor HH*Time		0.014**	0.036*
		(0.0056)	(0.020)
Treat*Poor HH*Timesq			-0.0021
			(0.0019)
Treat*Noot Poor HH	0.019	-0.014	0.024
	(0.025)	(0.045)	(0.058)
Treat*Not Poor HH*Time		0.0085	-0.018
		(0.0071)	(0.026)
Treat*Not Poor HH*Timesq			0.0025
			(0.0023)
Observations	2,768	2,768	2,768
R-squared	0.778	0.784	0.784
Equality F-stat	3.04	0.42	2.35
Equality p-val	0.081	0.52	0.13

[1] Results are from OLS regressions on endline outcomes. Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.17. This seems to be the case for both samples, although the results are more striking for the *board* sample. This fits with the story that individuals with relatively weak labour market attachment rely the most on active (costly) job search, and stand to gain the most from subsidies.

Effects by Education

Table A.18 in the Appendix looks at heterogenous treatment effects by education. While I find no evidence of differential impacts on search, education is crucial for the margin at which job seekers can improve their employment outcomes. Individuals with degrees are most likely to find permanent jobs, these jobs are out of reach for individuals with post-secondary education. The treatment has a large impact on the probability of finding permanent work for high educated individuals.

Individuals with diplomas seem to be have fared worse than those with degrees. They have some chance of finding good permanent jobs, but who may find it more difficult, or take longer to find these jobs. The treatment does not lead to higher employment rates for this group, but they are more likely to still be searching at endline. I find large impacts on finding jobs among individuals with only primary school, but these are not permanent jobs.[51]

1.7.2 *Persistence*

In this Section I test for whether the treatment effect of transport subsidies persisted after the subsidies were removed. The end of subsidy date differed from individual to individual by randomize design. If the treatment effects were due purely to a change in the relative price of search, we would expect the effects is dissipate immediately after the subsidies end. However the model of Section 1.6 predicts that treatment prevents individuals from running down their savings, and thus could leave some individuals in a position to keep searching after the subsidies end. This would lead to persistent treatment effects.

In the Table 1.14, I replicate the week 16 (endline) specific treatment effect in Panel A, to show that the effect on search is still present 5 weeks after the treatment ended.

Next I exploit the randomized variation in when treatment ended, with some individuals stopping the program in week 8, three weeks before the others ended it. In each week 9-11 I can compare those who were still receiving treatment to those who had finished it. I create a dummy variable P_{it} equal to one only if participant i was eligible to receive the treatment in the week t. Once the treatment period ended for an individual, this treatment variable "switches off", while T_{it} stays on. In estimates presented here I estimate the impact on T_{it} as

51 They do seem to be better, more formal jobs, in areas further away from respondents' place of living.

Table 1.14: Persistence of treatment effects after subsidies have ended

	(1) work	(2) work perm	(3) searchnow	(4) searchboards	(5) discouraged	(6) days search
Panel A: Average Impacts at Follow up Survey (Week 16 only)						
Treated ever	0.061*	0.040	0.076*	0.068	-0.051*	0.042
	(0.034)	(0.026)	(0.041)	(0.044)	(0.029)	(0.14)
R-squared	0.065	0.081	0.059	0.084	0.073	0.075
Panel A: Average Impacts over weeks 8 to 12						
Treated now	-0.050	0.002	-0.022	-0.006	0.024	-0.200
	(0.047)	(0.027)	(0.047)	(0.040)	(0.029)	(0.20)
Treated ever	0.061	0.006	0.10***	0.085**	-0.082***	0.37**
	(0.040)	(0.025)	(0.037)	(0.040)	(0.023)	(0.17)
R-squared	0.059	0.099	0.066	0.130	0.090	0.046
Panel C: Heterogenous Impacts at Follow up Survey (Week 16 only)						
Treated ever board	0.043	0.080**	0.12**	0.094	-0.019	0.200
	(0.051)	(0.037)	(0.054)	(0.069)	(0.033)	(0.17)
Treated ever city	0.087*	-0.005	0.023	0.045	-0.096*	-0.140
	(0.044)	(0.033)	(0.064)	(0.046)	(0.050)	(0.22)
R-squared	0.067	0.095	0.063	0.107	0.082	0.077
Panel D: Heterogenous Impacts over weeks 8 to 12						
Treated now board	-0.022	-0.011	-0.066	-0.017	0.021	-0.410
	(0.063)	(0.046)	(0.051)	(0.062)	(0.028)	(0.30)
Treated now city	-0.076	0.022	0.030	-0.010	0.029	0.038
	(0.076)	(0.018)	(0.087)	(0.051)	(0.052)	(0.24)
Treated ever board	0.023	0.028	0.15***	0.13**	-0.066***	0.66***
	(0.054)	(0.040)	(0.048)	(0.059)	(0.023)	(0.24)
Treated ever city	0.11*	-0.023	0.039	0.034	-0.100**	0.010
	(0.058)	(0.024)	(0.056)	(0.053)	(0.043)	(0.21)
R-squared	0.062	0.110	0.076	0.178	0.095	0.059
Observations	2202	2202	2202	2202	2202	2202

[1] Dependent Variables are listed at the top of each column. Results are from OLS regressions on phone survey outcomes, with different treatment effects estimated as the average of groups of 4 weeks.
[2] Each coefficient gives the estimate for the treatment effect of *transport* with the sample restricted to the weeks denoted in the first column. The total number of observation used all regressions in each *row* is given in the last column (N)
[3] Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level
[4] Two types of treatment effects are presented: "on" denotes having received the treatment at any time in the past or currently. "here" indicates the impact of the treatment being available in that specific week.

the treatment effect of *ever* receiving treatment compared to P_{it}, the effect of receiving subsidies *now*, in week t.

$$y_{it} = \alpha_t + T_i\lambda + P_{it}\delta + X_{i0}\beta + \epsilon_{it} \qquad \forall t \geqslant 8 \qquad (1.17)$$

Panel B represent the results from Equation 1.17. The coefficient on *Treat Ever* (λ) is large and highly significant while that on *Treat now* (δ) is not. This shows that there is no additional effect of having the treatment in current period, for the last few weeks of the study (weeks 9, 10 and 11). Individuals who have stopped the treatment go on searching as intensely as those who are still receiving it. This shows that the treatment effects are significant, as predicted by the model of cash constraints. These results hold when the results are estimated for the two sample separately, as presented in Panels C and D.

1.7.3 *Alternative Explanations*

The results presented thus far are consistent with a story of credit constraints preventing poor job seekers from being able to invest in job search at an individually optimal level. The effect of the treatment works by lowering the cost of search relative to remaining unemployed, allowing job seekers to take the risk of searching when they are lowing on savings, and thus allowing them to retain savings to search for longer.

The evidence that the treatment effects were persistent after the end of the subsidies would seem to support this theory. However, there are a number of alternative explanations that could be driving these results that deserve further investigation. This Section will consider some of these competing explanations for the results in this paper, not just for the results on the persistence on the impacts, but also explanations other than those related to the price of search, such as those involving learning, aspirations or beliefs that might be driving the results.

The transport subsidy may have increased job search in the short run simply by nudging, or hinting to respondents that they should search, or reminding them to do so on a regular basis. Then, if a short period of search lead to learning about the nature of job search, or more information about the availability of jobs, or the salaries that they pay, this could lead to respondents searching more after the end of the subsidies because of changes in their information set. Searching more may make someone a "better" job seeker, because they learn how to go about doing it.

Further, since the treatment reduced participation in temporary work during the weeks of the study, it could be the case that individuals were less likely to get "stuck" in temporary jobs, that were hard to leave because of the short run benefits of the pay that they offered. This fits into a category of "scarring" explanations. The treatment may have prevented discouragement by keeping job seekers looking for work for slightly longer. If behaviour is persistent, and discouragement is a mental state that is hard to break out of, this would lead to treated individuals searching more after the end of the treatment period.

The decision to search for a permanent job is one not taken lightly. It is time consuming, and involves a certain fixed cost in getting acquainted with the market, preparing a CV and applications, and keeping up with vacancies, possibly while freeing oneself up from other work obligations, such as in temporary employment. Treated individuals may have built up that critical level of search intensity, that could carry through into later periods.

Hawthorne Effects

Could the weekly phone calls that were given to all treated individuals (and half of the control group) be causing improved job outcomes? The phone calls might have primed individuals to think that search was a worthwhile activity, or offered a higher promise of employment than it actually did, or just induced shame in respondents for not searching. Furthermore, it may have induced false reporting at endline because of the pressure of being asked questions so often.

This experiment was designed to test this mechanism, by randomly assigning half of the control group to not get the phone calls. The main results in this paper are robust to pooling this split control group together or looking at each separately.[52]

If the individuals who were not given subsidies but were called experienced better employment outcomes than those were not contacted at all, this would be cause for concern. It would suggest that the calls had some independent impact on job search that is confounding the main results. It test for an impact of the phone calls by looking at endline outcomes, with data for both individuals who were called during the study, and those who were not.

In Table 1.15 I estimate the coefficient on a dummy variable indicating that respondents had received the phone calls. This coefficient estimates the impact of the calls, while the usual coefficient on "trans" estimates the *additional* impact of receiving the subsidies, over and above the effect of the calls. I find that, at endline, there are few if any statistically significant differences between those with phone calls to those who did not receive them, across a range of specifications. The impact of the subsidies is still large and significant. In this regression, the coefficient on subsidies compares individuals who received the subsidies to individuals who were called (it does not include those were not called).[53]

52 For the search trajectories, of course, we do not have choice: the impacts are estimated only by looking at the sampe who were called.

53 One competing hypothesis could still explain these results; which is that phone calls, in combination with the transport subsidies, together induced the transport group to search more intensively, but without the phone calls, the transport treatment alone could induce increased search effort. I cannot reject this outright, since budgetary and sample constraints prevented me from assigning some individuals to a transport treatment group, without the phone call.

Table 1.15: Impact of the phone call survey on outcomes at endline

	(1) searchnow	(2) searchboards	(3) discouraged	(4) work	(5) work perm
Panel A: Average Impacts at Endline					
TE trans	0.096**	0.081	-0.059*	0.053	0.034
	(0.048)	(0.055)	(0.030)	(0.045)	(0.033)
TE call	-0.029	0.00085	0.011	0.011	-0.010
	(0.049)	(0.047)	(0.044)	(0.053)	(0.035)
Panel B: Average Impacts at Endline by Sample					
TE trans boards	0.13*	0.10	-0.050	0.060	0.10**
	(0.072)	(0.093)	(0.044)	(0.059)	(0.046)
TE trans city	0.050	0.053	-0.073*	0.048	-0.045
	(0.061)	(0.053)	(0.042)	(0.067)	(0.037)
TE call boards	-0.0037	0.0072	0.043	-0.028	-0.067
	(0.069)	(0.075)	(0.044)	(0.064)	(0.055)
TE call city	-0.071	-0.012	-0.035	0.064	0.057*
	(0.072)	(0.052)	(0.087)	(0.087)	(0.029)
Obs	658	658	658	658	657

[1] Dependent Variables are listed at the top of each column. Results are from OLS regressions on phone survey outcomes, with different treatment effects estimated as the average of groups of 4 weeks.
[2] Each coefficient gives the estimate for the treatment effect of *transport* with the sample restricted to the weeks denoted in the first column. The total number of observation used all regressions in each *row* is given in the last column (N)
[3] Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Reservation Wages

Standard search-theoretic models in which reservation wages are endogenous, would be ambiguous about the impact of subsidies on unemployment durations. Lowered searched costs could lower reservation wages thus increasing time in unemployment. This reservation wage model is not entirely applicable to Ethiopia. While respondents may be rejecting job offers, they are far more likely to be doing so the basis of the type of work offered. Specifically, they are after more prestigious, stable jobs.

The result in this paper are consistent with individuals holding out for better quality, permanent, jobs, but I find no impact of the subsidies on reservation wages. Figure A.9 shows no significant change in reservation wages throughout the survey, nor at the endline survey. There is an increase in reservation wages for the treated *boards* individuals of about 6 percentage points at the endline (week 16), but the increase is not statistically significant, even when the sample is restricted to those who who haven't found work. The size of the effect is consistent with the difference in wages between permanent and temporary jobs, suggesting that job seekers decided to pursue those jobs, and adjusted their wages up accordingly.

I do find a significant negative impact on perceptions of average and fair market wages, among the *city* sample. Job seekers in this sample had over-inflated expectations about the wages that they could earn in the market.[54] This result is consistent with the unemployed learning about the true value of jobs through additional search. I find no other effects on other variables such as optimism, expectations about the probability of finding work, or general well-being.

1.8 CONCLUSION

This paper looks at the impact of high search costs on labour market outcomes for cash constrained youth in Addis Ababa, Ethiopia. The job market in this city is characterized by high levels of unemployment, and a growing supply of labour wanting to work in those professions. This growth has been driven by rapid urbanization as well as the the enormous expansion of the secondary and tertiary education system in Ethiopia.

The labour market is plagued by search frictions. Gathering information about job vacancies and applying for those vacancies is time consuming and expensive. But the costs are particularly high for finding the highly sought after jobs that are in short supply. These are the permanent jobs that are found predominantly at the job boards near the center of town. Job seekers must make decisions between paying the costs to search for the jobs that they really want, or taking temporary working that is more easily available closer to home.

54 respondents said they expected to earn about 1500 birr on average at baseline, when in reality, those that found jobs earned little over 1000. There was no such discrepancy for the *board* sample, who seemed to understand the market better to begin with.

I test whether these high costs of job search cause poor labour market outcomes for disadvantaged youth living in particularly dislocated parts of the city. A randomly selected group of individuals were given a weekly transport subsidy covering the costs of two return trips from their place of living in around Addis Ababa to the center of town where the vacancy information boards are located.

My split sample approach allows me to compare how job seekers with different backgrounds, looking for different types of jobs, respond differently to reduced job search costs. The *board sample* is comprised of active job seekers, often of high educational attainment, surveyed in areas around the vacancy boards where they were searching for work. These are respondents who are most likely to self-select into any youth employment programme initiated by government or NGOs.

Four months after participants were first surveyed, individuals in *both* samples receiving the transport money are positively impacted in their labour market outcomes, but these impacts differ across the samples, in line with the types of work available to different types of job seekers. I show that *board* sample participants were more likely to find permanent work, particularly in the professions they want to work in, while those in *city* sample are more likely to be working generally, and the work they are doing tends to more formal and less likely to be part time, or casual. Furthermore the transport subsidies increase job search intensity, for those with and without work, throughout the study.

This paper supports the hypothesis that labour market frictions are constraining the ability of the young and unemployed to enter the labour market. "Flattening" spatial distance seems to have improved their access to employment opportunities that might otherwise have been denied to them, as a direct result of their place of living and financial constraints.

The results suggest that labour markets could be made more efficient, as well as accessible and equitable, to a growing and aspirant urban population by policies that reduce the costs of finding work. This could be done either through improved and subsidized transport for the poor, or more direct measures to make access to information about vacancies and employers more readily available such as online or mobile-phone based matching services or job search assistance programs. The paper also highlights the vulnerability of the youth while in unemployment, and suggests that access to more reliable sources of income could significantly ease the transition from school to work.

BIBLIOGRAPHY

Acemoglu, D. and Shimer, R. (1999). Efficient unemployment insurance. *Journal of Political Economy*, 107(5).

Ardington, C., Case, A., and Hosegood, V. (2009). Labor supply responses to large social transfers: Longitudinal evidence from South Africa. *American economic journal. Applied economics*, 1(1):22–48.

Banerjee, A., Galiani, S., and Levinsohn, J. (2007). Why has unemployment risen in the new South Africa. *IPC Work Paper Series Number 35*.

Beam, E. (2014). Incomplete information in job search : Evidence from a field experiment in the Philippines. *Working Paper, National University of Singapore*, pages 1–66.

Betcherman, G., Olivas, K., and Dar, A. (2004). Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries. *World Bank Social Protection Discussion Paper Series*.

Blattman, C. and Dercon, S. (2015). More sweatshops for Africa? A randomized trial of industrial jobs and self-employment. *Mimeo*.

Broussard, N. and Teklesellasie, T. G. (2012). Youth unemployment: Ethiopia country study. *International Growth Center: Working Paper*.

Browning, M., Crossley, T. F., and Smith, E. (2007). Asset accumulation and short-term employment. *Review of Economic Dynamics*, 10(3):400–423.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5):1671–1748.

Buera, F. J. (2009). A dynamic model of entrepreneurship with borrowing constraints: theory and evidence. *Annals of finance*, 5(3-4):443–464.

Cameron, C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90:414–427.

Card, D., Chetty, R., and Weber, A. (2007). Cash-on-hand and competing models of intertemporal behavior: New evidence from the labor market. *The Quarterly Journal of Economics*.

Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy*, 116:1197–1197.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Quarterly Journal of Economics*, 128(2):531–580.

Danforth, J. P. (1979). On the role of consumption and decreasing absolute risk aversion in the theory of job search. In McCall, J. J. and Lippman, S. A., editors, *Studies in the Economics of Search*, pages 109–131. North-Holland, New York.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, 59(5):1221–1248.

Dinkelman, T. (2011). The effects of rural electrification on employment: new evidence from South Africa. *American Economic Review*, 101(7):3078–3108.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in Peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Fields, G. S. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. *Journal of development economics*, 2(2):165–187.

Franklin, S. (2012). Enabled to work: The impact of housing subsidies on slum dwellers in South Africa. *Mimeo*.

Frison, L. and Pocock, S. (1992). Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Statistics in medicine*, 11:1685–1704.

Groh, M., Krishnan, N., McKenzie, D., and Vishwanath, T. (2012). Soft skills or hard cash? the impact of training and wage subsidy programs on female youth employment in Jordan. *Policy Research Working Paper Series*.

Haile, G. (2005). The nature of self-employment in urban Ethiopia. In *Conference on the Ethiopian Economy*.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American Economic Review*.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, pages 153–161.

Holzer, H. J. (1991). The spatial mismatch hypothesis: What has the evidence shown? *Urban Studies*, 28(1):105–122.

Ibarraran, P., Garcia, B., and Ripani, L. (2012). Life skills , employability and training for disadvantaged Youth: Evidence from a randomized evaluation design. (6617).

Ihlanfeldt, K. R. (1997). Information on the spatial distribution of job opportunities within metropolitan areas. *Journal of Urban Economics*, 41(2):218–242.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics*, 127(2):753–792.

Kain, J. F. (1992). The spatial mismatch hypothesis: three decades later. *Housing policy debate*, 3(2):371–460.

Kling, J. R., Liebman, J., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75:83–119.

Krueger, A. B. and Mueller, A. (2011). Job search, emotional well-being, and job finding in a period of mass unemployment: Evidence from high-frequency longitudinal data.

Kumar, A. and Barrett, F. (2008). Stuck in traffic: Urban transport in Africa. *AICD, Background Paper, World Bank*.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2):139–191.

Mains, D. (2013). *Hope is Cut: Youth, unemployment, and the future in urban Ethiopia*. Temple University Press.

Maloney, W. F. (2004). Informality revisited. *World Development*, 32:1159–1178.

Marshall, A. (1890). *Principles of economics*. Macmillan, London.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2):210–221.

Moretti, E. (2014). Cities and Growth. *International Growth Centre*.

Phillips, D. C. (2014). Getting to work : experimental evidence on job search and transportation costs. *Labour Economics*, 29:72–82.

Pissarides, C. A. (2000). Equilibrium unemployment theory. *MIT Press Books*.

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, 50(1):203–219.

Serneels, P. (2007). The Nature of unemployment among young men in urban Ethiopia. *Review of Development Economics*, 11(1):170–186.

UN-Habitat (2005). *Addis Ababa urban profile*.

Vereshchagina, G. and Hopenhayn, H. a. (2009). Risk taking by entrepreneurs. *American Economic Review*, 99(5):1808–1830.

Verick, S. (2011). Giving up job search during a recession : The impact of the global financial crisis on the South African labour market. *Journal of African Economies*, pages 1–29.

Wheeler, C. H. (2006). Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics*, 60(2):162–184.

World Bank (2012). *World Development Report 2013: Jobs*. World Bank, Washington, D.C.

Yntiso, G. (2008). Urban development and displacement in Addis Ababa: The impact of resettlement projects on low-income households. *Eastern Africa Social Science Research Review*, 24(2):53–77.

Zenou, Y. (2009). Urban Labor Economics. *Cambridge Books*.

A

APPENDICES FOR CHAPTER 1

A.1 WHAT ARE JOBS LIKE, AND WHO FINDS THEM?

Over the 16 weeks that I follow my survey of job seekers I observe enormous changes in their lives and job market outcomes.

Of the 658 individuals interviewed in the endline survey, 359 of them were working, compared to only 168 of those individuals at the baseline survey just 16 weeks earlier. Only 183 of the individuals working at the endline had ever held a any kind of job before in their lives. And among the 186 individuals working at baseline (28% of the sample) less than one third kept those jobs through to the endline 4 months later. A similar proportion were no longer working at the endline, and the remaining 40% of were working in different jobs. Half of those who weren't initially working, had found work by the endline. These transitions provide evidence of considerable volatility of the labour outcomes of individuals in this population. The high-frequency data shows even more volatility in week to week employment outcomes.

Table A.2 provides a picture of the jobs available in Addis Ababa. This is not meant a representative sample of the labour market in Addis Ababa, rather it provides a picture of entry level jobs found by young people. It gives an overview of what jobs are like, what attitudes are to different types of labour, and who gets what types of jobs.

Table A.1: Descriptions of job market outcomes and characteristics by individuals characteristics

	% of sample	Job	Perm job	Casual job	Monthly Wage	Hourly wage	Hours work	Paid Monthly	In Office	Firm Size	Dissatisfied	Referral	Board
Full Sample	100%	54.6%	14.3%	10.7%	1287	10.9	40.4	25.5%	8.72%	71.6	22.3%	13.6%	16.4%
Panel A: By Sample													
City	44%	48.1%	5.88%	15%	1107	11.5	37	15.4%	2.2%	51.5	16.6%	16.5%	2.93%
Board	56%	59.6%	20.9%	7.12%	1401	10.6	42.6	34.1%	14.2%	85.7	26.8%	11.1%	27.9%
Panel B: By Gender													
Male	78.1%	58.4%	14.4%	12.9%	1367	11.8	40	23.7%	8.84%	73.9	25.5%	15.9%	18.1%
Female	21.9%	41%	13.9%	3.03%	882	6.72	42.5	31.8%	8.33%	60.6	11.1%	5.3%	10.6%
Panel C: By Period of Migration to Addis Ababa													
Born AA	36.2%	49.6%	9.66%	10.2%	1178	11.9	36.8	18.1%	6.19%	64.4	14.7%	16.8%	8.85%
Since Birth	19.5%	51.6%	10.2%	13.3%	1496	10.3	43.1	21.7%	5.83%	49.1	24.2%	11.7%	7.5%
Last 5 Yrs	24%	62%	16.5%	12.2%	1175	9.22	43.5	30.9%	9.35%	67.2	27.8%	13.7%	23%
Last 1 Yr	20.2%	57.5%	24.1%	7.21%	1419	12.2	39.9	37.8%	16.2%	113	27.6%	9.01%	33.3%
Panel D: By Education Level													
Grades 0-9	19.3%	49.6%	8%	14.7%	1180	9.53	41	13.8%	3.45%	38.2	15.2%	12.1%	6.03%
Secondary	22.8%	51%	9.52%	10.9%	1326	11.2	40.4	19.7%	1.46%	54.2	23.8%	19.7%	7.3%
Vocational	9.44%	55.7%	9.84%	8.47%	1078	8.95	41.1	25.4%	5.08%	38.8	26.2%	20.3%	11.9%
Diploma	22.8%	59.9%	12.9%	14.5%	1261	10.6	40	31.3%	10.7%	82.8	21.8%	14.5%	20.6%
Degree	25.7%	55.1%	26.5%	5.63%	1439	12.2	41.4	34.5%	20.4%	126	24.6%	4.93%	31%
Panel E: By Year of last Education attendance													
Last 1 Yr	37.2%	53.1%	17.2%	8.56%	1293	12.9	39.5	29.3%	13.5%	104	24.9%	9.01%	22.1%
13-36 Months	27%	58.8%	15.3%	13%	1212	9.79	41.8	29.6%	8.02%	75.3	23.7%	14.8%	17.3%
+ 3 Years	35.8%	52.8%	10.2%	11.4%	1343	9.92	40.3	18%	4.27%	35.2	18.7%	17.5%	9.48%

The first two Columns *Job* and *Perm Job* give the percentage of respondents have jobs or permanent jobs, respectively, whereas the later columns give the average statistics for respondents of a certain type who *have employment*. These results are broken down by job seeker characteristics in the various Panels. So for individuals working in construction, of course everyone has a job, but only 6.45% of these jobs are permanent, and 38.2% were found via a referral. For individuals born in Addis Ababa (Born AA) 49.3% had jobs, and 18.3% of the jobs found by these individuals were found at the job boards.

A few notable statistics are facts mentioning in Table A.2. *Boards* individuals with jobs are far more likely to have found them at the vacancy boards, or got them by applying for through formal channels (getting the job with an interview). Many still find out about their jobs through social networks, but far fewer than those in the *city sample*. But as the panel describing jobs by the method that was used to find them shows, the jobs found at the boards look a lot better. They are more likely to be permanent, pay more, and often require formal applications.

Table A.2: Descriptions of job market outcomes and characteristics by job type

	% of sample	Perm job	Casual job	Monthly Wage	Hourly wage	Hours work	Paid Monthly	In Office	Firm Size	Dissatisfied	Referral	Board
All Jobs	100%	26.3%	19.9%	1287	10.9	40.4	47.4%	16.2%	71.6	40.9%	25.2%	30.5%
Panel A: By Job Activity												
Construction	29.5%	5.56%	41.1%	1388	12.6	37.2	12.2%	1.11%	25.8	57.8%	38.9%	13.3%
o/ Daily Labour	6.23%	0%	68.4%	813	9.35	28.8	10.5%	5.26%	67	57.9%	21.1%	0%
Factory Work	6.23%	26.3%	10.5%	860	5.35	47.9	78.9%	5.26%	249	42.1%	15.8%	26.3%
Basic Services	23.6%	18.1%	8.22%	935	9.48	43.5	58.9%	5.48%	25.9	47.9%	23.3%	28.8%
Vocational	11.5%	17.1%	2.86%	1389	14.9	39.1	40%	5.71%	52.5	25.7%	37.1%	20%
Civil Service	5.57%	94.1%	0%	1458	8.42	42.2	82.4%	88.2%	206	47.1%	0%	94.1%
o/ Skilled	17.4%	47.2%	9.43%	1459	11.1	41.1	83%	49.1%	104	32.1%	11.3%	60.4%
Panel B: By Job status												
Permanent	27.7%	100%	0%	1575	10.2	45.5	85.1%	43.2%	139	27.7%	9.46%	63.5%
Temporary	45.7%	0%	0%	1216	9.27	40.8	51%	11%	62.8	47.7%	29%	29.7%
Casual	18.9%	0%	100%	1162	13.7	34.2	7.81%	4.69%	50.4	56.3%	32.8%	3.13%
Self Empl	7.67%	0%	0%	1053	13.1	36.7	19.2%	0%	9	38.5%	30.8%	11.5%
Panel C: By Method job was found												
At Boards	37.4%	48%	2.04%	1397	8.55	45.3	85.7%	39.8%	140	36.7%	1.02%	100%
Networks	62.6%	12.2%	25.6%	1195	11.3	38.5	31.7%	7.32%	40.8	50.6%	45.7%	0%
Panel D: By Job Hiring Method												
Formally	48.7%	53.2%	3.9%	1271	7.75	45.6	94.8%	40.3%	107	31.2%	0%	79.2%
Referral	51.3%	8.64%	25.9%	1296	11.9	38.3	22.2%	2.47%	30.6	50.6%	100%	1.23%
Panel E: By Job Education Requirement												
None	51.3%	8.45%	31.7%	1223	12.8	36.1	22.5%	2.82%	38.1	56.3%	34.5%	6.34%
Secondary	32.9%	25.3%	17.6%	1087	7.95	43.5	63.7%	18.7%	84.8	39.6%	16.5%	39.6%
Degree	15.9%	59.1%	0%	1685	11.3	41.6	79.5%	54.5%	139	38.6%	4.55%	84.1%

Panel D provides a breakdown of labour outcomes by respondent education level. Surprisingly, better educated individuals in my sample do not seem to earn considerably more than those without higher education. While those with degrees do earn more, the difference is not especially large. Evidence from representative surveys of Ethiopia suggest that wages grow with tenure for educated respondents. However, those with degrees are far more likely to have permanent employment, and to have found their jobs formally or at the job boards. Indeed, jobs that have holding a degree as requirement for employment are overwhelmingly 87.2% advertised at the job boards, and require formal applications.

Table A.2 shows detailed descriptions of job outcomes broken down by the different types of jobs observed in the sample. A few key facts are worth mentioning.

Permanent jobs: Individuals who found permanent jobs clearly earn a little more than other types of jobs, but the differences is small, particularly when looking at hourly wages instead of total monthly wages. Permanent jobs offer more hours of work per week,[1] and are undoubtedly less volatile in terms of the work being available from week to week: looking at the high frequency data, very few individuals (11%) holding down a permanent job had spells of unemployment (weeks when they worked one week, but then not the next) whereas 50% of those among those holding temporary jobs had spells of unemployment.

Construction work: One of the most striking and perhaps surprising findings of the survey data is the dominance of construction jobs as a means to make a living for young people in Addis Ababa. In the baseline survey about 25% of respondents and 60% of young men who had work were working in construction[2]. Almost half of these jobs were casual labour jobs (individuals were paid daily, or piece rate salaries) and few were considered permanent jobs. Very few construction jobs were found on the job boards, they tended to be found by going to visit worksites, or hearing about them through social networks.

Interestingly, the wages paid in construction are surprisingly high. On average, these wages were hardly lower than much sought after civil service jobs, with only Other Skilled (non-government jobs usually in specialized occupations such as lawyers or teachers) paying higher hourly wages on average. This may reflect the high premium paid for the kind of difficult labour done in construction, and the enormous demand for this kind of work in the middle of the construction boom currently happening in Addis Ababa. Yet individuals working on con-

1 In an economy where many young workers consider themselves under-employment, in the sense of wanting more hours of work (Broussard and Teklesellasie, 2012), this is a sought after characteristic of a job.

2 almost no women were working in construction

struction sites were more likely (by 15pp) to be dissatisfied with their work, and more likely to be searching (by 12pp) for work while working, when compared to all other jobs. When asked what job they expected to work at, in 6 months time, less than half of all construction workers anticipated still working in construction. These construction jobs are exactly the kind that I have in mind as the kind used for temporary spells of income while workers are looking for better employment opportunities.

The public sector: Government jobs are sought after by the youth in Addis Ababa.[3] I distinguish between civil service jobs, usually office and administrative jobs, which are more prestigious and considered routes to a middle class life (Mains, 2013) from any other kind of government employment. In my baseline survey one third of all individuals with degrees expected to find work in government civil service jobs in the next six months. However, work in this sector is hard to find, and by the follow up survey only 15% of those with degrees were still expecting this type of work, and only 4% had found a civil service position. Discouragement set in quickly. Civil service jobs are almost all permanent positions in large government departments, and are almost exclusively found at the job boards. They are far more likely to be given after a formal job interview, and none were given on the basis of referral alone. However, claims of highly inflated civil service wages appear to be vastly overstated. In fact government jobs pay less than other jobs after controlling for education.

Yet, while not everyone is satisfied with civil service jobs, they are far less likely to be dissatisfied with these jobs than other permanent jobs on average. This satisfaction seems to be drive forces other than the wages paid by these jobs: government employees are more less likely to be dissatisfied with their jobs, but *more* likely to be dissatisfied with wages they are paid in these jobs.

3 For a detailed history of the civil service in Ethiopia see (Mains, 2013).

A.2 DESCRIPTION OF SEARCH BEHAVIOUR OVER TIME

This section looks into the search behaviour of job seekers over time. Here I use a new dataset of regular phone call interviews from ongoing fieldwork with a larger sample of 4000 job-seekers, over more than 35 weeks. I decompose the decline in job search over time into three different dynamics. 1) individuals who were always searching in every week start transitioning between searching and not searching on a regular basis. 2) Individuals who were searching on and off, do so less often and 3) Some individuals who were searching infrequently give up entirely (without finding a permanent job).

The patterns in this new dataset are similar patterns in the data used for the experiment. An individual not searching transitions to searching from one week to the next with probability 26% on average, with the probability of transitioning back to not searching is higher, at 33%. This is a high rate of oscillation, which the theoretical model explains through the movements around the steady-state critical value of savings. The percentage of sample searching in each week, among those who have not found permanent work, drops off from 65% to 35% of individuals searching in a each week over this 36 week period.

This drop off is driven by two effects: individuals who were searching in every week switching to searching only infrequently and individuals who were searching on and off searching less frequently than they were before. In the early weeks of the study, individuals searching in one week were observed searching in the next with probability of 75%. In the last month of the study, the probability of sticking with search in consecutive weeks falls to 62%.

I divide weeks into six distinct 6 week intervals. In the early intervals I see many individuals that search in every week of that interval (46% of the sample, in fact). In the final 6 weeks period, only 17% of individuals are searching in every week. In the interpretation of the theoretical model, that 17% are those who still above the critical threshold x^*. The drop off happens quickly: in weeks 31-36 only 18.7% are searching in every week.

Table A.3: Job Search consistency over time

	% Weeks Searched	*Never Search*	*Some Search*	*Always Search*
Weeks 1-6	64.2%	18.3%	35.5%	46.1%
Weeks 7-12	57.2%	21.4%	44.3%	34.3%
Weeks 13-18	55.1%	23.0%	45.6%	31.4%
Weeks 19-24	47.8%	26.9%	50.0%	23.1%
Weeks 25-30	44.9%	33.0%	44.1%	22.9%
Weeks 31-36	38.7%	40.2%	41.2%	18.7%

By contrast, the number of people who *never search* increases far more slowly.In the model these are the individuals who are permanently discouraged: their incomes never reach above x^*. After 18 weeks only 26% of the sample don't search at all for 6 weeks, up from 18% at baseline. Few individuals give up searching entirely in the early weeks. However after 30 weeks discouragement really sets in, where 40% are never searching.

So the proportion of individuals who transition between not searching and searching seems to increase over time at first, as individuals who previously had enough income to search run out of savings and start having to oscillate between not searching and searching. After almost 5 months some of those who are moving between not searching and searching begin to give up completely, possibly because their family stops giving them income.

Figure A.1: Map of Addis Ababa showing sampling frame and selected EAs

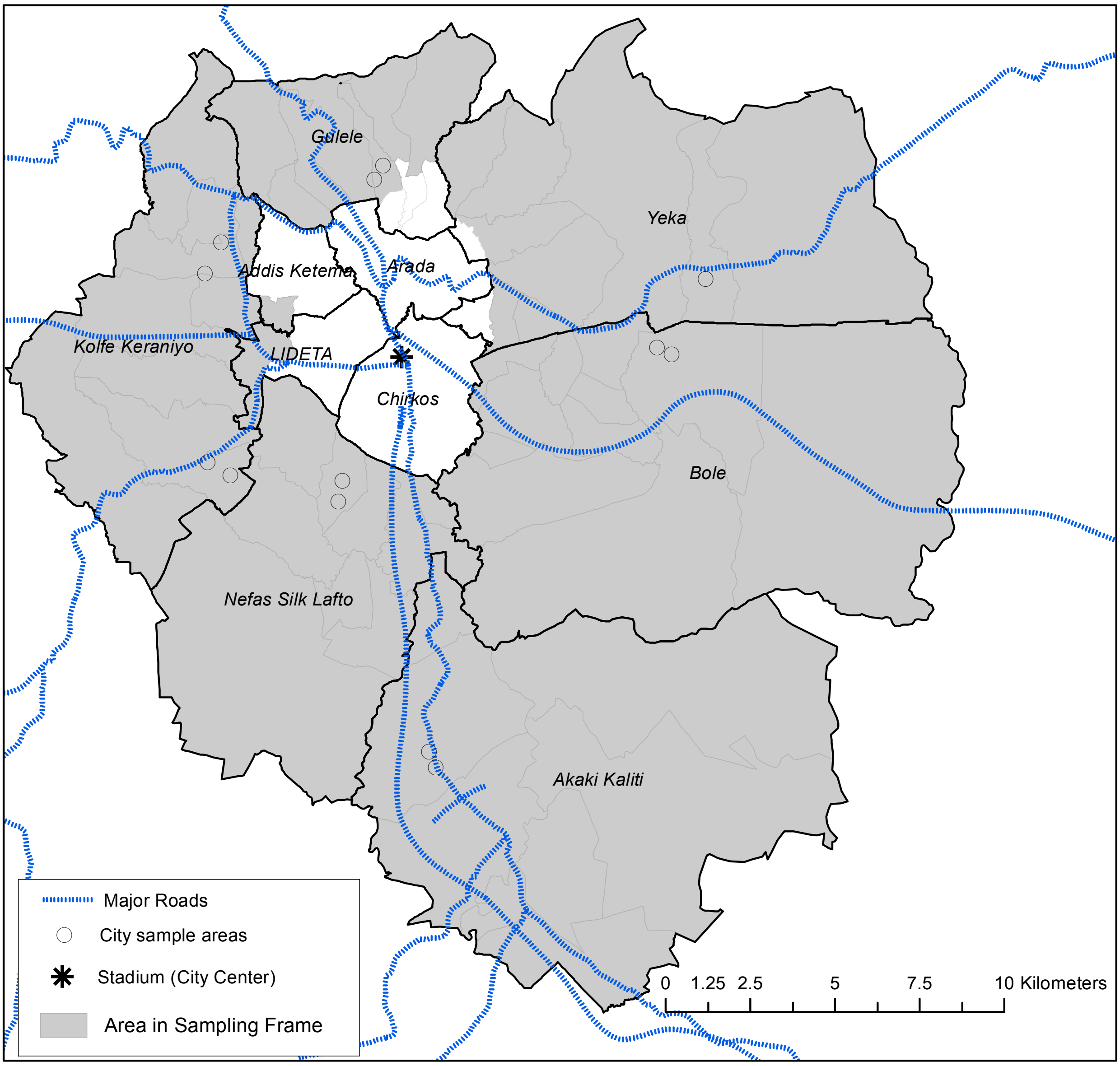


Figure A.2: Trajectory of Treatment effects across weeks in each sample

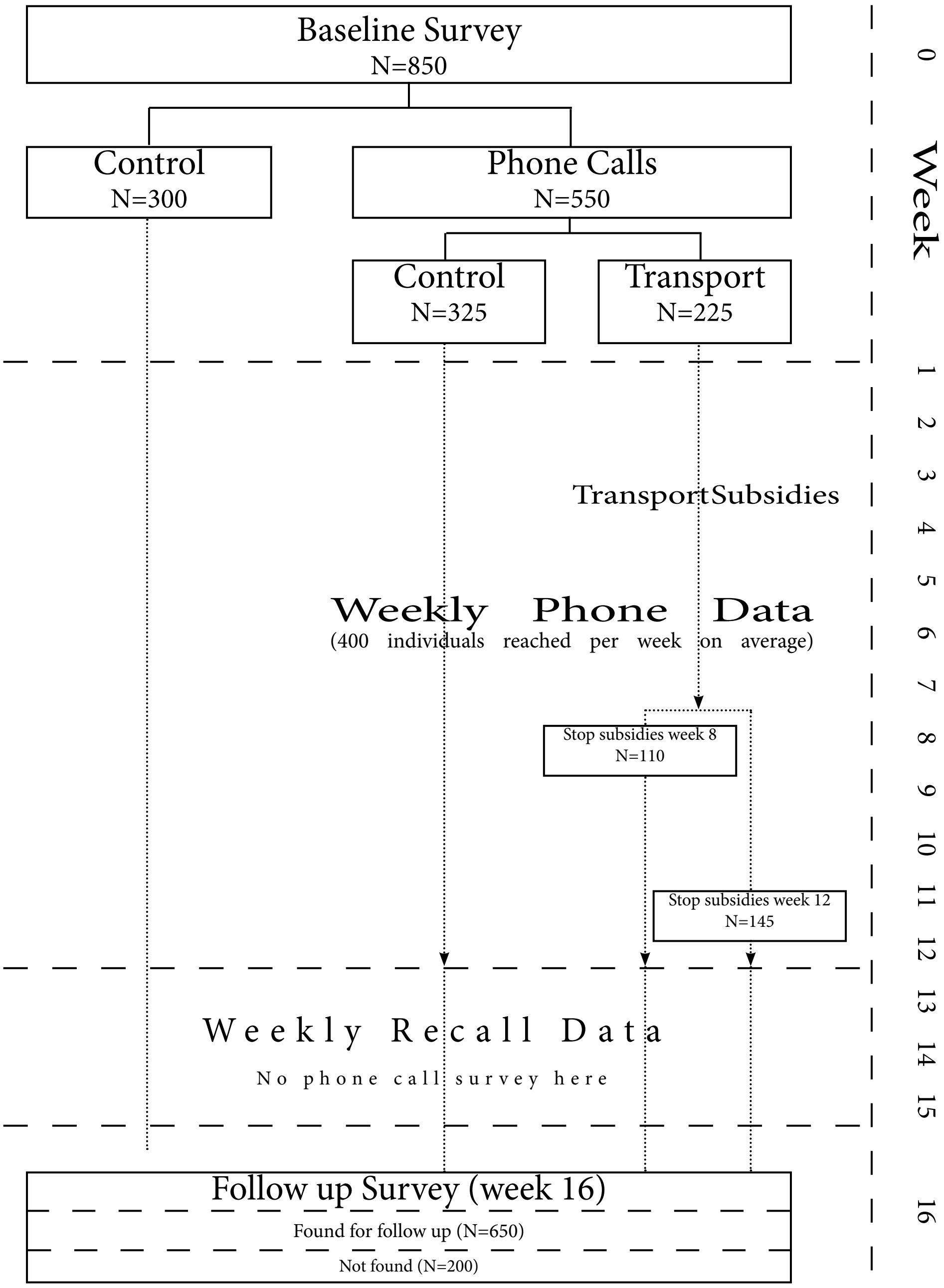


Figure A.3: Impact of treatment on distribution of occupations

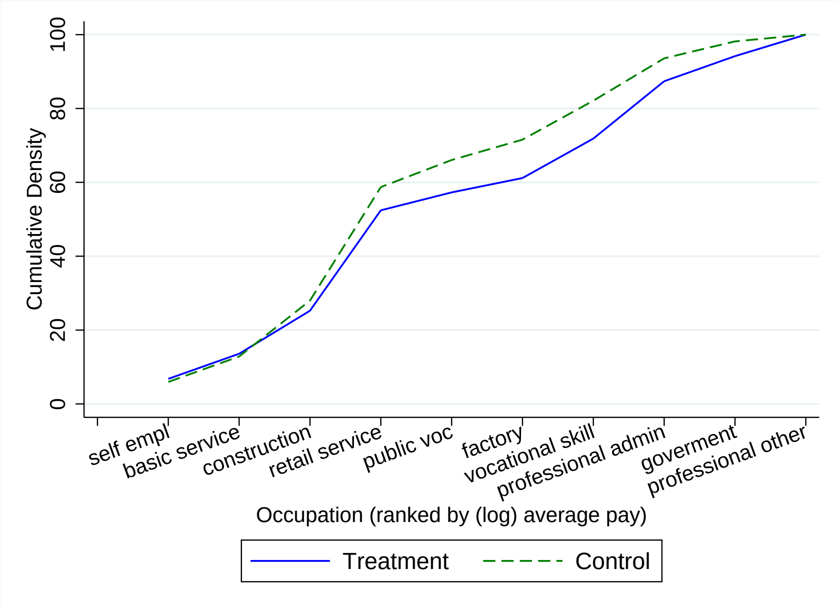


Figure A.4: Composition of the sample for each week by treatment and control: *Board Sample*

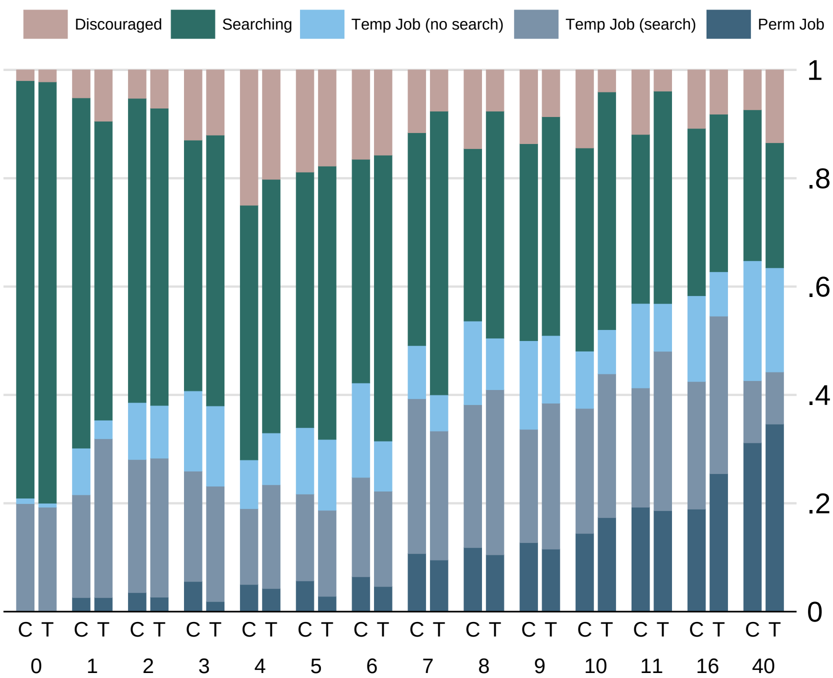


Figure A.5: Composition of the sample for each week by treatment and control: *City Sample*

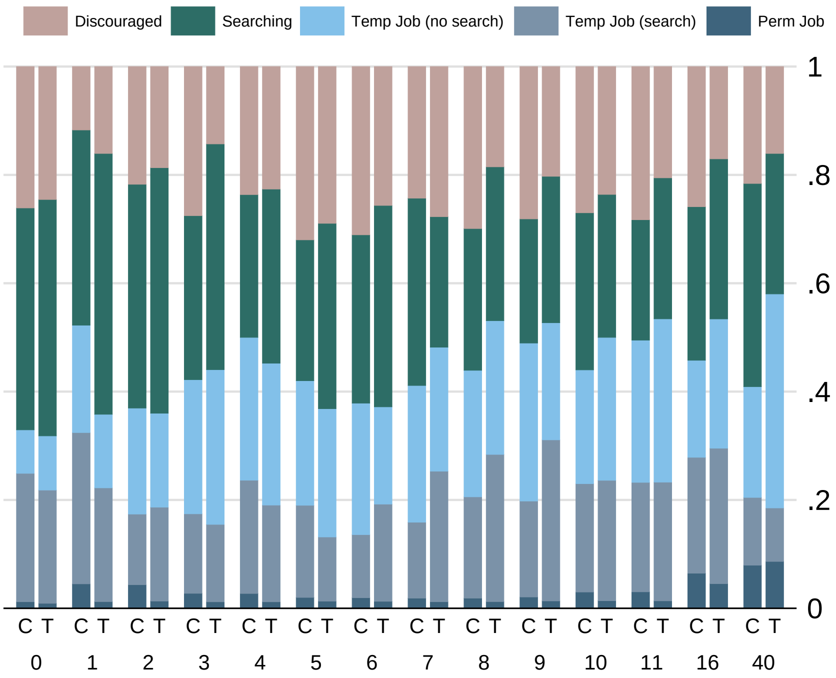


Figure A.6: Impact on visiting the job boards: Trends & treatment effects over time

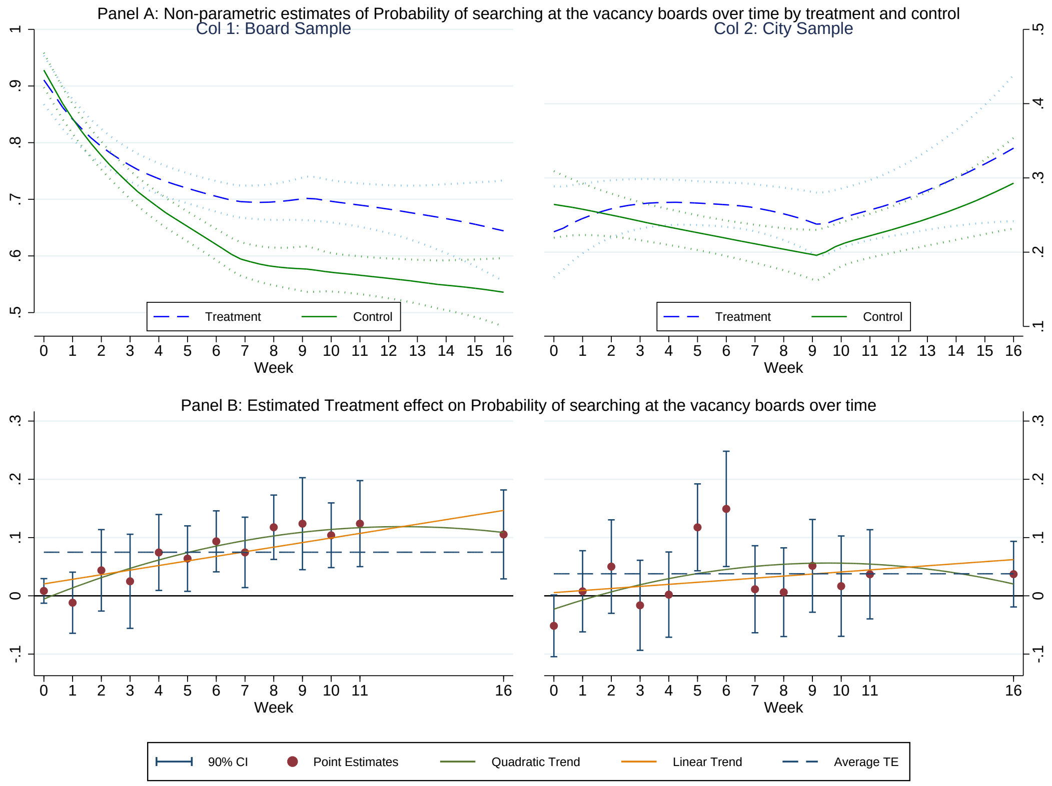


Figure A.7: Impact on having a job: Trends & treatment effects over time

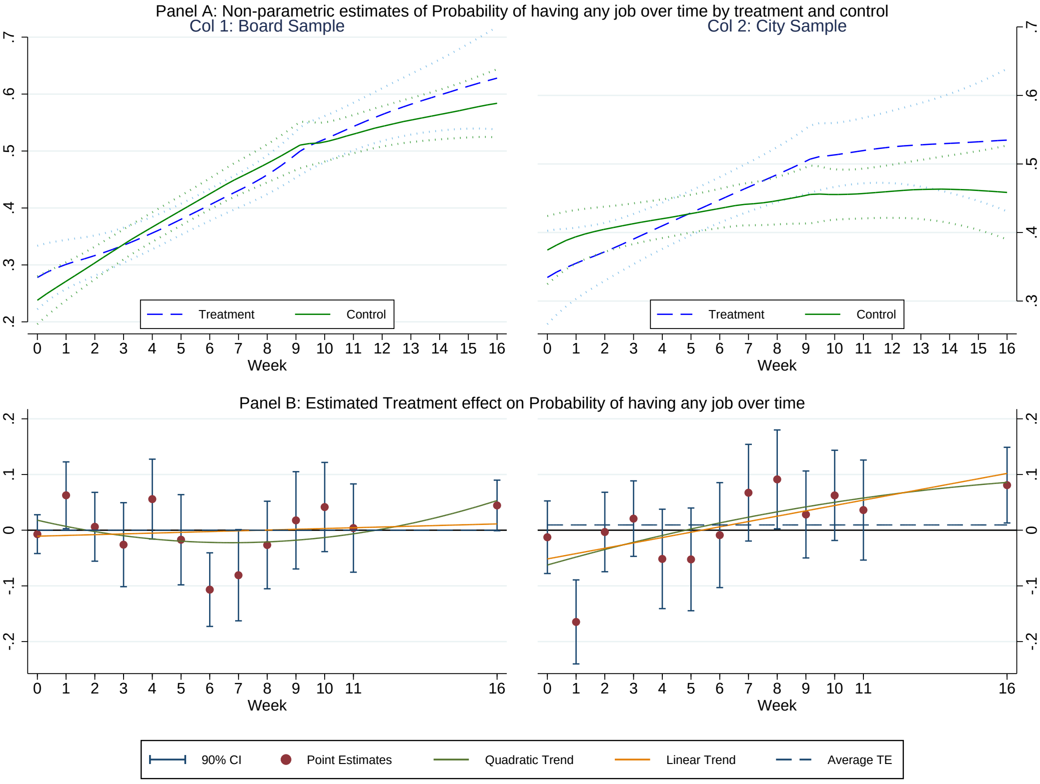


Figure A.8: Impact on discouragement: Trends & treatment effects over time

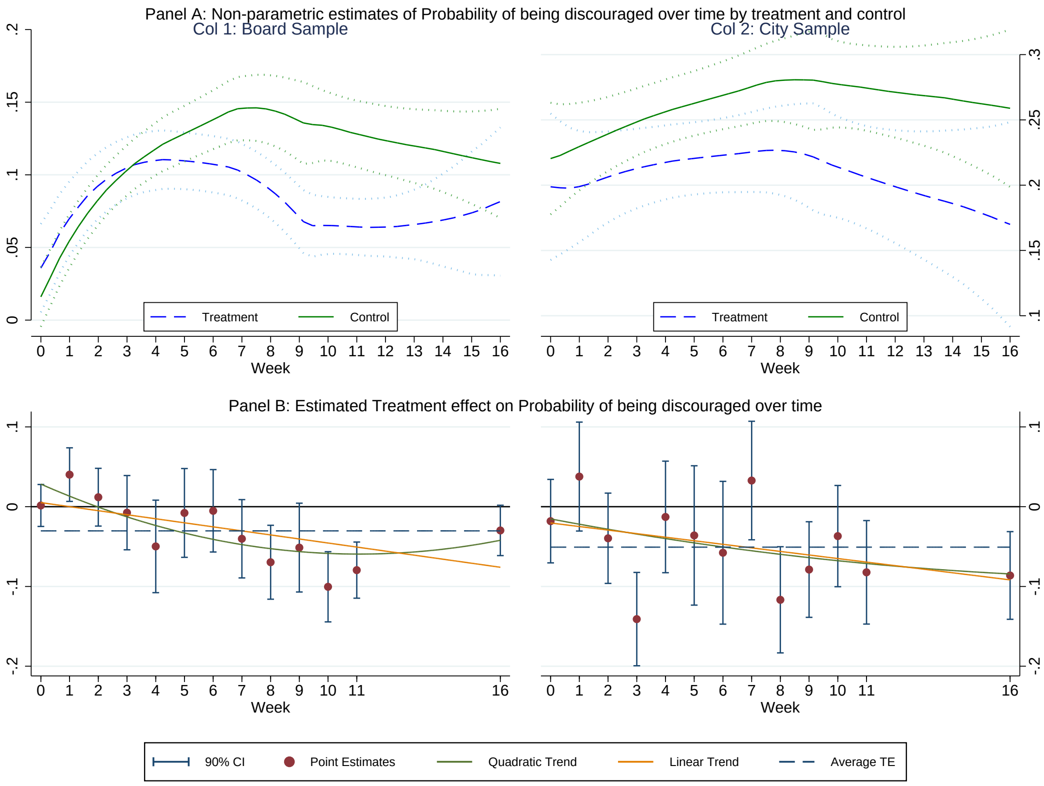
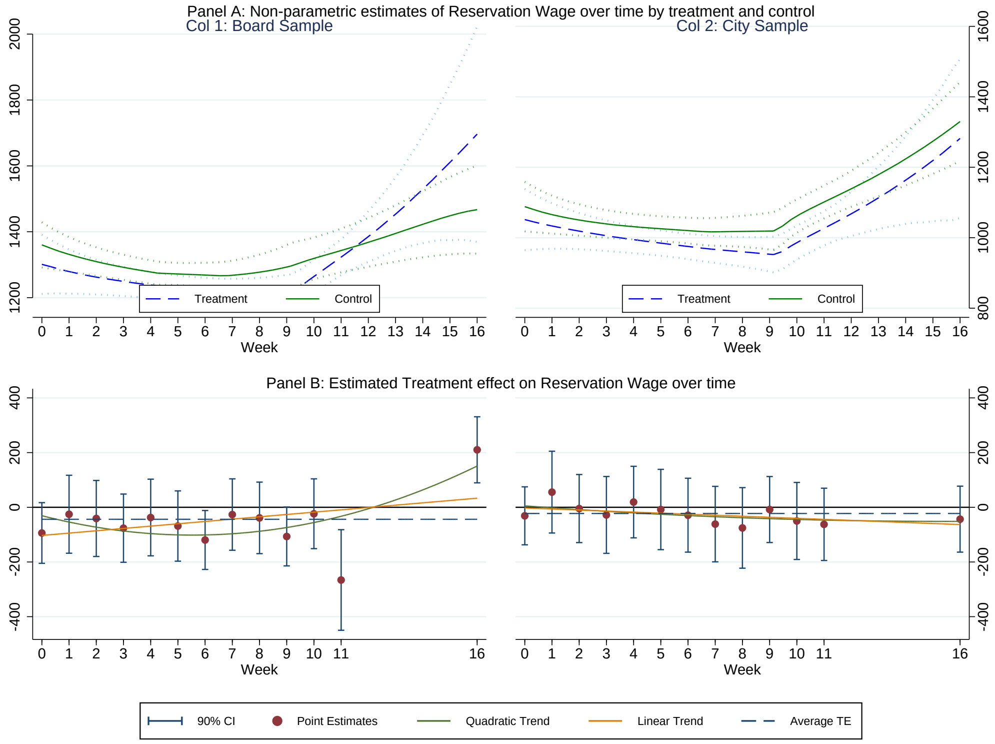


Figure A.9: Impact on reservation wage: Trends & treatment effects over time



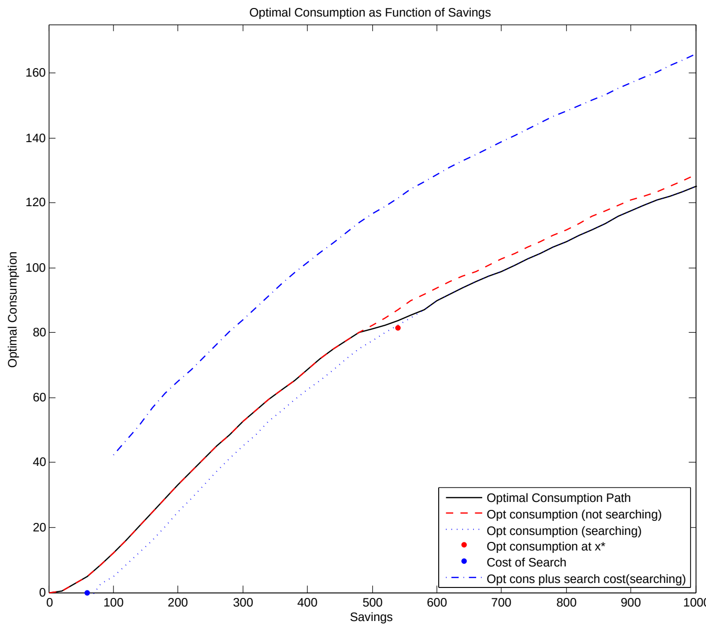


Figure A.10: Optimal consumption paths while searching, and while not searching

Table A.4: Test for balance in full sample and within board and city samples
Panel A: Entire Sample at Baseline

	Full Sample			Boards Sample			City Sample		
	treat	cont	p-val	treat	cont	p-val	treat	cont	p-val
Sample	.539	.54	.982	1	1		0	0	
Work	.256	.258	.934	.201	.201	.983	.319	.326	.892
Permanent Work	.0039	.0065	.643	0	0		.0084	.014	.642
Searching	.829	.829	.98	.971	.973	.912	.664	.66	.935
Visisted Boards	.624	.628	.902	.964	.958	.765	.227	.242	.744
Discouraged	.12	.129	.713	.0216	.018	.794	.235	.26	.609
Hours Worked	7.38	6.06	.197	6.89	5.15	.207	7.95	7.13	.588
Construction	.0891	.0905	.95	.0935	.0749	.497	.084	.109	.454
Female	.217	.223	.848	.129	.132	.948	.319	.33	.838
Diploma	.205	.183	.431	.302	.287	.749	.0924	.0596	.238
Degree	.236	.242	.853	.432	.44	.866	.0084	.0105	.845
Finish Gr 10	.783	.788	.858	.928	.955	.232	.613	.593	.703
Age	23.7	23.4	.162	23.9	23.6	.27	23.5	23.2	.371
Household Size	3.52	3.48	.8	2.76	2.89	.414	4.41	4.18	.321
Head of HH	.225	.223	.952	.302	.263	.392	.134	.175	.311
Amhara	.453	.496	.252	.446	.494	.343	.462	.498	.51
Oromo	.318	.3	.612	.388	.356	.509	.235	.235	.996
Orthodox	.705	.721	.652	.712	.698	.752	.697	.747	.303
Muslim	.101	.113	.595	.0432	.0719	.244	.168	.161	.869
Lives with Family	.256	.268	.706	.367	.383	.739	.126	.133	.844
Born out of Addis	.612	.612	.997	.813	.814	.971	.378	.375	.959
Recent Grad	.345	.401	.123	.468	.551	.0989	.202	.225	.613
Work Experience	.523	.499	.517	.417	.389	.571	.647	.628	.719
Weeks w/o Work	37.6	40.4	.409	37.3	34.4	.43	38	47.4	.1
HH Wealth index	-.0149	.0143	.695	-.112	-.0166	.382	.0985	.0506	.628
Own Room	.229	.223	.853	.23	.201	.472	.227	.249	.636
Kms from center	6.15	6.33	.467	6.4	6.86	.282	5.85	5.71	.481
Weekly expenditure	179	152	.0352	202	174	.115	152	128	.152
Money from fam	84.9	75.1	.395	113	105	.657	52	39.6	.371
Reservation Wage	1225	1282	.355	1326	1398	.379	1106	1146	.668
Observations	258	619		139	334		119	285	

Table A.5: Balance Table (cont): Test for balance across samples after attrition

Panel B: Sample resurveyed at Follow Up

	Full Sample			Boards Sample			City Sample		
	treat	cont	p-val	treat	cont	p-val	treat	cont	p-val
Sample	.556	.563	.859	1	1		0	0	
Work	.242	.263	.579	.182	.205	.616	.318	.338	.739
Permanent Work	.0051	.0065	.824	0	0		.0114	.0149	.812
Searching	.828	.826	.946	.964	.969	.787	.659	.642	.778
Visisted Boards	.631	.65	.647	.955	.954	.971	.227	.259	.571
Discouraged	.116	.126	.723	.0273	.0193	.632	.227	.264	.514
Hours Worked	6.84	6.45	.741	6	5.45	.724	7.9	7.74	.93
Construction	.0859	.087	.963	.0818	.0656	.58	.0909	.114	.554
Female	.207	.224	.632	.127	.135	.839	.307	.338	.601
Diploma	.202	.185	.606	.282	.278	.94	.102	.0647	.269
Degree	.247	.252	.899	.436	.44	.947	.0114	.01	.914
Finish Gr 10	.818	.807	.727	.927	.961	.165	.682	.607	.227
Age	23.8	23.6	.301	23.8	23.7	.653	23.9	23.4	.326
Household Size	3.45	3.43	.869	2.68	2.88	.275	4.42	4.12	.299
Head of HH	.258	.25	.838	.336	.282	.296	.159	.209	.325
Amhara	.449	.509	.164	.445	.498	.356	.455	.522	.29
Oromo	.348	.302	.242	.409	.34	.206	.273	.254	.736
Orthodox	.717	.737	.6	.709	.71	.979	.727	.771	.425
Muslim	.0859	.102	.518	.0455	.0734	.321	.136	.139	.947
Lives with Family	.242	.261	.619	.345	.363	.749	.114	.129	.711
Born out of Addis	.616	.622	.893	.791	.803	.79	.398	.388	.877
Recent Grad	.328	.389	.139	.455	.541	.131	.17	.194	.637
Work Experience	.495	.511	.708	.391	.409	.743	.625	.642	.786
Weeks w/o Work	39	40	.788	37.7	35.3	.564	40.6	46.1	.417
HH Wealth index	-.0276	.0254	.547	-.171	.0025	.165	.152	.0549	.422
Own Room	.247	.224	.511	.264	.208	.247	.227	.244	.763
Kms from center	5.98	6.45	.106	6.09	6.94	.0709	5.85	5.8	.852
Weekly expenditure	183	155	.0422	206	166	.0327	156	140	.476
Money from fam	96.2	77.9	.197	123	107	.42	62.7	40.8	.236
Reservation Wage	1227	1288	.379	1323	1400	.434	1108	1145	.693
Observations	198	460		110	259		88	201	

Panel C: Sample Recontacted (at least once) in the Phone Surveys

	Full Sample			Boards Sample			City Sample		
	treat	cont	p-val	treat	cont	p-val	treat	cont	p-val
Sample	.557	.558	.982	1	1		0	0	
Work	.245	.264	.57	.197	.219	.62	.305	.322	.751
Permanent Work	.0042	.0062	.737	0	0		.0095	.014	.736
Searching	.823	.839	.587	.97	.967	.872	.638	.678	.484
Visisted Boards	.629	.655	.489	.962	.952	.641	.21	.28	.175
Discouraged	.122	.122	.986	.0227	.0222	.974	.248	.248	.999
Hours Worked	7.15	6.25	.407	6.62	5.32	.359	7.82	7.42	.812
Construction	.097	.093	.861	.0985	.0778	.485	.0952	.112	.647
Female	.232	.227	.886	.136	.137	.985	.352	.341	.843
Diploma	.207	.186	.507	.295	.289	.892	.0952	.0561	.196
Degree	.253	.246	.832	.447	.433	.796	.0095	.0093	.988
Finish Gr 10	.793	.795	.945	.932	.963	.168	.619	.584	.552
Age	23.7	23.4	.328	23.8	23.6	.397	23.5	23.3	.58
Household Size	3.51	3.49	.864	2.79	2.92	.45	4.43	4.21	.388
Head of HH	.224	.223	.988	.303	.256	.316	.124	.182	.185
Amhara	.456	.492	.364	.447	.504	.286	.467	.477	.867
Oromo	.333	.308	.49	.402	.356	.372	.248	.248	.999
Orthodox	.705	.725	.565	.705	.711	.892	.705	.743	.471
Muslim	.105	.114	.744	.0455	.0704	.333	.181	.168	.778
Lives with Family	.262	.273	.752	.364	.381	.729	.133	.136	.957
Born out of Addis	.62	.612	.822	.818	.811	.865	.371	.36	.84
Recent Grad	.359	.403	.253	.477	.556	.14	.21	.21	.988
Work Experience	.506	.506	.997	.409	.396	.806	.629	.645	.777
Weeks w/o Work	37.3	40.6	.349	37.7	33.9	.328	36.8	49	.0489
HH Wealth index	-.0057	.0321	.643	-.114	-.0028	.336	.131	.0761	.623
Own Room	.224	.211	.693	.227	.2	.529	.219	.224	.916
Kms from center	6.17	6.39	.41	6.38	6.93	.22	5.91	5.72	.383
Weekly expenditure	177	149	.0397	202	169	.0929	146	123	.222
Money from fam	90.4	74	.18	117	102	.39	56.5	38.7	.235
Reservation Wage	1207	1252	.448	1321	1370	.544	1064	1104	.64
Observations	237	484		132	270		105	214	

Table A.6: Determinants of staying in the survey at first follow up (Week 16)

	(1)	(2)	(3)	(4)	(5)
		Full Sample		Board Sample	City Sample
trans board	0.016	-0.0058	-0.012	-0.013	
	(0.044)	(0.052)	(0.052)	(0.050)	
trans city	0.034	0.0044	0.0025		-0.00075
	(0.045)	(0.041)	(0.038)		(0.037)
call board		0.038	0.042	0.045	
		(0.047)	(0.047)	(0.048)	
call city		0.063	0.078		0.083
		(0.058)	(0.055)		(0.056)
sample board	0.070	0.088	0.072		
	(0.042)	(0.068)	(0.073)		
Grade 0-9			-0.083	-0.083	-0.11
			(0.055)	(0.098)	(0.13)
Secondary			0.0045	0.043	-0.037
			(0.048)	(0.052)	(0.14)
Vocational			0.11*	0.033	0.12
			(0.064)	(0.11)	(0.13)
Diploma			-0.036	-0.044	-0.049
			(0.047)	(0.051)	(0.18)
household wealth i			0.015	0.0022	0.033
			(0.015)	(0.020)	(0.027)
hhsize			-0.0020	-0.0074	0.0024
			(0.011)	(0.015)	(0.014)
female			0.026	0.0048	0.038
			(0.037)	(0.058)	(0.047)
headofhh			0.12***	0.082	0.20**
			(0.046)	(0.058)	(0.077)
living relatives			-0.014	-0.017	0.0098
			(0.039)	(0.046)	(0.091)
amhara			-0.0044	-0.017	-0.0012
			(0.034)	(0.053)	(0.050)
orthodox			0.052	0.035	0.062
			(0.036)	(0.047)	(0.062)
birth migrant			0.013	-0.081	0.061
			(0.042)	(0.063)	(0.056)
age			0.0097*	0.0046	0.014*
			(0.0055)	(0.0091)	(0.0076)
experience			-0.012	0.0044	-0.025
			(0.034)	(0.042)	(0.050)
work			-0.0019	-0.025	0.015
			(0.034)	(0.044)	(0.051)
work perm			0.075		0.076
			(0.18)		(0.19)
married			0.019	-0.056	0.046
			(0.040)	(0.067)	(0.053)
Constant	0.71***	0.67***	0.40**	0.72***	0.26
	(0.034)	(0.060)	(0.16)	(0.23)	(0.21)
Observations	877	877	877	473	404
R-squared	0.006	0.009	0.045	0.028	0.076
F-test	0.36	0.53	1.41	0.85	5.76
Prob > F	0.70	0.71	0.15	0.64	0.0017

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.7: Determinants of staying in the survey at second follow up (Week 40)

	(1)	(2)	(3)	(4)	(5)
		Full Sample		Board Sample	City Sample
trans board	0.016	-0.0058	-0.012	-0.013	
	(0.044)	(0.052)	(0.052)	(0.050)	
trans city	0.034	0.0044	0.0025		-0.00075
	(0.045)	(0.041)	(0.038)		(0.037)
call board		0.038	0.042	0.045	
		(0.047)	(0.047)	(0.048)	
call city		0.063	0.078		0.083
		(0.058)	(0.055)		(0.056)
sample board	0.070	0.088	0.072		
	(0.042)	(0.068)	(0.073)		
Grade 0-9			-0.083	-0.083	-0.11
			(0.055)	(0.098)	(0.13)
Secondary			0.0045	0.043	-0.037
			(0.048)	(0.052)	(0.14)
Vocational			0.11*	0.033	0.12
			(0.064)	(0.11)	(0.13)
Diploma			-0.036	-0.044	-0.049
			(0.047)	(0.051)	(0.18)
household wealth i			0.015	0.0022	0.033
			(0.015)	(0.020)	(0.027)
hhsize			-0.0020	-0.0074	0.0024
			(0.011)	(0.015)	(0.014)
female			0.026	0.0048	0.038
			(0.037)	(0.058)	(0.047)
headofhh			0.12***	0.082	0.20**
			(0.046)	(0.058)	(0.077)
living relatives			-0.014	-0.017	0.0098
			(0.039)	(0.046)	(0.091)
amhara			-0.0044	-0.017	-0.0012
			(0.034)	(0.053)	(0.050)
orthodox			0.052	0.035	0.062
			(0.036)	(0.047)	(0.062)
birth migrant			0.013	-0.081	0.061
			(0.042)	(0.063)	(0.056)
age			0.0097*	0.0046	0.014*
			(0.0055)	(0.0091)	(0.0076)
experience			-0.012	0.0044	-0.025
			(0.034)	(0.042)	(0.050)
work			-0.0019	-0.025	0.015
			(0.034)	(0.044)	(0.051)
work perm			0.075		0.076
			(0.18)		(0.19)
married			0.019	-0.056	0.046
			(0.040)	(0.067)	(0.053)
Constant	0.71***	0.67***	0.40**	0.72***	0.26
	(0.034)	(0.060)	(0.16)	(0.23)	(0.21)
Observations	877	877	877	473	404
R-squared	0.006	0.009	0.045	0.028	0.076
F-test	0.36	0.53	1.41	0.85	5.76
Prob > F	0.70	0.71	0.15	0.64	0.0017

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.8: Impacts on having employment at both endlines (weeks 16 & 40)

Estimator	CM		Basic		Controls		First Diff	
			(1)	(2)	(3)	(4)	(5)	(6)
Week	16	40	16	40	16	40	16	40
Panel A: Average Treatment Effects At Follow Up (Pooled Sample)								
All	0.530	0.550	0.058*	0.063	0.062*	0.066*	0.081*	0.063
			(0.034)	(0.039)	(0.035)	(0.040)	(0.043)	(0.047)
Obs			657	605	657	605	657	605
R^2			0.003	0.003	0.066	0.074	0.062	0.105
Panel B: Treatment Effects At Follow Up by Sample								
Board	0.580	0.650	0.044	-0.013	0.043	-0.012	0.067	0.030
			(0.051)	(0.049)	(0.052)	(0.051)	(0.062)	(0.057)
City	0.46*	0.41*	0.076	0.17***	0.086*	0.17***	0.099*	0.110
			(0.046)	(0.053)	(0.044)	(0.057)	(0.057)	(0.079)
Obs			657	605	657	605	657	605
R^2			0.553	0.586	0.066	0.080	0.062	0.106

[1] The dependent variable is a dummy variable equal to one if the individual reported having work in the last 7 days, measured at endline (week 16). Results are from OLS regressions on endline outcomes.
[2] Panel A gives average ITT effect for the two samples together. Panel B shows results two different samples- "board" and "city"
[3] Standard errors are in parenthesis and are robust to correlation within clusters (Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level.

Table A.9: Week-specific treatment effects on searching for work

	(1) Pooled Effects		(2) Effects by Sample			
	Pool CM	Pool TE	Board CM	City CM	Board TE	City TE
week 0	0.820	0.004	0.970	0.640	-0.000	0.009
		(0.028)			(0.017)	(0.057)
week 1	0.750	0.024	0.840	0.660	-0.000	0.054
		(0.033)			(0.050)	(0.045)
week 2	0.700	0.039	0.820	0.550	0.000	0.083
		(0.037)			(0.042)	(0.067)
week 3	0.570	0.073*	0.690	0.450	0.026	0.12**
		(0.041)			(0.061)	(0.051)
week 4	0.550	0.043	0.630	0.470	0.051	0.032
		(0.042)			(0.067)	(0.052)
week 5	0.540	0.034	0.650	0.430	0.028	0.036
		(0.043)			(0.055)	(0.069)
week 6	0.520	0.12**	0.610	0.430	0.11*	0.130
		(0.053)			(0.059)	(0.091)
week 7	0.620	0.033	0.740	0.500	0.065	-0.014
		(0.039)			(0.049)	(0.058)
week 8	0.560	0.11***	0.650	0.460	0.12***	0.098**
		(0.033)			(0.047)	(0.047)
week 9	0.520	0.14**	0.610	0.420	0.120	0.15**
		(0.055)			(0.081)	(0.067)
week 10	0.590	0.051	0.670	0.500	0.098**	-0.015
		(0.043)			(0.046)	(0.075)
week 11	0.530	0.092	0.620	0.430	0.120	0.049
		(0.055)			(0.077)	(0.078)
week 16	0.580	0.079*	0.610	0.530	0.13**	0.012
		(0.041)			(0.053)	(0.063)
Obs	(5,752)		(5,752)			

[1] The dependent variable is a dummy variable equal to one if the individual reported having a searched for job in the last week (week 16). Results are from OLS regressions on endline outcomes.
[2] CM denotes the mean out the dependent variable for the control group. TE denotes the estimate treatment effect in that specific week, estimated by interacting the treatment variable with week dummy variables.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level.

Table A.10: Trends in the treatment effects on searching for work over all weeks

	(1) Pooled Samples			(2) Board Sample			(3) City Sample		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Treat	0.066***	0.029	0.007	0.073***	0.003	-0.015	0.059*	0.061**	0.035
	(0.021)	(0.021)	(0.025)	(0.027)	(0.028)	(0.026)	(0.034)	(0.030)	(0.044)
Treat X Time		0.0050*	0.015		0.0100***	0.018**		-0.001	0.010
		(0.0026)	(0.0089)		(0.0037)	(0.0086)		(0.0035)	(0.017)
Treat X TimeSq			-0.001			-0.001			-0.001
			(0.00054)			(0.00055)			(0.00099)
CM	0.590			0.680			0.490		
Obs	5,011	5,752	5,752	5,011	5,752	5,752	5,011	5,752	5,752
R^2	0.652	0.686	0.686	0.652	0.686	0.686	0.652	0.686	0.686

[1] For each sampled (Pooled, Board, and City) results from the following models are presented: (a) the constant average treatment effect over all weeks (b) a linear trend in the treatment effect (with the intercept given by "trans" (c) a quadratic function with linear, quadratic and intercept terms.
[2] Dependent Variable is a dummy variable equal to one if the individual reported having take some step to look for work in the last 7 weeks.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.11: Trends in the treatment effects on visiting the vacancy boards over all weeks

	(1) Pooled Samples			(2) Board Sample			(3) City Sample		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Treat	0.058**	0.015	-0.011	0.075**	0.019	-0.009	0.038	0.011	-0.014
	(0.025)	(0.025)	(0.024)	(0.029)	(0.034)	(0.033)	(0.041)	(0.036)	(0.036)
Treat X Time		0.0058**	0.018**		0.0080*	0.021*		0.003	0.014
		(0.0025)	(0.0086)		(0.0042)	(0.011)		(0.0022)	(0.014)
Treat X TimeSq			-0.001			-0.001			-0.001
			(0.00056)			(0.00077)			(0.00082)
CM	0.430			0.600			0.230		
Obs	5,011	5,752	5,752	5,011	5,752	5,752	5,011	5,752	5,752
R^2	0.576	0.620	0.620	0.576	0.620	0.620	0.576	0.620	0.620

[1] For each sampled (Pooled, Board, and City) results from the following models are presented: (a) the constant average treatment effect over all weeks (b) a linear trend in the treatment effect (with the intercept given by "trans" (c) a quadratic function with linear, quadratic and intercept terms
[2] Dependent Variable is a dummy variable equal to one if the individual reported having visited the job vacancy boards in the last 7 weeks.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.12: Trends in the treatment effects on discouragement over all weeks

	(1) Pooled Samples			(2) Board Sample			(3) City Sample		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Treat	-0.040**	-0.008	0.007	-0.030	0.005	0.029*	-0.051	-0.024	-0.021
	(0.019)	(0.021)	(0.017)	(0.024)	(0.024)	(0.017)	(0.032)	(0.037)	(0.030)
Treat X Time		-0.0047**	-0.011*		-0.0051**	-0.016*		-0.004	-0.005
		(0.0018)	(0.0067)		(0.0022)	(0.0082)		(0.0031)	(0.011)
Treat X TimeSq			0.000			0.001			0.000
			(0.00045)			(0.00054)			(0.00074)
CM	0.190			0.130			0.260		
Obs	5,011	5,752	5,752	5,011	5,752	5,752	5,011	5,752	5,752
R^2	0.237	0.241	0.241	0.237	0.241	0.241	0.237	0.241	0.241

[1] Dependent variable is a dummy variable equal to one if the individual reported having not worked and not searched for work in the last 7 days. Column (1) presents average effects across the full sample, while Column (2) estimates different coefficients for the two subsamples.
[2] CM denotes the mean out the dependent variable for the control group. TE denotes the estimate treatment effect in that specific week, estimated by interacting the treatment variable with week dummy variables.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.13: Week-specific treatment effects on the number of days searched per week in each week

	(1) Pooled Effects		(2) Effects by Sample			
	Pool CM	Pool TE	Board CM	City CM	Board TE	City TE
week 0	2.060	0.010	2.060	2.050	-0.100	0.160
		(0.12)			(0.089)	(0.24)
week 1	2.150	0.200	2.510	1.770	0.170	0.160
		(0.13)			(0.20)	(0.16)
week 2	2.010	0.140	2.240	1.740	0.030	0.260
		(0.16)			(0.20)	(0.24)
week 3	1.570	0.32**	1.840	1.290	0.220	0.41**
		(0.15)			(0.21)	(0.20)
week 4	1.410	0.130	1.550	1.290	0.230	0.029
		(0.14)			(0.22)	(0.18)
week 5	1.510	0.091	1.800	1.210	0.026	0.150
		(0.14)			(0.18)	(0.22)
week 6	1.400	0.45*	1.620	1.170	0.57*	0.250
		(0.23)			(0.33)	(0.27)
week 7	1.880	-0.250	2.460	1.270	-0.390	-0.160
		(0.23)			(0.43)	(0.14)
week 8	1.730	0.170	2.080	1.370	0.240	0.047
		(0.25)			(0.46)	(0.15)
week 9	1.400	0.30*	1.640	1.130	0.250	0.34*
		(0.16)			(0.25)	(0.19)
week 10	1.610	0.50**	1.830	1.380	0.80**	0.088
		(0.22)			(0.34)	(0.19)
week 11	1.400	0.38**	1.610	1.170	0.63**	0.011
		(0.18)			(0.27)	(0.18)
week 16	1.810	0.060	1.910	1.680	0.220	-0.120
		(0.14)			(0.17)	(0.22)
Obs	(5,752)		(5,752)			

[1] Dependent Variable is the number of days an individual reported searching for work out of the last 7 days. Column (1) presents average effects across the full sample, while Column (2) estimates different coefficients for the two sub-samples.
[2] CM denotes the mean out the dependent variable for the control group. TE denotes the estimate treatment effect in that specific week, estimated by interacting the treatment variable with week dummy variables.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.14: Trends in the treatment effects on the number of days searched in the last week

	(1) Pooled Samples			(2) Board Sample			(3) City Sample		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Treat	0.18**	0.120	0.054	0.24*	0.030	-0.063	0.110	0.24**	0.200
	(0.088)	(0.084)	(0.090)	(0.13)	(0.12)	(0.094)	(0.12)	(0.11)	(0.15)
Treat X Time		0.007	0.039		0.028**	0.071*		-0.019*	-0.003
		(0.0089)	(0.028)		(0.013)	(0.036)		(0.0098)	(0.042)
Treat X TimeSq			-0.002			-0.003			-0.001
			(0.0017)			(0.0022)			(0.0025)
CM	1.670			1.930			1.390		
Obs	4,949	5,690	5,690	4,949	5,690	5,690	4,949	5,690	5,690
R^2	0.481	0.502	0.503	0.481	0.503	0.503	0.481	0.503	0.503

[1] Dependent Variable is the number of days an individual reported searching for work in the last 7 weeks.. Column (1) presents average effects across the full sample, while Column (2) estimates different coefficients for the two subsamples.
[2] CM denotes the mean out the dependent variable for the control group. TE denotes the estimate treatment effect in that specific week, estimated by interacting the treatment variable with week dummy variables.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.15: Trends in the treatment effects on having a job over all weeks

	(1) Pooled Samples			(2) Board Sample			(3) City Sample		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Treat	0.004	-0.030	-0.020	-0.000	-0.012	0.016	0.009	-0.052	-0.065
	(0.029)	(0.030)	(0.032)	(0.038)	(0.040)	(0.042)	(0.043)	(0.043)	(0.047)
Treat X Time		0.0052*	0.001		0.002	-0.011		0.0097**	0.015
		(0.0029)	(0.011)		(0.0040)	(0.016)		(0.0041)	(0.013)
Treat X TimeSq			0.000			0.001			-0.000
			(0.00062)			(0.00096)			(0.00070)
CM	0.450			0.460			0.450		
Obs	5,011	5,752	5,752	5,011	5,752	5,752	5,011	5,752	5,752
R^2	0.493	0.478	0.478	0.493	0.478	0.478	0.493	0.478	0.478

[1] For each sampled (Pooled, Board, and City) results from the following models are presented: (a) the constant average treatment effect over all weeks (b) a linear trend in the treatment effect (with the intercept given by "trans" (c) a quadratic function with linear, quadratic and intercept terms
[2] Dependent Variable is a dummy variable equal to one if the individual reported having a any kind of paid work in the last 7 weeks.
[3] Standard errors are in parenthesis and are robust to correlation within clusters (70 Woredas within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.16: Iterative 4 week average treatment effects (one regression per coefficient)

	(1) work	(2) work perm	(3) searchnow	(4) searchboards	(5) discouraged	(6) days search	N
weeks 0-3	-0.019	-0.012	0.034	0.0090	-0.011	0.19*	1227
	(0.034)	(0.016)	(0.026)	(0.033)	(0.023)	(0.10)	
weeks 1-4	-0.0044	-0.0091	0.042	0.025	-0.037	0.17	1191
	(0.035)	(0.016)	(0.029)	(0.034)	(0.025)	(0.11)	
weeks 2-5	-0.013	-0.011	0.041	0.040	-0.038	0.15	1186
	(0.039)	(0.016)	(0.032)	(0.034)	(0.030)	(0.12)	
weeks 3-6	-0.028	-0.0060	0.057	0.081**	-0.024	0.20	1175
	(0.039)	(0.018)	(0.036)	(0.039)	(0.035)	(0.14)	
weeks 4-7	-0.028	-0.0068	0.050	0.080**	-0.016	0.063	1194
	(0.040)	(0.019)	(0.036)	(0.038)	(0.033)	(0.11)	
weeks 5-8	-0.011	-0.0027	0.087***	0.080**	-0.044	0.11	1208
	(0.040)	(0.022)	(0.032)	(0.038)	(0.029)	(0.14)	
weeks 6-9	0.012	-0.0025	0.10***	0.076**	-0.058**	0.080	1194
	(0.040)	(0.024)	(0.031)	(0.038)	(0.029)	(0.17)	
weeks 7-10	0.028	-0.0017	0.12***	0.090**	-0.081***	0.33**	1161
	(0.039)	(0.025)	(0.032)	(0.037)	(0.027)	(0.13)	
weeks 8-11	0.023	-0.0062	0.11**	0.100**	-0.077***	0.41***	1141
	(0.040)	(0.028)	(0.042)	(0.038)	(0.026)	(0.14)	
weeks 9-12	0.026	-0.0081	0.092**	0.099**	-0.084***	0.45***	757
	(0.042)	(0.031)	(0.044)	(0.040)	(0.026)	(0.15)	

[1] Dependent Variables are listed at the top of each column. Results are from OLS regressions on phone survey outcomes, with different treatment effects estimated as the average of groups of 4 weeks.
[2] Each coefficient gives the estimate for the treatment effect of *transport* with the sample restricted to the weeks denoted in the first column. The total number of observation used all regressions in each *row* is given in the last column (N)
[3] Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.17: Heterogenous effects on endline outcomes by respondent background

	Board Sample			City Sample		
	(1) work perm	(2) work	(3) discouraged	(4) work perm	(5) work	(6) discouraged
Heterogeneous Treatment Effects by Duration of Unemployed (Long = 4 months+)						
long duration	0.11**	0.14**	-0.086**	-0.008	0.100	-0.19**
	(0.051)	(0.064)	(0.041)	(0.026)	(0.098)	(0.088)
not long duration	0.058	-0.063	0.059	-0.016	0.046	-0.008
	(0.063)	(0.081)	(0.049)	(0.053)	(0.088)	(0.057)
R^2	0.085	0.086	0.054	0.062	0.087	0.114
Equality F-stat	0.310	3.920	5.320	0.022	0.120	3.070
Equality p-val	0.580	0.053	0.025	0.890	0.740	0.100
Heterogeneous Treatment Effects by Migration Status (Migration to Addis since birth)						
birth migrant	0.089*	0.053	-0.010	-0.010	0.030	-0.061
	(0.046)	(0.055)	(0.040)	(0.053)	(0.097)	(0.069)
not birth migrant	0.049	0.001	-0.086*	-0.011	0.110	-0.120
	(0.084)	(0.14)	(0.051)	(0.039)	(0.079)	(0.084)
R^2	0.080	0.075	0.037	0.061	0.075	0.091
Equality F-stat	0.150	0.130	1.400	0.001	0.300	0.230
Equality p-val	0.700	0.720	0.240	0.980	0.590	0.640
Heterogeneous Treatment Effects by Experience						
experience	0.076	0.002	-0.015	-0.037	0.100	-0.046
	(0.076)	(0.095)	(0.047)	(0.046)	(0.079)	(0.055)
not experience	0.083	0.069	-0.033	0.035	0.037	-0.19**
	(0.059)	(0.070)	(0.049)	(0.034)	(0.11)	(0.085)
R^2	0.080	0.075	0.035	0.065	0.075	0.095
Equality F-stat	0.004	0.280	0.063	1.810	0.160	2.050
Equality p-val	0.950	0.600	0.800	0.200	0.690	0.180
Observations	368	369	369	289	289	289

[1] Results are from OLS regressions on endline outcomes, details of the specifications titled are in the REF
[2] Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) ∗ denotes significance at the 10%, ∗∗ at the 5% and, ∗ ∗ ∗ at the 1% level

Table A.18: Heterogenous effects on endline outcomes by respondent education

	(1) work	(2) work perm	(3) searchnow	(4) searchboards	(5) discouraged	(6) work satisfied
Average Treatment Effects At Follow Up (Pooled Sample)						
Pooled Sample	0.057*	0.030	0.082*	0.080*	-0.054*	0.029
	(0.034)	(0.026)	(0.041)	(0.044)	(0.029)	(0.032)
Heterogeneous Treatment Effects by Education Level Completed						
Grades 0-9	0.16**	0.044	-0.130	-0.015	0.003	0.22***
	(0.079)	(0.060)	(0.11)	(0.11)	(0.085)	(0.064)
Secondary	-0.066	-0.051	0.14*	0.110	-0.022	0.018
	(0.084)	(0.043)	(0.081)	(0.086)	(0.052)	(0.061)
Diploma	-0.044	-0.013	0.17**	0.091	-0.095**	-0.056
	(0.075)	(0.046)	(0.079)	(0.079)	(0.046)	(0.061)
Degree	0.23***	0.15**	0.067	0.110	-0.074	-0.001
	(0.073)	(0.074)	(0.080)	(0.071)	(0.049)	(0.066)
Observations	658	657	658	658	658	596
R-squared	0.021	0.055	0.022	0.030	0.020	0.022
Mean of Dependent Variable for Control Group by Education Level						
Grades 0-9	0.390	0.040	0.620	0.280	0.220	0.110
Secondary	0.440	0.090	0.570	0.380	0.190	0.170
Diploma	0.490	0.110	0.650	0.530	0.120	0.200
Degree	0.410	0.180	0.700	0.620	0.080	0.140
All Levels	0.440	0.110	0.640	0.470	0.140	0.160

[1] Results are from OLS regressions on endline outcomes, details of the specifications titled are in the REF

[2] Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level

Table A.19: Impact of the subsidies on finances and aspirations at endline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	log of:								
	total savings	formal savings	total money	expenditure	fair wage	market wage	job prospects	kept occ pref	expected offers
Panel A: Impacts on Aspirations at week 16									
TE Ave	0.029	-0.120	-0.068	0.064	-0.048	-0.036	-0.054	0.085*	-0.061
	(0.20)	(0.23)	(0.12)	(0.062)	(0.043)	(0.048)	(0.041)	(0.049)	(0.32)
Heterogeneity by Sample									
TE board	0.160	-0.340	-0.064	0.087	0.030	0.025	-0.056	0.065	0.150
	(0.28)	(0.28)	(0.17)	(0.087)	(0.057)	(0.064)	(0.054)	(0.063)	(0.31)
TE city	-0.130	0.200	-0.073	0.038	-0.14**	-0.110	-0.050	0.110	-0.300
	(0.28)	(0.29)	(0.17)	(0.093)	(0.058)	(0.070)	(0.063)	(0.076)	(0.59)
Panel B: Heterogenous Impacts on Aspirations by work status week 16									
TE work	0.260	-0.150	0.160	-0.043	-0.040	-0.016	-0.095*	0.090	-0.300
	(0.23)	(0.26)	(0.24)	(0.086)	(0.056)	(0.063)	(0.052)	(0.065)	(0.34)
TE no work	-0.370	-0.037	-0.180	0.150	-0.065	-0.070	-0.010	0.084	0.210
	(0.30)	(0.43)	(0.15)	(0.10)	(0.077)	(0.078)	(0.072)	(0.080)	(0.55)
Heterogeneity by Sample									
TE work-board	0.360	-0.350	0.037	0.003	0.087	0.092	-0.064	0.077	0.067
	(0.32)	(0.31)	(0.32)	(0.11)	(0.084)	(0.089)	(0.072)	(0.078)	(0.41)
TE no work-board	-0.250	-0.360	-0.130	0.200	-0.057	-0.079	-0.047	0.065	0.270
	(0.41)	(0.66)	(0.22)	(0.16)	(0.095)	(0.088)	(0.089)	(0.12)	(0.52)
TE work-city	0.110	0.220	0.390	-0.120	-0.22***	-0.16*	-0.15**	0.120	-0.840
	(0.31)	(0.40)	(0.27)	(0.16)	(0.060)	(0.082)	(0.064)	(0.12)	(0.55)
TE no work-city	-0.500	0.330	-0.240	0.110	-0.073	-0.061	0.026	0.100	0.150
	(0.44)	(0.46)	(0.18)	(0.13)	(0.12)	(0.13)	(0.11)	(0.11)	(0.95)
N	440	225	286	590	594	594	658	450	571

[1] Dependent Variables are listed at the top of each column. Results are from OLS regressions on phone survey outcomes, with different treatment effects estimated as the average of groups of 4 weeks.
[2] Each coefficient gives the estimate for the treatment effect with the sample restricted to the weeks denoted in the first column. The total number of observation used all regressions in each *row* is given in the column (N)
[3] Standard errors are in parenthesis and are robust to correlation within clusters (subcities within Addis Ababa) $*$ denotes significance at the 10%, $**$ at the 5% and, $***$ at the 1% level.

2

ENABLED TO WORK: THE IMPACT OF GOVERNMENT HOUSING ON SLUM DWELLERS IN SOUTH AFRICA

For their assistance in giving me access to specific household GIS data and an early release of the 5th wave of CAPS data, I'd like to thank Jeremy Seekings, Brendan Maughan-Brown and David Lam. Thank you to Magdalen College for providing a great deal of the funding that made possible my research in Cape Town. Thank you Marcel Fafchamps for input, oversight and guidance of the project. Thank you to helpful seminar participants at the CSAE Conference, Oxford OXIGED housing workshop, as well various internal seminars in Oxford. In particular I thank Francis Teal, Stefan Dercon, Paul Collier, Somik Lall, John Muellbauer, Taryn Dinkelman, James Fenske, Simon Quinn, Julien Labonne, Bryan Coulter, Abi Adams, Lucas Brown, Andrew Kerr, Kate Vyborny. I also thank two anonymous MPhil examiners for their detailed and useful comments.

Abstract:

This paper looks at the link between housing conditions and household income and labour market participation in South Africa. I use four waves of panel data from 2002-2009 on households that were originally living in informal dwellings. I find that households experienced large increases in their incomes after receiving free government housing. This effect is driven by increased employment rates among female members of these households, rather than other sources of income. I take advantage of a natural experiment created by a policy of allocating housing to households that lived in close proximity to new housing developments. Using rich spatial data on the roll out of government housing projects, I generate geographic instruments to predict selection into receiving housing. I then use housing projects that were planned and approved but never actually built to allay concerns about non-random placement of housing projects. The fixed effects results are robust to the use of these instruments and placebo tests. I present suggestive evidence that formal housing alleviates the demands of work at home for women, which leads to increases in labour supply to wage paying jobs.

2.1 INTRODUCTION

Substandard housing conditions are considered to be one of the key deprivations suffered by the poor. Currently, it is estimated that over 860 million people live in slums in developing countries, and this number has been growing rapidly over the last decade (UN Habitat, 2003, 2010). Living in informal settlements is associated with a lack of access to running water, electricity, ventilation, security of tenure and access to economic opportunity. Improving or eradicating slums has been a key policy goal of governments, yet there is no clear consensus on what best practice should be.[1]

Slums can be thought to constitute a poverty trap (Marx et al., 2013). Living in a slum is not only an outcome of poverty, a growing body of evidence shows that slum conditions themselves have a detrimental impact on households. Yet improving housing conditions at the cost of relocating households further away from jobs and existing networks could do more harm than good (Barnhardt et al., 2014; Lall et al., 2008). Policies aimed at improving the housing conditions of the poor need to take account the effect that they have on the economic and labour outcomes of slum-dwellers.

This paper examines the links between housing conditions and labour supply. Over the past 20 years the South African government has provided over 3 million free stand-alone houses to its citizens. While the government housing policy has been praised for the extraordinary scale of delivery of housing, it has also been criticised for providing low quality homes in areas far away from jobs. I study the impact of free government housing in South Africa, and find clear evidence of increased incomes among recipient households. The evidence suggests that poor housing conditions constrain the ability of households to take wage-paying work.

1 For an overview of some of the debates in this literature see (Marx et al., 2013; Collier and Venables, 2013; Davis, 2007; Werlin, 1999).

I use longitudinal household data from Cape Town over four waves from 2002 to 2009 to assess the impact of government housing on household income and labour market participation.[2] I test the hypothesis that receiving a government home allows substitution from labour at home to work in the labour market, which in turn leads to increases in household income.

I use the allocation procedure used by the local government to award housing to households as a natural experiment. Recipients of housing were selected because of their proximity to the new housing projects. I proceed in two steps: firstly I use the distance between households' original place of living and locations of newly built housing projects to instrument for individual selection into treatment. To do this I develop a unique maximum likelihood estimator which predicts the probability of receiving government housing from any one of multiple nearby housing projects.

Secondly, I control for non-random locations for the selected sites of new housing projects. I use a set of housing projects that were approved and planned but cancelled for reasons unrelated to the communities they were intended to benefit. I exclude from the sample households that had no projects planned nearby and use only those households that had planned but incomplete projects nearby as a control group. I then repeat the main estimates using only completed projects as instruments.

Both household fixed effects and instrumental variable estimates show that households receiving government housing experience large increases in income, relative to households that do not receive housing. These findings are robust to tests using the cancelled projects. In general the IV results are larger than the fixed-effects results. This is consistent either with measurement error, or a story whereby households with the greatest needs (those struggling economi-

2 It is estimated that between one quarter and one fifth Cape Town's entire population benefited from government housing since 1994 (Seekings et al., 2010).

cally) within communities are awarded the housing, leading to downward bias in the fixed-effect estimates.[3]

I investigate the channels through which improved housing increases household incomes. I find that the rise in household income is due to wage employment, rather than increases in income from rent or self-employment. The effect seems to be driven by increased female labour force participation. Women in treated households are more likely to be working in wage labour. These effects are not present for male households members. However, I do find a significant treatment effect on both both female and male household members' wage earnings.

The concentration of the effect on female labour supply suggests that poor housing conditions place particular burdens on the time use of women. I show that that women in South Africa, particularly in informal settlements, allocate significant time to housework and care. This is consistent with evidence from other developing countries (Berniell and Sánchez-Páramo, 2012). Due to data limitations, I cannot conclusively show that this is the main channel driving the results. I do find that government housing significantly increases electrification, direct access to running water, and modern home appliances, all of which could be saving significant amounts of time for women. In addition, living in the South African slums is hazardous- shack fires are common because of low quality stove and heating devices. I speculate that improved housing could reduce time mending and rebuilding after these kinds of disasters.

Field (2007) argues that improved tenure security frees up time that otherwise would have been spent at home defending the home from expropriation. Most South African households in informal settlements already enjoy de-facto tenure security (Payne et al., 2008), and eviction risks that do exist are unlikely to be

3 It could be because of local average treatment effect interpretation of the results, which I discuss in detail in the paper. Briefly, the instruments identify the effect on households that didn't have to move too far to take up housing, because they were treated by virtue of their proximity to housing.

bolstered by time spent at home.[4] I argue that these results are not driven by tenure security. Instead, South African households face a high rate of crime, and may need to spend time at home to deter intruders. I show that receiving government housing significantly increases feelings of safety in the home.

I contribute to the literature on the relationship between physical living conditions and female labour supply. Labour saving improvements to the lives of the poor can free up time to work in the labour market (Duflo, 2012; Greenwood et al., 2005; Devoto et al., 2011). Dinkelman (2011) and Field (2007) show that home electrification and improved tenure security, respectively, increase female labour supply and earnings by freeing up time from work at home. In the context of housing projects specifically, Keare and Parris (1982) find that provision of tenure and basic services in four countries had positive impacts on employment and income generation.

A related literature looks at the impact of *where* people live on their labour outcomes (Bryan et al., 2014; Ardington et al., 2009; Barnhardt et al., 2014; Franklin, 2015). The housing project studied in this paper moved households very slightly further away from job opportunities. In a setting where distance from jobs is thought to be an important contributor to poor labour market outcomes for black South Africans (Banerjee et al., 2007), I find evidence that government housing is not improving the class and race-based segregation of the city. My study benefits from an IV estimator that estimates the impacts of housing on those that did not have to move too far. In this way my study isolates the impact of housing on household outcomes from the effect of relocation.[5]

Secondly, I contribute to the broader literature on the ways in which slum living constitutes a poverty trap. A growing literature looks at the impact of

4 Ironically, evictions that have come from large townships, as opposed to "squatter" settlements on private land, are made by the government to make way for government housing projects. While these evictions often come with the promise of replacement government housing, some households could miss out, or have been placed in temporary shelters for long periods of time.

5 If the movement induced by housing did have negative impacts on household outcomes, it seems that household incomes rise in spite of this additional distance from jobs.

slum upgrading on health and well-being (Cattaneo et al., 2009; Galiani et al., 2014), considerable evidence shows large impacts on health of improved access to services such as sanitation and running water (Zwane and Kremer, 2007; Pitt et al., 2006; Jalan and Ravallion, 2003; Duflo et al., 2012). Marx et al. (2013) discuss three channels through which slums could act as a poverty trap: human capital and health effects, poor incentives for policy, and under-investment due to weak property rights.[6] I add a forth channel to this list by showing that poor housing conditions can limit the ability of those living in them to participate in the labour force.

Thirdly, this paper provides a rigorous evaluation of a large scale government program that is the subject of much debate and criticism. The scale and political sensitivities of projects like these make them difficult to randomize. A large developed country literature generally draws negative conclusions about the impacts of large scale public housing projects (Olsen and Zabel, 2014).[7] This is the first rigorous evaluation, to my knowledge, of a housing project that provides a complete housing unit free of charge in developing country context.

Indeed, delivery of housing on the scale of millions of households is usually considered infeasible or not cost-effective for most developing countries (Gilbert, 2004). Donors and researchers have tended to focus on evaluations of upgrading and land-titling programs. However, standard policy approaches have had little success at mitigating the expansion of slums (Marx et al., 2013). Housing projects of the kind of implemented by South Africa are popular with governments, because they are so popular with electorates. Ethiopia and Columbia (Gilbert, 2014) are embarking on a housing projects of a similar scale. Rigorous evaluations of

6 Galiani and Schargrodsky (2010) provide strong evidence for the last of these channels in Argentina.

7 The flavour of these arguments are still best summarized by Jane Jacobs (1961), in her seminal text on urban planning: "The method fails. At best it merely shifts slums from here to there, adding its own tincture of extra hardship and disruption. At worst, it destroys neighbourhoods where constructive and improving communities exist and where the situation calls for encouragement rather than destruction".

projects like these are important to guide policy makers, to inform best practices for dealing with informal housing conditions.

Finally I contribute methodologically to a growing literature that attempts to estimate the effects of housing policies and urban policy more generally (Baum-Snow and Ferreira, 2014; Field and Kremer, 2006). Local area instrumental variables are harder to use in a setting of continuous population density. I extend a literature that uses proximity data to instrument for individual selection into projects (Attanasio and Vera-Hernandez, 2004; McKenzie and Seynabou Sakho, 2010). I overcome the challenge of multiple weak instruments and improve the efficiency of my first stage IV estimate by creating a unique maximum likelihood estimator for the effect to create a time-varying set of instruments top predict selection into housing.

The rest of the paper is organized as follows. In Section 2.2 I discuss the context of housing policy in South Africa, as well the role of work at home for women in South Africa. I present a simple model of how housing could increase labour supply for women. Section 2.3 describes the GIS and survey data used in the paper. In Section 2.4 I describe the instrumental variables strategy in detail, and show results from the first stage to show that proximity to housing predicts selection into treatment. Section 2.5 show the main results for household earnings. The mechanisms driving the impacts on income are discussed in Section 2.6, while further robustness checks are in Section 2.7. Section 2.8 concludes.

2.2 SETTING AND CONTEXT

Informal settlements in South Africa grew in the context of the apartheid system of enforced segregation. Relocation of non-white populations to the periphery of cities or remote rural areas has left a persistent pattern of segregation. Poor non-white areas are located far from more prosperous city centres. Migrant labour in cities was highly regulated, making access to urban employment a constant battle, and secure rights to adequate housing in the cities almost impossible (Royston, 1998). Public investment in urban infrastructure and housing in black areas was minimal.

As the architecture of apartheid was dismantled, starting in the 1980's with the repeal of the Group Areas Act, families that had previously been prevented from doing so began to move to the cities in vast numbers. This rate of migration, combined with the poor existing housing stock and South Africa's extremely high and rising employment rate (estimated to be around 24% for those actively seeking work) has led to a housing crisis. Many new urban migrants moved into shacks built in the backyards of existing formal dwellers (Seekings et al., 2010).

When the first democratic government was elected in 1994 there were an estimated 12.5 million people without adequate housing. Only 65% of the total population was housed in formal (cement and brick) dwellings, and high household formation rates have made this problem even more acute. It is estimated that the number of informal dwellings in Cape Town grew from 28 000 in 1993 to roughly 100 000 in 2005 under the pressure of migration and urban population growth (Rodriques et al., 2006).

2.2.1 *Housing Policy*

The first democratically elected government embarked on a number of policies to improve the lives of South Africans.[8] The new South African constitution included the right to adequate housing (see Section 26, Constitution of South Africa).[9] The South African government promised to deliver 1 million houses in the 5 years between 1995 and 2000. This ambitious target was more or less met.

8 Important policies include electrification of 1.75 million home, improved access to running water to nearly 5 million people in rural areas, the extension of free basic health care to 5 million people, and a childcare grant and pension program. Some evidence on the positive impacts of these policies are documented in Duflo (2003); Case and Deaton (1998); Dinkelman (2011).

9 The South African constitutional court has consistently upheld individual rights to housing, and these constitutional changes have insured rights to land for millions of households that were previously categorized as "squatters". For a description of the watershed legal case involving land rights see Sachs (2003) on the Grootboom case.

By 2008 it was estimated that 2.3 million houses had been built, and in May 2013 the government announced that it had passed the 3 million mark (South Africa, 2013).[10]

This housing policy, originally referred to as the Reconstruction and Development Program (RDP) aimed to provide as many low cost houses as quickly as possible.[11] The value of the house provided is very small in comparison to similar projects in countries such as Chile (Gilbert, 2004).

The program gives individual capital subsidies to eligible households. However, the vast majority of the subsidies have been product linked; they had to be used to purchase houses commissioned by the government and built by private construction companies. These came to be known as the RDP houses, small stand alone units built on large empty land just outside existing informal settlements. Government housing policy was updated with the "Breaking New Ground" policy document of 2004, which placed increased emphasis on minimum building standards, *in situ* approaches to upgrading, rental housing and densification (Charlton and Kihato, 2006). But by and large, the housing scheme has continued to be characterized by the construction of large greenfields projects. Small houses on separate plots with the orange roofs that have come to characterize the South African urban landscape. For a detailed and up-to-date outline of issues relating to the housing policy, see Tissington (2011). There is no evidence to my knowledge that government housing projects have been rolled out in conjunction with other welfare or urban improvements projects.

To be eligible to receive housing an individual applicant needs to be married or otherwise supporting dependents in a household with total income of less than R3500 per month, cannot own a registered property, and must be a South

10 According to the census of 2011 there are approximately 14.5 million households in total in South Africa.

11 This policy of the national housing subsidy scheme was outlined in "A New Housing Policy and Strategy for South Africa" (Republic of South Africa, 1994).

African citizen (Department of Human Settlements, 2009).[12] I discuss issues to do with allocation of housing to applicants when I discuss the identification strategy used in this paper in Section 2.4

The success in the delivery of housing units has been a cornerstone of the African National Congress's electoral campaigns since 1994. More than 10 million people are estimated to have benefited directly from the program. Yet in a period of increasing poverty, unemployment and urbanization, the number of households living in informal housing has actually increased, especially among the African population.

The policy has been criticised for not doing enough to deal with the housing backlog, providing low quality substandard housing that hardly improves living conditions of the poor (Tomlinson, 1998; Lipman, 1998), not being accompanied with other infrastructure and neighbourhood investments (Huchzermeyer, 2003) and for not doing enough to deal informality by ensuring transfer of title deeds (Huchzermeyer and Karam, 2006). Housing programs have been said to have contributed to forced evictions (Chance, 2008), particularly to areas further away (Centre on Housing Rights and Evictions, 2009) and to have been biased towards particular racial or political groups (Seekings et al., 2010). The most pervasive criticism of the policy has been the location of housing which has often been determined by private construction companies that choose to build on the cheapest possible land. In most cities housing has been built far away from the city centres in a way that has reinforced the spatial segregation of South African cities (Huchzermeyer, 2006; Bundy, 2014; Charlton and Kihato, 2006).

In the setting where this study is conducted, the Western Cape Occupancy Study (Vorster and Tolken, 2008) finds that resale rates of housing are high, at around 20%, mostly on an informal market. Rental of these houses does not appear to be at all common, while the practice of building a small "backyard"

12 It has been frequently observed that these eligibility requirements were often unverified, and for the vast majority of slum dwellers, are not likely to be binding anyway.

shack or shelter is. More than one third of households had a backyard structure within a few years of receiving the house.[13]

While many studies have evaluated government housing using observational data or qualitative analysis, there is no study, to my knowledge, that attempts to estimate causal impacts of government houses on the outcomes of households receiving them.

2.2.2 *Theoretical Framework*

In this section, I provide a theoretical basis for causal links between housing and household labour participation. I provide evidence on the patterns of time use for female household members in South Africa and argue that slum dwellers' time is constrained by their physical living environment. A simple model of home production predicts that upgrading housing would induce substitution of time away from home production into wage labour.

I conceive of home production as time consuming activities related to the production of goods and services consumed at home. This includes maintaining the physical structure to ensure safety, security, warmth and shelter, household activities such as cooking and cleaning, and rebuilding of structures after damage from fires or flooding.

The UN-Habitat report of 2003 outlines a full taxonomy of the basic characteristics of slums. Many of the deprivations of slum living relate to issues of home production. Cooking and bathing is likely to be considerably easier in a home with running water and electricity, as opposed to a informal dwelling where other carbon fuel sources are often collected, and water has to be fetched from communal taps. Maintaining a sanitary home environment is also likely

13 These structure were sometimes used to accommodate other members of the household that could not fit in the original structure. In the cases when the structures were occupied by non-household members, only about half of paid rent. Some households still owned their previous (informal) dwelling, and were renting it out, but in most cases their informal dwellings had been demolished when they left, or they had given it to a friend or family member.

to be far easier in a cement floored home without leaking roofs or permeable walls.[14] In Cape Town most slum dwellers have to use badly maintained communal toilets located some distance from their homes, or buckets which must be emptied outside of the home every morning. Paraffin is a common use of fuel for cooking and heating, and is known to be a cause of fires and respiratory disease (Schwebel et al., 2009). In Cape Town, as in slums around the world, formal electricity connections are rare for shack dwellers, with more than 50% having fire-prone illegal connections, or no electricity at all (City of Cape Town, 2005). Electricity greatly aides home production if it facilitates the use of fridges, stoves and microwaves.

Time use surveys of poor South African indicate that a considerable amount of time is consumed by domestic activities, particularly for female members of households, who are primarily responsible for chores at home. South African women spend on average three times as long (3.5 hours a day) as men on unpaid work (Budlender et al., 2001).[15] Crucially, the evidence suggests that these activities take far longer in informal housing. In the national accounts individuals living in informal housing in urban areas spent 25% more time on non-labour market work than other urban households (Budlender et al., 2001). Shack dwellers in Cape Town report *more than twice* as much time (17.1 hours per week) spent on housework than their formally housed counterparts (7.5 hours per week).[16]

Many of the issues related to time use are likely to do with access to labour saving appliances. In 2000 only 28% of informal dwellings in urban areas used electricity for cooking, versus 77% of urban households. Similarly they were far

14 Cattaneo et al. (2009) have looked at how cement floors improve health from improved sanitary conditions.

15 These patterns are consistent estimates for other developing countries (Berniell and Sánchez-Páramo, 2012).

16 These were calculations based on the CAPS datasets used for this paper, outlined in Section 2.3. Unfortunately this data was no collected for periods of the survey beyond the first wave, which makes it impossible to estimate the impact of housing on time use in this setting.

more likely to use gas or paraffin stoves for cooking and heating and lighting. Only 46% of shack dwellers have access to a refrigerator, compared to 90% of families in brick houses.

Households living in the slums on the Cape flats are extremely vulnerable to township fires and, during winter months, storms and flooding.[17] Fire hazards are due, in part, to the types of appliances used for cooking and heating outlined above. These events are common, and often lead to widespread destruction of housing infrastructure, which takes time and money to rebuild.

Pharoah (2012) provides an overview of some of the risks facing informal dwellers in Cape Town. The greatest impacts come from the health problems and losses of days worked and at school because of the disruption caused by fires.[18] In that study 83% of shack dwellers had experienced some kind of flooding while living in Cape Town.

In addition, slum dwellers have considerably less security, since their homes can easily be broken into. This could impose limits on tenants' ability to commute into the city to look for work for fear of theft.[19]

2.2.3 *A Model of Home Production, Work and Leisure*

All told, the time burdens of living in informal dwellings are considerable. In the empirical analysis I seek to evaluate the total effect of receiving housing, which affects the way in which home production happens through many channels. In what follows, I present a simple model of how changes in housing quality could

17 In 2005 a particularly damaging fire razed over 3000 shacks in Joe Slovo informal settlement just outside of Cape Town ("Shack-dwellers have nothing left after blaze" (iolnews, January 17 2005)) The victims of the fire were promised government housing after being displaced, but many remain in temporary relocation camps years later (Centre on Housing Rights and Evictions, 2009).

18 Some 20% of live in high flood risk areas, and roughly 40 000 people were directly affected by townships fires in Cape Town between 1995 and 2004.

19 This threat of invasion seems more urgent than that of expropriation risk (Field, 2007). While security of tenure is a great issue for informal dwellers in South Africa (Royston, 2002) this risk is related more to formal eviction to make way for new housing or urban development projects, rather than contestation of property right by other private agents.

influence home production, which in turn predicts increases in labour hours due to the effect of formalized housing. I do not distinguish between the different channels through which housing could improve home production, which were outlined in the previous section.

The model I use is of the lineage of Becker (1965), since it specifies utility as a function of an unobserved home production input $H(T_h, b)$, which is produced through time at home T_h in combination with the physical housing infrastructure b. Importantly, I assume that home produced goods and services are *not* perfect substitutes with other forms of consumption, as opposed to many of the other models in this literature (Gronau, 1977, for example). This fits with the way I have conceived of home production in informal settings, where the basic needs provided for by housing cannot be taken for granted.

In this model, production at home cannot be traded on the market, it is used within the household. Household utility is a function of home production, consumption and leisure $U(H, C, L)$. Consumption is given by time spent working for wage labour T_w times the prevailing wage w. Leisure, time on home production, and time at work sum to one. With prices normalized to one, household utility is given by:

$$U(H(T_h, b), wT_w, 1 - T_h - T_w,)$$

If the household maximizes utility with respect to its allocation of time between labour, leisure and work at home, the first order conditions are simple:

$$\begin{aligned} U_H \cdot H_{T_h} &= U_L \\ U_C \cdot w &= U_L \end{aligned}$$

The optimizing household would thus choose its optimal time on work at home and labour, given by $T_h^\star = T_h^\star(b, w)$ and $T_w^\star = T_w^\star(b, w)$, respectively. I want to find $\frac{dT_w^\star}{db}$: the impact of upgrading the physical housing infrastructure on wage labour supplied. While one could speculate intuitively about the direction of impact from the FOC's, total differentiation with respect to b, gives a more complete picture, in the general case. With some manipulation this eventually yields:

$$\frac{dT_w^\star}{db}\left[U_{LL} - \frac{wU_{CC} + U_{LL}}{U_{LL}}\left(U_{LL} + U_{HH}H_T + H_{TT}U_H\right)\right] = -[U_{HH}H_bH_T + H_{Tb}U_H]$$

$$\frac{dT_w^\star}{db} = \frac{[U_{HH}H_bH_t + H_{Tb}U_H]}{wU_{CC} + \left(\frac{U_{CC}w + U_{LL}}{U_{LL}}\right)\left(U_{HH}H_T + H_{TT}U_H\right)}$$

Assuming diminishing marginal utility for all inputs into the utility function, and a diminishing marginal product of time at home, renders the denominator unambiguously negative. Turning to the numerator, the first term is clearly negative due to the diminishing marginal returns on home production and the positive returns to housing quality from time spent at home. The sign of the second term hinges on whether or not the marginal utility of time in the home increases or decreases with an improvement in housing quality. If $H_{Tb} = \frac{\partial^2 H}{\partial T \partial b} \leqslant 0$ the numerator would be negative, and the response of hours in the labour market would be unambiguously positive.

I do not take a definite position on the sign of the H_{Tb} or even $\frac{dT_w^\star}{db}$ at this stage. This paper tests empirically the relationship between home production and labour supply. However, I would argue that under my definition of home production, H_{Tb} is likely to be negative in this context. While both time at home and housing quality influence home production, they are close substitutes. Take for instance the basic issue of shelter. Providing a better roof and walls would reduce the value of work done on maintaining the home, because nothing really needs to be done to make the structure more secure anymore. Before, time spent fixing holes in the roof greatly improved the quality of housing. In the case of perfect substitutes, $H = h(T_h + b)$, $H_{Tb} = h''(.) < 0$. Of course, in the additively separable case, $H_{Tb} = 0$ and the result holds.

Alternative channels

I cannot rule rule out other channels that could lead to changes in household labour supply. Health could be leading to a positive impact on labour supply through the productivity of household members. There is a large literature looking at the links between health and productivity (Strauss, 1986; Strauss and Thomas, 1998). The links between housing and health are also firmly established (Pitt et al., 2006; Cattaneo et al., 2009). This channel is relevant in this setting, but I am unable to estimate the impact of housing on health using the data available. Given that the impacts of improved health are likely to accrue more to female members of households who spend the most time using the stoves and appliances that are most detrimental to health, this effect might be considered part of the full effect of informal housing on female capabilities.

In addition there could be additional effects of receiving government housing such as changes in household composition, new rental income, and household location. New household members might move into the additional space that a larger house and plot affords. These new arrivals could bring with them sources of income if they are employed, or government grants. Alternatively they could be alleviating the burden of work in the home, allowing other members of the household to seek employment. Recipients of housing could see large increases in incomes due to rental incomes- either from the shacks they have moved out of, or backyard structures constructed on their properties, which is common practice.[20] In my empirical analysis I will show that the results are not influenced by

20 Rent and one off wealth increases from the sale of housing should not be included in measures of include, but it is possible that these were incorrectly reported by households.

including controls for changes in household size and composition, nor are they effected by looking at per capita measures of household income and earnings. I argue that the labour supply channel fully explains the impacts on household income.

2.3 DATA

My empirical analysis uses the CAPS panel survey of Cape Town metropolitan area, with four waves collected in 2002, 2005, 2006, and 2009.[21] The sample was randomly selected using probability proportional to size sampling and stratification by racial group using 1996 local area census data.[22]

Of the households surveyed in the first wave, roughly one third were living in informal dwellings. For the purposes of evaluating the impact of receiving a housing subsidy, it was necessary to drop all those who weren't eligible, and therefore not a valid control group. As a result I have dropped all households that were not living in an informal dwelling (or "shack" as coded in my data). I also drop the few households who report that they have already received government housing before wave one. Households who have received government housing are usually not still living in shacks, but some have moved out or lost their houses. These individuals are no longer eligible for housing and

21 The Cape Area Panel Study Waves 1-2-3 were collected between 2002 and 2005 by the University of Cape Town and the University of Michigan, with funding provided by the US National Institute for Child Health and Human Development and the Andrew W. Mellon Foundation. Wave 4 was collected in 2006 by the University of Cape Town, University of Michigan and Princeton University. Major funding for Wave 4 was provided by the National Institute on Aging through a grant to Princeton University, in addition to funding provided by NICHD through the University of Michigan (Lam et al., 2006). Further information can be found on the CAPS website at http://www.caps.uct.ac.za

22 This survey was conducted with the primary motivation of tracking young adult's behaviour, sexual attitudes, labour force participation and health. However household questionnaires were also conducted, with an extensive household roster questionnaire which surveyed the entire household. However, this does impose some limitations on the analysis that can be conducted.

thus should not be in the sample. This leaves me with a sample of 1350 eligible households. 1097 of these are found at least once in subsequent waves.[23]

Table 2.1 provides an overview of my sample of houses that were living in shacks in 2002. Over a period of just 7 years nearly 40% of the sample has received a government house. I show the proportion of households that received housing for each wave of data ("Treated Here") as well as cumulative proportion that have received housing to that point. It is the latter outcome that will be used as the dependent variable in the analysis because I expect the impact of housing to be present in all periods after which it is received. More households are treated between the first and second periods (19.7%) than any other. The high proportion of households receiving housing in this data is testament to the scale of the roll out of the housing program in Cape Town.

However, it is also striking how rapidly all households improved the quality of their living and housing conditions. Some of this effect is undoubtedly due to government housing, but households that did not receive government housing managed to either improve their housing conditions (move out of shacks) or gain access to important infrastructure and amenities. This could be due to both improved government service delivery during this time, and a natural process whereby new migrants to cities manage to improve their living conditions over time, consistent with a "modernization theory of slums" (Marx et al., 2013). Indeed many of the households in informal housing in the first wave of the panel were recent migrants to the city.

The scale of rollout of housing, the effects of which are clear in my sample, provides a perfect setting in which to evaluate the effects of government housing on labour outcomes. My data includes information on housing conditions in each wave of the survey, and detailed information on labour market decisions of

23 Although in each specific follow up wave about 80% of households are reached on average, conditional on having been found at least once in the follow up waves.

Table 2.1: Evolution of sample household characteristics

Wave	1	2	3	4
Year	2002	2005	2006	2009
Treated	0.0%	19.7%	25.9%	38.6%
Treated Here	0.0%	19.7%	8.87%	12.09%
Shack	100.0%	65.6%	62.4%	45.6%
Flush Toilet	70.8%	79.2%	85.6%	90.7%
Piped Water	12.3%	25.3%	28.4%	42.0%
% Female	54.0%	55.1%	54.0%	53.3%
Dist To City (km)	23.51	23.69	23.65	23.55
Head of Household Background				
Coloured	15.0%			
African	83.2%			
Moved to Cape Since 1985	56.2%			
Born Cape Town	19.0%			
Born Eastern Cape	75.1%			
Lived in backyard dwelling	10.6%			

one young adult member of the household. Other labour data comes from the household rosters.

In Table B.2 in the Appendix, I compare the sample mean for households that received housing to those that did not at both baseline and and endline (at baseline I look at househld that are going to receive housing). There are clear differences in observables between treated and control individuals. This differences are consistent with a story of housing allocation whereby poorer households were more likely to get housing, as discussed in Section 2.4. Backyarders (those living not in large informal settlements but in shacks in the yards of a more formal dwellings) seem far less likely to get housing, as are coloured households. Migrant status does not seem to make a significant difference. Importantly, we observe that households that were treated lived far further away from the city center, which is due to to fact that projects were built further away from the city, where there was more cheap available land. These differences in the characteristics of the population targeted by housing motivate many of the robustness checks discussed in Section 2.7.

2.3.1 *Housing Project Data*

During fieldwork conducted during 2011, I gathered datasets on the rollout of government housing from the Provincial Department of Human Settlements and Local Government Planning departments in Cape Town. I built a comprehensive and accurate dataset of RDP housing roll out in Cape Town over last 15 years. I used three main sources to generate this data. The first was a database of projects that originally came from the National Housing administrative records, with geographical coordinates of the projects, along with approval status and date of approval. However I found that a great deal of the coordinates were highly inaccurate.[24] The data lists dates for when programs were proposed or approved, rather than when they were actually completed. I used project ID numbers to match this data with a second database of projects which listed more accurate dates of housing roll out, as well as a detailed breakdown of housing subsidy numbers by building date, but that lacked any geographical information.

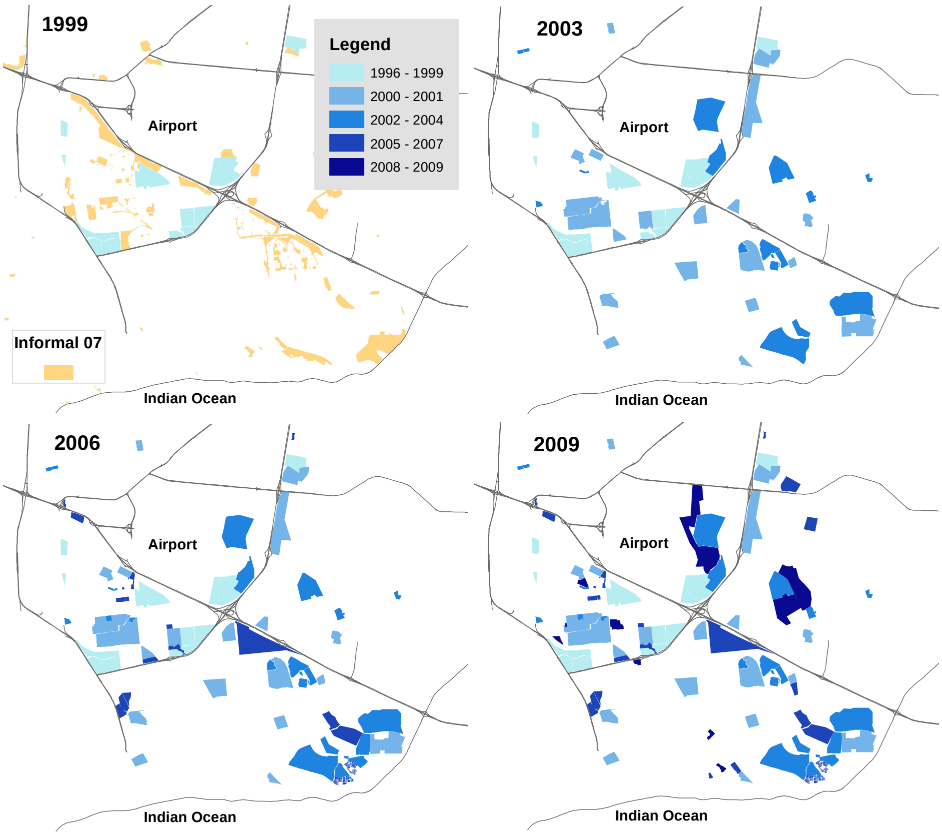
Finally I combined this data with an invaluable geographical ArcGIS map acquired from the Cape Town City Housing Department, which provided polygons outlining the location of housing projects.[25] By linking the three datasets together I was able to generate a georeferenced panel of the number of households built per project in each year.

This data is presented in Figure 2.1 showing the expansion of housing projects over the years from 1999 to 2009 for areas in Cape Town where housing was built. Figure B.2 in the Appendix shows a broader overview of housing projects for the whole City in 2009.

24 There were housing projects placed in the ocean or on the mountain.

25 This dataset was built by Rehana Moorad at the local government department with great accuracy. In some cases planning department construction blueprints had been used to individually identify housing units in great detail.

Figure 2.1: Housing roll out in Cape Town



I aggregated this yearly data into blocks of years corresponding to the time between waves of the CAPS data to get a measure of how many houses were built, at each location, between each wave of survey data.

2.3.2 *Location and Proximity Measures*

I used confidential datasets in order to track households as they moved.[26] I used original enumeration areas maps to locate the original living location of households in the first wave of the sample, then used household addresses from survey tracking sheets to update household locations as they moved.

I used ArcGIS maps of the original EAs sampled to map the approximate locations of the households at the start of the survey. I then used household's

26 These were provided with the help of Jeremy Seekings of the Centre for Social Science Research, University of Cape Town, and David Lam of Population Studies Center, University of Michigan, after discussions in January 2011.

addresses in later waves, transcribed from the survey documents, to identify households that had changed address. I then geocoded the new addresses. In this way I tracked households throughout the four waves by their GPS coordinates.[27] I was then able to generate a range of distance and geographic outcomes for each household. In each wave of data I calculated the distance from schools, roads, the city centre, and the distance of move from the original place of living at the basline (if there was any move at all).

Summary statistics of the migration data are presented in table 2.2, along with the housing distance data described in the next section. Roughly 30% of the sample moved at some point during the survey.[28] The average move distance is small, under 1 km.[29] This data gives an idea of how far informal households live from the city centre- 26kms on average.

Table 2.2: Proximity data in wave 3 (2006)

	Mean	Min	Max	N	Control	Treat	Diff
City Distance	25.8	4.05	53.3	968	24.7	27.8	3.07***
School Distance	0.48	0.019	3.43	970	0.50	0.45	-0.041
Moved	0.36	0	1	970	0.34	0.40	0.060
Move distance	0.94	0	36.8	970	0.95	0.92	-0.025
Cumulative dist moved	1.53	0	36.8	968	1.37	1.83	0.46
Distance Proj1	0.88	0	16.7	970	1.13	0.41	-0.73***
Distance Proj2	2.41	0	28.6	970	2.82	1.69	-1.13***
Distance Proj3	3.24	0.046	31.2	970	3.60	2.57	-1.04***
Rank Proj1	0.39	0	1	970	0.35	0.48	0.13***
Rank Proj 2	0.11	0	1	970	0.089	0.16	0.068***

27 I used Google maps for this. Their batch geoprocessing tools could not always be used because of the considerable variation in spellings of streets and areas name, especially when in different languages, or in newly developed areas were street names had not been formalized. Most of these GPS coordinates had to be found by hand.

28 Most of these moves were within the boundaries of the City of Cape Town, but there were a few households that moved back to rural areas in the Eastern Cape or KwaZulu Nata, some hundreds of kilometers away. For the purposes of urban relocation analysis, such outliers were excluded from the sample.

29 This may be an underestimate because households that moved further were less likely to be found, and there were sometimes mistakes with updating address data during the fieldwork.

Housing Project Distances

Most importantly I was able to generate distances for each household, in each wave, to all of the government housing projects on which houses had been constructed during the years since the last survey. For the reasons that become clear in the next section, I focus only on the distance between housing projects and enumeration areas (EAs) that the household was living in the first wave.[30] I created dummy variables for EAs that were contained within housing projects, as they were most likely to be upgraded.

In addition each EA was given a rank (among all other EAs) to each project nearby, such that each household-project distance pair had a corresponding rank assigned to it. A housing ranking might not necessarily correspond closely with its distance to a project, if is located in a densely populated area where many households are competing for treatment.

Table 2.2 shows these measures, the average distance from the closest housing project, then the second and third closest. I also show a dummy variable for whether the household lived in the top 3 closest EAs to the housing project nearest to that household.

2.4 EMPIRICAL STRATEGY

The paper uses three key strategies for identifying the causal effect of government housing on household outcomes. Firstly I look at OLS regressions with household fixed effects to estimate the effect of receiving housing. This gives a basic estimate for the difference-in-difference treatment effect of housing. Secondly, using a natural experiment that I will explain in detail in this Section, I instrument for individual selection into treatment (receiving a house) by us-

30 Importantly, EA centroid locations, instead of their boundaries, were used to calculate distance. This only makes a noticeable difference for those EAs right next to, or inside projects. I wanted to distinguish EAs that were completely surrounded by projects from those that simply had part of their boundary overlapping with the boundary of a housing project.

ing proximity to government housing projects. Thirdly, I use a set of housing projects that were planned but not built in order to control for selection at the geographic level, by dropping from the control group those areas that never had projects planned nearby.

Before turning to the formal identifying specifications and assumptions, I describe the natural experiment that I exploit as part of my identification strategy. In what follows, I use 'treated' or 'treatment households' to refer to households that received government housing as a result of the policy.

2.4.1 *Natural Experiment: allocation by proximity*

This paper uses the government's proximity-based allocation policy as a natural experiment. I focus on the procedures used by the local government in Cape Town. Observations on the workings of allocation procedures came from numerous meetings and discussions with officials in the local government in early 2011.[31] Additional policy documents, reports and research papers on the methods of allocation corroborate this story (Tshangana, 2009; Seekings et al., 2010; Tissington et al., 2013).

While the official eligibility rules for housing stipulate that households must earn less than R3500 per month to be eligible for housing, this cut off seems not to be enforced in practice.[32] Once a household has (rightly or wrongly) been deemed eligible, it joins a national housing waiting list. This list is supposed to work like first-come-first-served queue, but in reality housing construction at the local level determines the order of delivery, and even within communities there is evidence that households often jump the queue (Tissington et al., 2013).

31 I refer to discussions I had with Paul Whelan (Western Cape Provincial Department of Housing), and Heinrich Lotze (Head Housing Development Co-ordinator, City of Cape Town Government).

32 Indeed I looked for a discontinuity in the probability of receiving housing at the cut-off in baseline income. While the probability of receiving housing was definitely lower for very wealthier households, the discontinuity at the eligibility cut-off was almost non-existent and statistically insignificant. In addition, there are relatively few individuals (only 12%) in my sample of slum-dwellers who fall above that cut off: the eligibility constraint did not find for them.

As a result of the project-by-project nature of the roll out, households were selected into projects according to catchment areas around the projects. From these areas a number of "source areas" (particular informal settlements, or communities within settlements) were selected. These stakeholders were allocated a certain quota of housing units from the project (Tshangana, 2009).[33] That group of communities would establish project committees responsible for allocating housing to their members, with the one restriction (not always enforced) that all selected candidates much be on the housing waiting lists.[34]

In this way, households that were living close to housing projects that were built between 2002 and 2009 were more likely to be treated than those living further away. It is this relationship that I exploit as an identification strategy. Of course location of housing projects itself was not exogenous, making it crucial to understand how housing site locations were selected. The location of the housing projects was not generally determined by members of communities. In most cases it was not determined by the government either. The role of private developers in the housing process meant that land availability and affordability were the main forces determining construction locations. This meant that housing projects were generally developed in areas where land was relatively abundant or cheap, or in parcels of undeveloped within the city.

In this way I argue that geographic proximity to new projects was uncorrelated with *changes* in household outcomes, except through the channel of improved housing. I will return to this argument shortly. Firstly, I use a set of the set of

33 In some cases a certain number of units would be reserved for households outside of the catchment area, usually communities that had been waiting for a houses for a particularly long time, or had been recently relocated. An example is the Joe Slovo informal settlement near Langa, which was allocated housing in the N2 Gateway Project due to a fire that affected that community.

34 Street committees are a common characteristic of most townships in Cape Town and are often those involved in the management of the communities housing quota allocations. Committee representatives that I met in Cape Town had a list of their community members who were eligible for housing, which they used to make allocations.

distance-from-project measures to predict selection into treatment, as the first stage of an instrumental variables estimator, discussed in more detail below.

2.4.2 *Identification*

The basic OLS regression of household outcome y_{it} on having government housing T_{it}, including controls for household observables X_{it} is given by:

$$y_{it} = \alpha_0 + \alpha_i + \lambda_t + X_{it}\beta + T_{it}\tau + \delta_{it} + \epsilon_{it} \tag{2.1}$$

This estimator likely to be biased due to correlation between household unobservables α_i and the housing treatment. In order to account for household unobservables that might be driving selection into housing, as well as outcomes of interest, I estimate a fixed effects model which estimates the difference-in-difference impact of receiving government housing:

$$\begin{aligned} y_{it} - \bar{y_i} &= \lambda_t - \bar{\lambda} + (X_{it} - \bar{X_i})\beta + (T_{it} - \bar{T_i})\tau + (\epsilon_{it} - \bar{\epsilon_i}) \\ \widetilde{y_{it}} &= \widetilde{\lambda_t} + \widetilde{X_{it}}\beta + \widetilde{T_{it}}\tau + \widetilde{\epsilon_{it}} \end{aligned} \tag{2.2}$$

where $\widetilde{y_{it}}$ represents the demeaned version of the outcome of interest. The fixed effects estimates correctly identify the effect of housing under the assumption of common trends. That is, households that were treated would have had the same changes in y over time had they not been given the housing, that is: $E(\delta_{it}|T_{it} = 1) = E(\delta_{it}|T_{it} = 0)$.[35] This requires of course that treated households were not effected by different trends or shocks unrelated to housing over time.

35 For a more detailed discussion of the problem of unobserved time trends in panels, and the resulting bias of difference-in-difference estimators see (Bertrand et al., 2004).

Sources of bias

There are a number of reasons to doubt the assumption that treatment is uncorrelated with individual time shocks. There have been widespread reports of manipulation of the housing allocation lists, with certain individuals receiving preferential treatment based on political connections or other means to access housing (even paying bribes) (Seekings et al., 2010; Tissington et al., 2013). If households who received windfalls or good new jobs were able to leverage their increased incomes to access housing, this could bias the estimates upwards.[36]

On the other hand, it may be the case that housing allocation is more pro-poor such that housing is allocated by local politicians and communities to households that have the least ability to improve their own circumstances. Alternatively, households that suffer negative income shocks might be more likely to be awarded housing. This would bias the estimates of the impact of housing downwards, as households are less likely to experience increases in their incomes are most likely to get housing. In the data used in this paper I find that housing is more likely to go to households that were poorer at baseline.

In addition, the long waiting lists for houses could cause downward selection bias. Many of households that get treated are likely to be the ones who have remained in informal dwellings the longest, making them high up the community waiting lists. Those who were able to get out of poverty and upgrade dwellings on their own are, by definition, off the waiting lists (or at leats out of my sample of eligible individuals). Thus households would be selected into treatment due to their relative inability to improve their housing on their own. Finally, measurement error could be a source of downward bias: the extent of measure-

36 Given the roll-out of numerous government programs at the same time as the housing project, it is possible that households that managed to get government housing, also received other benefits simultaneously, which might improve their economic outcomes

ment error in the sample could be substantial, especially in the measurement of incomes, and even in the treatment variable.[37]

Instrumental variables estimator

I deal with non-random selection into housing at the individual level through use of an instrumental variables (IV) estimator. The natural experiment outlined in the previous section allows me to use distance from housing projects as instrument for selection into housing projects. In this way I follow McKenzie and Seynabou Sakho (2010), Attanasio and Vera-Hernandez (2004) and Ravallion and Wodon (2000), who use distance from tax registration offices, community centres and schooling project, respectively, to control for selection into social programs.[38]

Call the relevant distance instrument Z_{it} to estimate a fixed effects-two stage least squares (FE-2SLS) estimator, given by

$$\widetilde{y_{it}} = \widetilde{\lambda_t} + \widetilde{X_{it}}\beta + \widetilde{T_{it}}\tau + \widetilde{\epsilon_{it}} \tag{2.3}$$

$$\widetilde{T_{it}} = \widetilde{\lambda_t} + \widetilde{X_{it}}\pi_1 + \widetilde{Z_{it}}\pi_2 + \widetilde{\epsilon_{it}} \tag{2.4}$$

where Equation 2.4 gives the first stage prediction of the probability of switching to be treated (receiving a house) from non-treated in time period t. The fitted values for $\widetilde{T_{it}}$ are then used as regressors in Equation 2.4.2. The identifying assumption (exclusion restriction) of this model is that distance from housing projects is uncorrelated with the change in the outcome of interest: $\widetilde{Z_{it}} \perp \widetilde{\delta_{it}} + \widetilde{\epsilon_{it}}$. I turn to discuss this assumption in Section 2.4.4.

37 Sometimes the interviewed household member might not be able to remember if the household had received the house from the government. Alternatively households might have moved out of housing after selling it or renting it out, such that they would mistakenly report not having received government housing.

38 This fits with a larger literature of using geographic instruments. Dinkelman (2011) and Klonner and Nolen (2010) use terrain data to instrument for the placement of electrification programs and mobile phone antennas, respectively. These papers follow a methodology pioneered in Duflo and Pande (2007) to evaluate the growth impact of dams. Similarly Banerjee et al. (2012) uses distances from major roads built across China to evaluate the impact of these roads on local growth.

In this framework, fixed effects estimation addresses the problem of endogenous time invariant household unobservables, while endogenous time varying "shocks" to the household are dealt with through the instrumentation.[39]

First stage

It is the distance from multiple housing projects that matters for the probability of receiving government housing. This presents an econometric challenge since the distance from a single (closest) housing project is not particularly informative about the probability of treatment. It is the cumulative effect of numerous housing projects, including the number of houses built in that project, over the years that predicts selection. After all, if a household was not given a house by the closest project, it may stand a good chance of winning housing in the next closest project, especially if it was moved up the waiting list after neighboring households got houses. Furthermore the number of other households in the neighbourhood of a project will also influence the probability of receiving housing for a fixed supply of new housing.

In Section B.1.1 in the Appendix I discuss in more detail some of the challenges arising from this issue, including a discussion of why alternative measures summarizing the total distance of households from multiple projects are problematic in terms of the parametric assumptions that they place on the relationship between distance and selection. In the robustness checks, Section 2.5.3 I look at the results IV estimates where I simply use a full set of distance measures linear predictors of treatment in the first stage, and show that the results are consistent with the rest of the results in the paper, but are estimated imprecisely and with a severe problem of too many weak instruments.

39 Murtazashvili and Wooldridge (2008) present a more thorough discussion of what I have presented here. They investigate a more general version of the model I have introduced, using time varying and permanent individual slopes, and show the conditions required for this model to give consistent estimates of the 2nd stage parameters.

Instead, I need a flexible estimator to predict selection into treatment that involves multiple nearby housing projects. I follow Wooldridge (2002) by estimating the probability of treatment by a non-parametric function $G(x, z; \rho) = P(T = 1|x, z)$, which uses multiple instruments z and a common coefficient ρ determining the impact of distance on the probability of treatment. Importantly, the fitted probabilities of the probability of treatment $\widehat{G}$ cannot be used as regressors in Equation in the usual 2SLS estimator. These are unlikely to uncorrelated with the error term as they are in the linear case.[40] Such an estimator will not be consistent. In addition, inference with this method will produce incorrect standard errors because of the non-linear form of the regressors and error correction methods would need to be applied.

I use an IV estimator adapted from Wooldridge (2002) and applied to the fixed effects case. Firstly I generate fitted probabilities of treatment $\widehat{G_{it}}$ for each individual in each period using a non-linear specification based on a full set of proximity instruments. I then use those predicted values as an instrument for treatment status T_{it} in the FE-SLS given by Equation 2.4. In other words I use a linear projection of T_{it} onto $[x, G(x, z; \hat{\rho})]$ as the first stage of a 2SLS procedure. Wooldridge (2002) refers to this as using generated instruments as opposed to generated regressors. This linear projection will not be correlated with the error term under a valid exclusion restriction. This follows intuitively from the logic of 2SLS; if the instruments Z are informative and valid, then $G(x, z; \hat{\rho})$ will be too.

Wooldridge (2002) shows that in the IV framework, we can ignore the method of estimation of ρ in the first stage. Inference in the 2SLS with $\widehat{G_{it}}$ as instruments is consistent, and no standard error corrections are required. But this non-linear

40 The only condition under which such a method would yield an efficient estimator is if data generating process is perfectly specified by G. We can never really know this and is far too strong an assumption in almost any case (Angrist and Pischke, 2008).

form is more efficient than the linear 2SLS, and thus more likely to provide valid inference (Newey, 1990).

First Stage Specification

In this section I define the function that determines selection into housing $G(x, z; \rho)$ and the estimation of ρ (the set of coefficients that capture the effect of distance on receiving housing). I use maximum likelihood methods to estimate a unique binary outcomes estimator which assume a latent variable structure for the impact of each distance instrument on the probability of treatment.

Imagine a household surrounded by a number of housing projets: the aim here is to predict treatment as a joint function of distance from all of the nearby housing projects as efficiently as possible. Firstly I use a binary outcomes model to for an expression for the probability of household i being selected by a particular project a, for each project-household pair. This is not the same as actually getting housing from that project, since a household cannot receiving housing twice. I then combine the probability of being selected by each project into an expression for the joint probability of a household being selected by any housing project. I do not observe T_{ia}: that is which households received housing from which projects. I observe only T_i, the combined effect of being selected by any project.

Here I use the set of instruments $\mathtt{dis}_{ia}$- the distance between the household and a project built since the last survey wave.[41] The probability of household being *selected* housing from a specific project is given by:

$$y^\star_{ia} = x_i\beta + \mathtt{dis}_{ia}\rho \tag{2.5}$$

$$T_{ia} = \mathbf{1}(y^\star > 0)$$

$$T_{ia} = \mathbf{1}(x_i\beta + \mathtt{dis}_{ia}\rho + \nu_i > 0) \tag{2.6}$$

I assume that the error term ν takes on the logistic distribution $F(y^\star) = \Lambda(y^\star) = \frac{\exp(y^\star)}{1+\exp(y^\star)}$ such that

$$P(T_{ia} = 1) = \Lambda(x_i\beta + \mathtt{dis}_{ia}\rho) \tag{2.7}$$

Imagine the case where there are only two projects, and note that a household can receive housing from only one project. The likelihood of a household being treated by *either* project is:

$$P(T_i = 1) = P(T_{i1} = 1) + P(T_{i2} = 1) - P(T_{i1} = 1)P(T_{i2} = 1)$$

where $P(T_{ia} = 1)$ is calculated by (2.7). Notice the adjustment for the fact that a household cannot be treated more than once. In this framework, the expression $P(T_{ia} = 1)$, given by (2.7) has to be interpreted as the project specific contribution to being treated, *not* the probability of being treated by that project. For

41 In the application to the real data we will use a more full set of instruments, all relating to the relationship between households and individual projects. These are excluded at this point, for ease of exposition.

many projects, the probability is most simply expressed as complement of the probability of being selected by none of the projects:

$$P(T_i = 1) = 1 - \prod_{a}^{A}(P(T_{ia} = 0)) \tag{2.8}$$

$$= 1 - \prod_{a}^{A} \Lambda(-x_i\beta - dis_{ia}\rho) \tag{2.9}$$

This expression, when estimated, gives a single solution to the coefficient ρ, a common effect of distance for all housing projects, no matter how many different housing projects are used in the estimation. I use this model to predict the probability of being treatment for a single period, based on the housing projects that were built in that period.

The problem is complicated further by the use panel data: we'd like efficient estimates for the probability of receiving housing in each period. Households cannot receiving housing more than once, so the predicted probability of treatment should decline in a period after a household had a high predicted probability of treatment, all things equal. We want to derive an expression for the probability that a household has received housing at point in time up until the specific period. I development a functional form that conditions the probability of receiving housing in a particular time period on the probability of having received housing in previous periods.In the interests of space, this method relegated to the Appendix, Section B.1.2. There I also develop a multinominal estimator that predicts, the period in which a household will most likely be treated.

In addition, I present Monte Carlo simulations using simulations with calibrated parameters for ρ which gives the effect of proximity on the probability of receiving housing. I find that the estimator developed here does a good job of recovering the true parameter value for ρ, even in the presence of considerable noise and individual fixed effects. The average predicted probabilities of treatment from this model match the rates of treatment in the simulated data.

I am able to estimate the equation given by the 2.8 and the time (wave) specific probability of being treated using maximum likelihood techniques. It is to the results of these estimates that I now turn.

2.4.3 *First Stage Results*

I have outlined the key elements on the first stage of my instrumental variables strategy. Taken together I will estimate a system of equations taking the following form:

$$\widetilde{y_{it}} = \widetilde{\lambda_t} + \widetilde{X_{it}}\beta + \widetilde{T_{it}}\tau + \widetilde{\epsilon_{it}} \tag{2.10}$$

$$\widetilde{T_{it}} = \widetilde{\delta_t} + \widetilde{X_{it}}\delta_1 + \widetilde{\widehat{G_{it}}}\pi + \widetilde{\nu_{it}} \tag{2.11}$$

$$\widehat{G_{it}} = G(X_{it}, Z_{it}; \widehat{\rho}) \tag{2.12}$$

In this section, I show the results for the estimates of the selection equation (2.12). The model is estimated by maximum likelihood, where a single likelihood function describes the probability of getting housing in each period using all housing projects built during the time period. In this section show that this method of predicting selection into housing is highly informative and efficient. I also discuss other interesting predictors of selection into treatment to shed light on the way in which allocation to housing opportunities happens in practice.

Table 2.3 shows the estimates for the two different models outlined in the estimation section and obtained by maximum likelihood programming methods. I show two estimates: "L" denotes the use of the binary form estimator with a different likelihood function for each time period given by Equation (2.13). In this case the dependent variable is T_{it}- whether the household had received a government house by time period t. By contrast the multinominal "MNL" (described in detail in Equation (2.14) in the Appendix) denotes the model with dependent

Table 2.3: Maximum likelihood estimation of treatment status

	(1) L	(2) L	(3) L	(4) L	(5) MNL	(6) MNL	(7) MNL
Project Dist	-0.499***	-0.515***	-0.396***	-0.672***	-0.534***	-0.514***	-0.561***
	(0.100)	(0.105)	(0.106)	(0.142)	(0.137)	(0.135)	(0.164)
Rank	0.687***	0.735***	0.662***	0.350**	0.844***	0.897***	0.426*
	(0.126)	(0.127)	(0.130)	(0.143)	(0.203)	(0.202)	(0.245)
Project Dist sq	.0062***	.0065***	.0044***	.0079***	0.007***	0.007***	0.007***
	(.0015)	(.0015)	(.0017)	(.002)	(.245)	(.244)	(.288)
In situ	1.113***	1.149***	1.191***	1.802***	0.707***	0.761***	1.28***
	(0.176)	(0.177)	(0.186)	(.196)	(.0019)	(.0019)	(.0023)
Female Head		0.202**	0.166*	0.120		0.312**	0.269*
		(0.0895)	(0.0910)	(0.0957)		(0.136)	(0.142)
HH Size		0.0859***	0.0666***	0.0733***		0.0219	0.0178
		(0.0158)	(0.0173)	(0.0180)		(0.0272)	(0.0319)
Sex Ratio		-0.610***	-0.637***	-0.544**		-0.679**	-0.591**
		(0.197)	(0.201)	(0.212)		(0.283)	(0.295)
Age Ratio		0.153	0.205	0.0658		0.744**	0.667*
		(0.224)	(0.226)	(0.237)		(0.325)	(0.355)
City Distance			0.04***	0.05***			0.04***
			(0.00655)	(0.00699)			(0.0111)
Max Education			0.0267**	0.0319**			0.0132
			(0.0126)	(0.0134)			(0.0194)
From Cape Town				-0.515**			-0.168
				(0.239)			(0.344)
Coloured				-0.713			-0.713
				(0.643)			(0.942)
Back yard				0.0492			0.0418
				(0.183)			(0.277)
Migrant				-0.70***			-0.69***
				(0.0969)			(0.138)
Informal settlement				1.35***			0.89**
				(0.243)			(0.348)
Obs	2,694	2,654	2,648	2,648	1,074	1,074	1,074
Time	Yes	Yes	Yes	Yes	No	No	No
LL	-1430	-1390	-1364	-1284	-838.2	-829.9	-801.9

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. Dependent var in "L" is a dummy for treated in current period. Dependent var in "MNL" is categorical: 0 for never treated $t = 2, 3, 4$ for treated in period.

Each coefficient on the project-household pair variables estimates the common parameter specifying the effect of that variable on the probability of being selected for a project. Project dis: the coefficient of the distance from a housing project to the household. Rank 3: dummy variable = 1 if the enumeration area of the household is among the three closest EAs to the project. Project dis sq: project distance squared. In Project: dummy variable = 1 if the enumeration area of the household was contained within a housing project that was build within the last year. Informal settlement indicates the household was contained within an area recognized as an informal settlement by the state. Age ratio is the ratio individuals under 15 to individuals over 15 living in the household.

variable TD_i - indicating the period in which the household got the house (or 0 if not at all).

In the estimation I add a range of additional project-household variables to predict treatment to the basic specification . The distance between the household and the project ($\mathbf{ProjectDist_{ia}}$), a dummy variable indicating that the household was actually located within the boundary of the project ($\mathbf{Insitu_{ia}}$), and dummy variable indicating that the household's enumeration area was ranked among the three closest EAs to the project ($\mathbf{Rank_{ia}}$).[42] I also include a square term in the distance from projects, to capture non-linearities in the effect of distance.

Figure 2.2: Kernel density of predicted treatment for treated and untreated groups

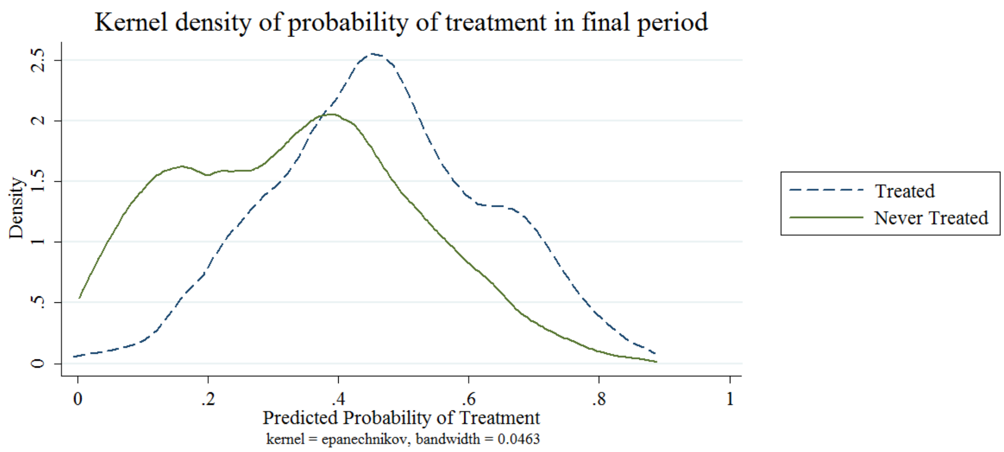


Table 2.3 shows large and significant impacts of distance from housing on the probability of receiving housing. Living in an area that was fulling upgraded "in situ" is also a significant predictor of treatment, as is being one of the closest three households. I discuss some marginal effects interpretation in the Appendix, Section B.1.2.

The combined effect of the distance has enormous predictive power on the probability of treatment in the data. I use the coefficients from the estimates in Table 2.3 to predict treatment in each time period ($\widehat{G_{it}}$). Using the coefficients from the estimation in Column 1 in Table 2.3, I plot the kernel density of pre-

42 Such a household might have been upgraded *in situ*, which was the case for some areas in Cape Town. This dummy variable indicates that the entire enumeration area (EA) was located within a project, not just that the boundary of the EA overlapped with a project.

dicted treatment by those that actually received housing, and those that did not, by the final wave of data from 2009. The results show a very clear right shift in the distribution for treated households.

I note a few other facts about observables that predict access to housing: female headed households are more likely to get housing, households living further away from the city are more likely to getting housing even when controlling for distance from projects, perhaps indicating that higher demand for housing closer to the city leads to housing being allocated differently.[43] Recent migrants (arrived in Cape Town in the last 5 years) are less likely to be treated, perhaps reflecting the fact that they joined the housing lists later and are there further down the waiting lists. Households living in communities classified as informal by the city government were also more likely to get housing: perhaps reflecting that formal slum recognition matters for ensuring access to services in this context.

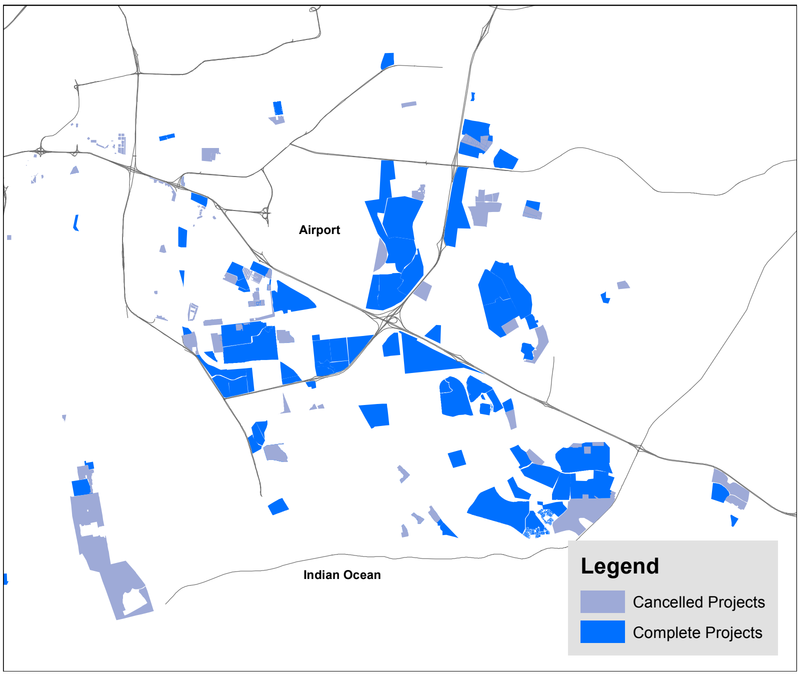
2.4.4 *Cancelled Projects*

The identification strategy using proximity instruments attempts deals with issues of selection into treatment based on individual characteristics. It does not necessarily account for non-random selection at the geographic level.

The identifying assumption of the IV strategy with fixed effects is that the chosen location of projects is not correlated with changes in households outcomes over time, except through the channel of government housing. This assumption may not hold for a number of reasons. Although all anecdotal reports suggest that communities had very little power in initiating new housing projects, which were mostly driven by the land demand needs of private construction companies, it is possible that certain connected individuals were able to lobby effectively for housing on the part of their communities. These individuals have been

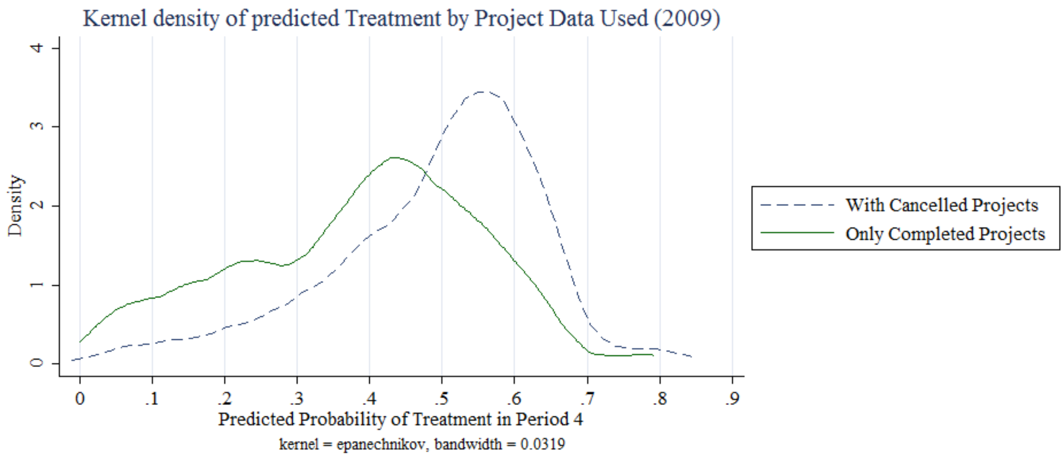
43 Perhaps ineligible, wealthier households were more likely to jump the queues and get these houses.

Figure 2.3: Comparison of complete and cancelled housing projects



successful at lobbying for other services or employment projects. Further, the government may have prioritized the development of certain neighborhoods or areas for political reasons, and simultaneously awarded those areas other social programs.

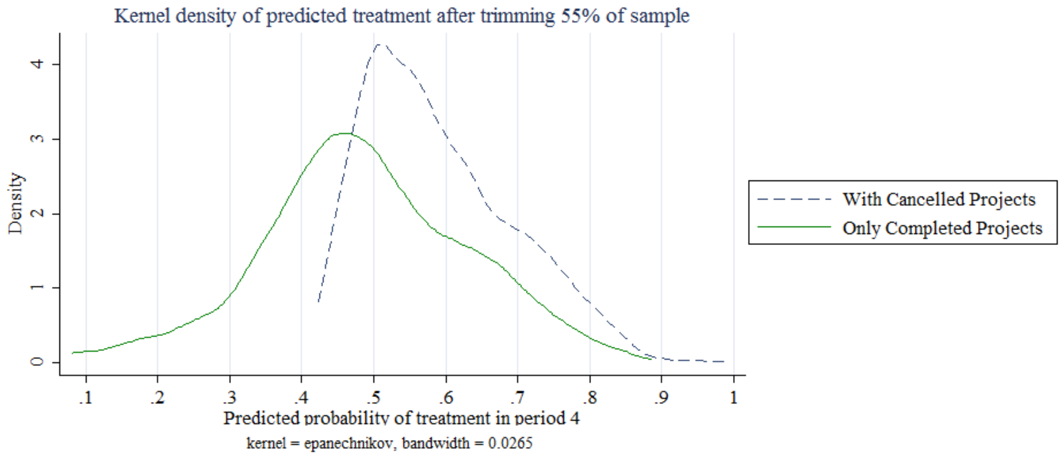
Figure 2.4: Predicted probability: comparison with and without incomplete projects



I present a set of robustness checks to show that the results are not driven by larger geographical variation in treatment: I restrict the estimation to certain areas and townships and find that the results hold within those sub-samples.

Similarly I argue that the results were not driven by targeting of certain ethnic groups geographically, by restricting the analysis to certain groups in turn. I also show that receiving government housing did not lead to the receipt of other, additional government grants. Finally, I argue that housing projects did not stimulate local employment by creating construction jobs. Construction was always done by external construction companies that brought in their own permanent labour force to the sites.[44]

Figure 2.5: Predicted probability of treatment in the trimmed sample



Most importantly, I use a natural experiment using housing projects that were planned and approved but cancelled for bureaucratic reasons. According to discussions with City officials, this was often due to bureaucratic and budget issues, or problems with the construction companies, as opposed to something inherent in the local communities in the area.[45] Table 2.3 shows a map of the cancelled projects along with completed projects, which gives an idea of the variation in treatment probabilities induced by project cancellations.

The idea of using cancelled projects is to compare households that lived near planned and cancelled projects to those that lived near housing projects that were

44 In addition, my results show the biggest impact on labour supply of women, who are unlikely to have got jobs on construction sites.

45 During the time of the study, there were numerous reports of housing projects that were cancelled or put on hold because the holding companies had become bankrupt

actually built, while excluding from the sample those households that lived in areas were projects were not even planned.

To do this, I generate a new dataset of distances from housing projects, *including* the projects that were cancelled.[46] This new dataset of distances, combined with coefficient estimates from Table 2.3, are then used to predict the probability of having been treated had *all* planned projects been completed. These predicted probabilities are the counterfactual probabilities of treatment had all housing been built.

Naturally, this new set of instruments has a smaller average distance to housing- there were a number of areas that would have been very close to housing projects had their nearby projects not been cancelled. As a result, the predicted probability of treatment (I call this the "placebo probabilities") with the completed projects is considerably higher with the cancelled projects included. Figure 2.4 shows the predicted probabilities of treatment with and without cancelled projects, and a clear right shift for the "placebo probabilities", and considerably less mass with probabilities less than 30%.

In order to concentrate on the differences between areas that were near cancelled projects and those near completed projects, I trim the sample by dropping those far away from both. That is, I drop individuals with a low probability of treatment with cancelled projects included. Since many individuals with a relatively high probability of treatment when cancelled projects are included have a relatively low probability of treatment with only completed projects, the distribution of the predicted treatment once the sample has been trimmed still has support over the full range of predicted probabilities. This is illustrated in Figure 2.5 where all households with probability of treatment less than 40% are dropped from the sample.

46 Of course, the distance instruments used until now included only distances from completed projects. I carefully verified that cancelled projects had indeed been cancelled, and that completed projects had indeed been completed, by using satellite imagery from the time.

In the main empirical results, I will re-estimate the impacts of receiving housing using trimmed samples, by iteratively dropping quintiles of the "placebo probabilities" and estimating both the fixed effects and IV models for those restricted samples.

2.5 MAIN RESULTS

In this section I present the main results of the impact of government housing on household outcomes.

As explained in the empirical strategy, I proceed in three steps: I estimate regular OLS models with household fixed effects, then use instruments to deal with selection on individual unobservables, and finally re-estimate both the FE and IV result with a trimmed sample that excludes areas that were far away from both cancelled and completed projects.

I start by looking at the impacts housing on log of household income. The income variable comes directly from the CAPS household data.[47] Throughout this section I use the label "house" to denote this coefficient of interest. Column 1 of Table shows the results from fixed effects regressions without any controls for changes in household composition, characteristics, place of living, or other sources of income shows large and significant increases in incomes for households receiving housing. This effect on total income is about 24%.

Columns 2 and 3 in Table 2.4 introduce controls for changes in household composition. I control for household size, changes in the sex and age (young-old) ratio in the house, as well as whether households were receiving household grants. Adding these controls reduces the size of the estimated effect of housing, but not significantly. The results are similar and still significant even when controlling for changes in household composition and size. This goes some way

47 This is the most comprehensive income variable which includes data from a one-shot total household income question, but excludes income from rent. Later I address issues due to sources of non-wage income that might be included in this measure.

Table 2.4: Effects of government housing on total household income

	(1) FE lginc	(2) FE lginc	(3) FE lginc	(4) IV lginc	(5) IV lginc	(6) IV lginc	(7) IVbal lginc
house	0.245***	0.192***	0.190***	0.552**	0.617**	0.559*	0.657**
	(0.0539)	(0.0509)	(0.0512)	(0.2695)	(0.310)	(0.313)	(0.313)
femalehd		-0.188***	-0.188***		-0.172**	-0.174**	-0.144*
		(.0477)	(.0480)		(.0712)	(.0743)	(.0798)
hhsize		0.146***	0.122***		0.142***	0.120***	0.112***
		(.0085)	(.0126)		(.0137)	(.0177)	(.019)
sexratio		-0.174	-0.228*		-0.178	-0.231	-0.317*
		(0.114)	(0.122)		(0.140)	(0.144)	(0.166)
youngratio		-0.60***	-0.56***		-0.59***	-0.55***	-0.63***
		(0.0921)	(0.104)		(0.115)	(0.129)	(0.163)
femadults			0.0353			0.0360	0.0603*
			(0.0272)			(0.0325)	(0.0337)
citydis			-0.00705			-0.0128	-0.0212*
			(.0079)			(.0089)	(.0121)
maxedu			0.0209***			0.0203***	0.0145*
			(.0060)			(.0078)	(.0084)
maxage			0.00159			0.000842	0.000417
			(.0017)			(.0024)	(.0025)
govgrants			0.0679*			0.0744	0.0899
			(0.038)			(0.049)	(0.059)
Obs	3,590	3,590	3,570	3,572	3,572	3,547	2,526
R^2	0.191	0.284	0.291	0.1828	0.264	0.277	0.272
Groups	1,077	1,077	1,076	1,059	1,059	1,053	661
AvGroup	3.333	3.333	3.318	3.373	3.373	3.368	3.821
McKinnonF	.	.	.	2.62*	5.51***	4.31***	2.26***
WeakIVF	.	.	.	70.24	96.98	96.95	91.94

Notes: Clustered standard errors in parentheses.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. house=1 if household reported getting a subsidized house at any point in the past. Dependent variable is the log of total household income. The column IV-Bal is a replication of the results with a fully balanced panel- households that appear in very wave of the data. All IV regressions use a non-linear predicted probability of treatment as an instrument.

to showing that the effects are not driven by new incomes from new household members.

I then turn to the IV results. As outlined in the estimation section I use the predicted probabilities from Equation (2.13) and Table 2.3 (Column 1) in each time period as instruments in a two staged least squares estimator. The estimation results of this first stage of the 2SLS is presented in the Appendix B.4.

All IV regressions report the Kleibergen-Paap Wald F statistic for weak instruments.[48] There is little reason to suspect a problem of weak instruments, given the very high Wald F-statistics on the test for insignificant instruments in the first stage, reported in the last row of each table.

I find that the IV results estimate large and significant effects of housing on household income. The coefficients are generally larger than the OLS results, which indicates either correlation between the probability of treatment and *negative* income shocks or a problem of measurement error.[49] As discussed in the conceptual framework, this could be due to individuals in worsening circumstances (like the aged or recently unemployed), or victims of recent shocks, being assigned houses by their communities, which leads to downward bias of the treatment effects of housing. We might have expected them to have done considerably worse without the treatment. Finally, under the LATE interpretation of my results (see Section 2.6.5) my IV results identify the effects for households who did not have to move far to get housing, and thus do better than households that were required to relocate.

48 This is the analogue of Cragg-Donald Wald F statistic, under the assumption of heteroskedastic or serially correlated errors. Note that in the single endogenous regressor case, such as this one, this statistic is equivalent to standard Wald test on the first stage coefficient on the single instrument (Baum and Schaffer, 2007).

49 I report post estimation tests of endogeneity, which, in the fixed effect IV setting is the Davidson-MacKinnon F statistic. This test rejects the null that the IV estimates are the same as the OLS model. See Davidson and MacKinnon (1993). For the execution of the Fixed Effects Instrumental Variable model with clustered standard errors and 2-Step GMM I use the stata command xtivreg2, and the *dmexogxt* post estimation command for exogeneity test of OLS vs XT-IV estimation (Baum and Stillman, 1999).

Table 2.5: FE and IV Impacts on different earnings measures

	(1) total salary	(2) log total salary	(3) log income/person	(4) log head's salary
Panel A: OLS with Fixed Effects Impacts on Different Income Measures				
house	348.8*	0.242***	0.199***	0.137**
	(195.5)	(0.0649)	(0.0717)	(0.0664)
Observations	2,273	1,837	2,438	1,159
R-squared	0.171	0.158	0.236	0.197
Number of hhs	574	574	636	416
Av Group Size	3.960	3.200	3.833	2.786
Panel B: Instrumental Variables Impacts on Different Income Measures				
house	568.0	0.520*	0.721**	0.547*
	(801.6)	(0.304)	(0.344)	(0.312)
Observations	2,518	1,837	2,438	1,159
R-squared	0.099	0.117	0.185	0.157
Number of hhs	637	574	636	416
Av Group Size	3.953	3.200	3.833	2.786
Weak IV F	79.21	63.30	75.94	44.08

Notes: Clustered standard errors in parentheses. All IV regressions use a non-linear predicted probability of treatment. house=1 if household reported getting a subsidized house at any point in the past. Dependent Variable is the log of total household income from wage earnings
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Next, I confirm that these results are due to higher average wage earnings for members of the household. The data on household income used thus far came from a combination of questions, including a one-shot question on household income when individual earnings were missing. This may have contained sources of income not related to labour market activities. I use estimates of the sum of household income from earnings data from household rosters. However, the data is more often missing for these variables, which yields less precision for the estimates. Panel A shows the OLS fixed-effects results, Panel B the IV results. The coefficients on the impacts on household income in logs are significant and similar in magnitude to those in Table 2.4.

I also estimate the impact of housing on the per capita income of household members, to rule out that the effects on incomes were being driven by increases in household size. I find that receiving housing had a significant impact on the total salary earned by the household, and on the average earnings per person in the household. In Column 1 I estimate the impact on the total salary in levels. This is estimated with less precision, but estimates an average impact on household income of about R350. Average household earnings in the final wave around R2500.

2.5.1 *Cancelled Projects*

In this section I use the data on cancelled projects to deal with non-random selection of project sites. As outlined by the identification strategy, I use the predicted probability of receiving housing using the projects that were cancelled (I call this the "placebo probability") in order to drop areas that were never considered for housing projects. Since these areas might be systematically different from areas in which housing was planned.

I trim the sample by dropping different quintiles of the "placebo probability". So I start by dropping the 20% of household least likely to be treated, then 40%, and so on. The remaining variation in the probability of receiving housing under the true predicted probability of treatment (excluded cancelled projects) is then driven by the projects that were built versus those that weren't built. The more of the sample that is trimmed, the more of the remaining variation is due to project cancellations alone (although the sample sizes get considerably smaller).

I find that the results are robust to the trimming the sample in this way. Table 2.6 shows the impacts on total household earnings for the usual FE (Panel A) and IV (Panel B) estimates. The coefficients are similar to those in Table 2.4 and significant, and are stable as I gradually drop more of the sample. Only when I have trimmed a whole 80% of the sample (Column 4) are the impacts no longer

Table 2.6: Impacts on total income with trimming using cancelled projects

	(1)	(2)	(3)	(4)
Sample Trimmed %	**20**	**40**	**60**	**80**
Panel A: OLS with Fixed Effects Impacts on Log Total Income				
house	0.163**	0.203***	0.163**	0.116
	(0.0650)	(0.0750)	(0.0793)	(0.0799)
Observations	2,874	2,219	1,507	743
R-squared	0.272	0.271	0.328	0.486
Households	831	625	417	206
Panel B: Instrumental Variables Impacts on Log Total Income				
house	0.886**	1.313**	1.409**	0.639
	(0.422)	(0.571)	(0.659)	(0.544)
Observations	2,880	2,224	1,510	745
R-squared	0.209	0.123	0.108	0.420
Households	829	623	416	206
Av Group Size	3.474	3.570	3.630	3.617
Weak IV F	53.35	37.73	14.92	15.83

Notes: Clustered standard errors in parentheses. All IV regressions use a non-linear predicted probability of treatment. house=1 if household reported getting a subsidized house at any point in the past. Dependent Variable lginc is the log of total household income from all sources *** $p<0.01$, ** $p<0.05$, * $p<0.1$

significant. The estimated coefficients are slightly smaller, but still large and positive. I show that these results hold for the FE estimates for total household salaries in Table 2.7.[50]

These results provide evidence that the the results in this paper are not driven by non-random project site choices, under the assumption that projects were cancelled were not cancelled for reasons related to the outcomes of households in those areas. So particularly well organized or motivated communities that were able to work hard to get their projects completed, these communities might have also have been able to improve their communities in other ways, and get

50 The IV estimates for these measures quickly become imprecise, with many missing outcomes and the trimmed sample reducing the sample size further. The results are not presented here.

jobs for those nearby. However, discussions with City officials and urban planners suggested that the reasons for project cancellations were rarely to do with communities living there. They were more likely to be related to changes in budgetary issues, or disputes with developers and contractors.

Table 2.7: Impacts on log total salaries with trimming using cancelled projects

	(1)	(2)	(3)	(4)
	FE	FE	FE	FE
Sample Trimmed %	**20**	**40**	**60**	**80**
house	0.189**	0.210**	0.191**	0.0751
	(0.0741)	(0.0797)	(0.0867)	(0.0910)
Observations	2,167	1,720	1,171	590
R-squared	0.130	0.158	0.199	0.352
Number of personid	719	557	373	184
Av Group Size	3.014	3.088	3.139	3.207

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. Dependent var is the log of the sum of all monthly salaries earned by members of the household. All regressions include controls for time varyings household characteristics and household fixed effects.

2.5.2 *Restricting to Treated Areas Only*

The fixed effects regression test whether treated households did better than untreated households, while the IV results test whether households living closer to projects did better than those living further away. The trimming with cancelled projects tests whether households living near completed projects did better than households living near cancelled projects. The skeptical reader may still be unconvinced that the project placement or cancellation decisions were endogenous, so that the effects of housing may be due to differences in the growth rates of certain *areas* of the city.

I test whether treated households living in areas that were close to housing projects, and where *most* of the households were treated, and compare the outcomes of households that were treated to those that were not. I look at clusters (primary enumeration areas) where more than 20% of households received a government house over the period of 4 years (this is exactly half of all clusters, 45% of clusters had no-one receiving housing). The results are presented in the Appendix, Table B.7. I find that the results are very similar to the results presented on total household incomes and salaries. Note that of course I only present the FE estimates here, the aim is to compare households within clusters, and the distance instruments do vary within clusters.[51]

This should rule out a story under which certain high growth neighborhoods and areas saw particularly high growth, and were also targeted for housing project investments at the same time. I return to further robustness checks of this kind in Section 2.7.

2.5.3 *Linear IV Results*

The results presented thus far have used the predicted probabilities of treatment from my maximum likelihood estimator as instruments in a linear 2SLS estimator. This strategy was motivated by the need to more efficiently predict the probability of treatment. In this section I motivate this strategy further by showing the results for a simpler 2SLS estimator. I use multiple project-distance and project-rank measures directly in the first-stage of the IV estimator. I show that the estimate treatment effects are of similar magnitudes to those when the more complicated first stage is used. This is reassuring and fits with the intuition of the (Wooldridge, 2002) that my results are not driven by non-linearities in the functional form of the first stage.

These results are presented in the Appendix, Table B.6. For instruments I use distance to the closest 5 projects, the ranking of household among the first 3 closest projects, and whether or not the EA was inside a project as instruments

51 The results are also robust to the specification of cluster fixed effects, instead of individual fixed effects.

(2SLSa) (Columns 1-3). In the second specification (2SLSb) I use squared project distance instruments as well (Columns 4-6). I show results for some of the main specifications in the paper and find similar coefficient estimates. However, I show that inference is not valid given precision of the first stage estimators. The F-Kleibergen-Paap F statistics for weak instruments are very low: we cannot reject the null of no effect of the instruments on treatment in the first stage. This problem is improved by the inclusion of the distance squared as a regressor, but Stock Yogo maximal bias critical values (Stock and Yogo, 2005) still show that we cannot reject the null of up to 20% bias in IV results. Further, I have estimated the results using limited information maximum likelihood (LIML) estimation techniques in which standard errors are correct in the presence of weak instruments: these estimators inflate the standard errors such that the effects are no longer significant. These results justify the use of the more precise first stage estimator.[52]

2.6 MECHANISMS

In this section I look for mechanisms through which households who received government housing were able to increase their household incomes. I have documented that these increases in household income were due to increases in the total wage earnings of individuals in the household.

In this section, I argue that housing enables household members to leave the home and go out in search of (more) work, because it alleviates the usual burdens of home production associated with living in informal housing. I document three main facts about labour outcomes in treated households. I find that female members of the households are more likely to be employed after receiving government housing, and that the same results do not apply to male household members. I find clear evidence that the household earnings from females *and*

52 My non-linear estiamte is considerable more efficient than all other linear first stage estimators, even when the distance measures are aggregated into a single composite measaure, which improves over the efficiency of the basic 2SLS presented here, but not by much.

males increased significantly after receiving government housing. For young adult members of the household, for whom I have more detailed labour data, I find significant impacts of hours worked.

I then look for mechanisms through which labour supply might be restricted by poor housing conditions. I find that receiving government housing significantly increases feeling of safety. Receiving housing seems to have moved households further away from the city centre and jobs. However, these effects are small, compared to anecdotal accounts of how far houses do move in Cape Town. I also find that improved housing is associated with a number of measures of labour saving technologies in the home. Finally, I discuss other mechanisms through which housing might be driving the results, and some limitations of the data to identify these.[53]

2.6.1 *Labour Supply*

I estimate the effect of government housing on the total number and proportion of women in the household working in last 7 days before the survey. The results show a significant increase in female labour supply in households that received government housing. The effects are there in the fixed effects and the IV regressions.

The impact on the proportion of females employed in the household is not significant in the fixed effects regression. I find that this is because the treatment effects do not seem to be present in the early waves of the data- waves 2 and 3. The effects are large and significant however, for the last wave of the survey. This could be because the effects of moving into new housing take a while to have an effect for female household members- the adjustment to living in a new home could delay the benefits of that housing. As a result I look at the fixed effects

53 The data places certain limitations on the mechanisms that can be explored. The CAPS data used in this data focused on the young adult members of the household, and most of the detailed questions about labour supply and employment is recorded only for those individuals. Similarly outcomes related to health and small scale enterprises are not well measured in the data.

Table 2.8: Impacts on male and female labour supply at the extensive margin

	(1)	(2)	(3)	(4)	(5)
	FE	FE	FD	IV	IV
	num employed	% employed	% employed	num employed	% employed
Panel A: Employment among Female Household Members					
subhere	0.102*	0.0285	0.0842*	0.488**	0.325**
	(0.0532)	(0.0291)	(0.0451)	(0.237)	(0.144)
Observations	2,471	2,471	1,237	2,459	2,458
R-squared	0.032	0.007	0.017	0.087	-0.017
Number of hhs	658	658	654	646	646
Av Group	3.755	3.755	1.891	3.807	3.805
HH Controls	Yes	Yes	Yes	Yes	Yes
Weak IV F				96.03	86.43
Panel B: Employment among Male Household Members					
subhere	0.00256	-0.0200	-0.0497	0.0348	-0.0846
	(0.0502)	(0.0365)	(0.0584)	(0.234)	(0.171)
Observations	2,471	1,906	947	2,459	1,836
R-squared	0.023	0.014	0.013	0.191	0.059
Number of personid	658	609	589	646	539
Av Group	3.755	3.130	1.608	3.807	3.406
HH Controls	Yes	Yes	Yes	Yes	Yes
Weak IV F				96.03	40.89

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable "num employed" is the number of men/women employed in the household (regressions include controls for the number of women/men in the household). % employed is the proportion of women/men of household members who are employed.

regressions with just the first and forth wave of the data, as a two period first difference estimate.

These results are presented in Table 2.8, Panel A. The fixed effects results show that households increase both the number of women working in the household (controlling for the number of women living there), and the proportion of women in that household working, by between 8% and 10%. Panel B then looks at the same set of results for men and finds no significant effects.[54] There is no impact of housing labour supply of male members of the household. Men were more likely to be employed than females, by 58% to 44% among adults in the sample.

Table 2.9: Effect of government housing on hours worked per day (young adults)

	(1)	(2)	(3)	(4)	(5)
	FE	FE	IV	IV	IV-Bal
house	0.642**	0.599**	2.289*	2.482*	2.717*
	(0.252)	(0.256)	(1.346)	(1.286)	(1.480)
HH Chars	No	Yes	No	Yes	Yes
Obs	1,630	1,605	1,295	1,293	969
Groups	821	813	502	501	356
Av Group	1.985	1.974	2.580	2.581	2.722
Weak IV F			17.20	21.38	9.232

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable is the number of hours worked for young adults in the household, conditional on having done some work. The column IV-Bal is a replication of the results with a fully balanced panel- households that appear in very wave of the data.

I look for impacts on employment rates among young adult members of the household. The impacts are small, positive, but not significant in the FE, although significant in the IV results (Appendix, Table B.5). However, I have data on hours worked per day by young adults in the sample (this data is not available for all household members). I find that the receiving housing increased the number of hours worked in wage labour, in both the OLS fixed effects regressions, and the IV regressions. This suggests that young adults increase their labour supply at the intensive margin since work hours are only observed for those working. The results suggest that young adults worked more than an ad-

54 The coefficients in these regressions are negative, perhaps suggesting some substitution away from labour for men, after women increase their labour supply, but these results are not significant and should not be over-interpreted.

ditional half an hour per day on average, over a mean of about 8 hours a day among those that work.[55]

2.6.2 *Female and Male Earnings*

Table 2.10 presents the main results on the impacts of housing on household male and female incomes. I look at two outcomes: in Columns 1 and 2, the impact on the sum of all male and female salaries earned by household members. In columns 3 and 4 and the average earners *per earner* for males and females. This allows me to capture the full effect of increased earnings in the household, including the effect of increased employment (the extensive margin) but also for increased earnings conditional on being employed.

The results show that receiving government housing seems to have increased earnings among both male and female members of households. The results are consistent with a story where females were more likely to be working, but do not seem to be earning more, conditional on earning (the coefficient is large but not significant).

For incomes earned by male household the fixed effect results suggest that earnings by male members increased, but that the impact is driven solely by increases in the average salaries of male household members, conditional having work (Panel B, Column 3 in Table 2.10). This finding is consistent with the result that government housing had no significant impact on the probability of employment among male household members. However, the IV results are less clear. The findings of an impact on male earnings are not robust to the use of the instrumental variables. The estimated treatment effects are smaller, and not significant, but still positive in the same direction as the fixed effects results.

55 These results are robust to using log-hours worked, with an estimated impact of about a 10% increase.

Table 2.10: Effect of government housing on earnings of female & male household members

	(1) FE	(2) IV	(3) FE	(4) IV
	sum earnings	sum earnings	average earnings	average earnings
Panel A: Impacts on Earnings of Female Household Members				
house	0.200**	0.832*	0.148	0.900*
	(0.0899)	(0.500)	(0.0940)	(0.470)
Observations	1,164	1,164	1,139	1,133
R-squared	0.230	0.087	0.115	0.027
Number of hhs	412	412	412	406
Av Group Size	2.825	2.825	2.765	2.791
Weak IV F		36.00		34.05
Panel B: Impacts on Earnings of Male Household Members				
house	0.245***	0.0554	0.255***	0.113
	(0.0888)	(0.308)	(0.0912)	(0.254)
Observations	1,114	1,113	1,098	1,090
R-squared	0.108	0.103	0.096	0.100
Number of hhs	398	398	398	391
Av Group Size	2.799	2.796	2.759	2.788
Weak IV F		35.96		34.21

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable "sum earnings" gives the total earnings brought into the household by females and males, in Panels A and B, respectively. Dependent variable "average earnings" gives the average wage income earned (conditional on earning) by female and male household males of the household, in Panels A and B, respectively.

2.6.3 *Home Production*

The theoretical framework in Section 2.2.2 postulated a link between female supply and the constraints imposed by work at home due to living in informal housing. In this section, I show that receiving government housing lead to significant improvements in housing quality and access to labour saving technology.

Table 2.11 presents fixed effects estimates of treatment on physical housing conditions. Unsurprisingly, government housing reduces the probability that households are living in a shack. Treated households are significantly more likely to own a stove, a fridge and a microwave. This is possibly because housing provides more space and security to keep such an appliance, or because access to electricity is more readily available. I find no significant impact on the probability that the household has *any* access to electricity. But housing does significant increase the probability of having access to piped water, and in particular piped water in the home. All of these are technologies that could provide significant time savings for women working in the home.

Government housing reduces the probability that households use paraffin- a form a fuel commonly used in informal settlements in South Africa. Many of the devastating fires that occur in townships in South Africa are attributed to the use of paraffin. There is a negative effect on the occurrence of fires in the home (although this coefficient is not significant, it was only measured once in the follow survey rounds) which could be driven by the reduction in paraffin use.

2.6.4 *Safety and Security*

One mechanism that could be driving the impacts on increased female labour participation and earnings could be the improved security that good quality housing gives to households. This could allow them to leave the home to take up employment or go in search of work without fear of burglary in their absence. This is related to the hypothesis of Field (2005) who argues households increase their labour supply when security of tenure is improved and they less likely to fear expropriation when they are out of the home.

While expropriation risk may be at play in this setting, de facto tenure security is already thought to be very good in South African informal settlements. On the

Table 2.11: Effects of government housing on living conditions in the home

	(1)	(2)	(3)	(4)	(5)
	FE	FE	FE	FE	FE
Panel A: Impact on Household Access to Services					
	Shack	Electricity	Toilet	Piped	Piped In
subhere	-0.579***	0.000640	0.00995	0.102***	0.242***
	(0.0206)	(0.0208)	(0.0210)	(0.0206)	(0.0219)
Observations	3,750	3,789	3,789	3,789	3,789
R-squared	0.480	0.041	0.077	0.028	0.157
Number of personid	1,095	1,095	1,095	1,095	1,095
Panel B: Impact on Ownershipo of Household Appliances and Fuel Use					
	Stove	Fridge	Microwave	Paraffin	Fire
subhere	0.0445*	0.0908***	0.0575**	-0.0439*	-0.0171
	(0.0267)	(0.0259)	(0.0250)	(0.0237)	(0.0193)
Observations	3,749	3,749	3,748	2,942	2,949
R-squared	0.149	0.064	0.158	0.002	0.036
Number of hhs	1,095	1,095	1,095	1,095	1,095

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variables are dummy variables if the household has: Panel A, Col (1) an informal dwelling, Col (2) access to Electricity, Col (3) access to a flushing toiler, Col (4) access to piped water inside the house or nearby, Col (5) access to piped water in the home. Panel B, Col (1) a stove, col (2) a refrigerator, Col (3) a microwave , Col (4) used a paraffin stove for cooking or heating, Col (5) experienced a fire in the home.

other hand security threats from burglary and other crime around the home are far more salient threats in this environment. If this is part of the mechanism driving these results, we should see some evidence that adequate housing has positive impacts on feelings of safety around the home. I provide evidence of this in this in Table 2.12. I find that households that received government housing were far less likely to report that they felt unsafe in their homes at night.

2.6.5 *Distance and Travel*

Government housing in South Africa has been criticised for reinforcing the spatial patterns of segregated living within South Africa cities (Bundy, 2014). Segre-

Table 2.12: Effects of government housing on feelings of being unsafe at home at night

	(1) FE unsafe	(2) FE unsafe	(3) IV unsafe	(4) IV unsafe	(5) IVBal unsafe
house	-0.325***	-0.134**	-1.711**	-1.169*	-0.714*
	(0.0548)	(0.0643)	(0.818)	(0.655)	(0.411)
HH Chars	No	Yes	No	Yes	Yes
Observations	1,408	1,376	1,352	1,344	816
R^2	0.039	0.103	-0.631	-0.210	-0.032
Av Group Size	2	1.955	2	2	2
Weak IV F			8.723	11.47	23.92

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable is a dummy variable = 1 if the household reports feeling unsafe at night in the home.

gation leaves households living far away from jobs and employment opportunities, which is argued to play a causal role in the high rates of urban unemployment for black South Africans (Banerjee et al., 2007; Rospabe and Selod, 2006). In fact government housing is often thought to have moved household *further* away from the original place of living, as new housing projects are built increasingly far away.

I confirm that government housing has moved households slightly further away from the city, on average. Beneficiary households have not had their situation improved with regards to distance from the city. Households that receive the government houses move on average 600m further away from the city during the course of the survey.

Further, this estimate is likely to be an underestimate of the impacts on distance, since for some households I was not able to geocode their new location, or their location was not updated by enumerators in the data files. I restrict the sample to households for which a move is recorded. Here I my point estimates indicate that treated households moved nearly 1.5kms further away from the

Table 2.13: Effect of government housing on distance from the city center (in kms)

	(1)	(2)	(3)	(4)	(5)	(6)
					Just Movers	
	FE	FE	FE	IV	FE	IV
	citydis	citydis	citydis	citydis	citydis	citydis
house	0.508**	0.540**	0.537**	0.0416	1.441**	0.917
	(0.221)	(0.228)	(0.229)	(0.525)	(0.587)	(2.205)
Obs	3,765	3,725	3,717	3,708	1,243	1,243
R^2	0.008	0.010	0.011	0.005	0.028	0.026
Households	1,077	1,077	1,077	1,068	362	362

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable citydis is the distance of the household from the city center in kms.

city than untreated households. Given that the average distance from the city is about 22kms in this sample, this is not a huge difference in relative terms. In this sense the IV results isolate the effects of improved housing without the usual large effect on displacement that comes with large housing projects of this kind. Government housing had a positive impact on household labour supply and earnings, in spite of whatever effect housing did have on distance.

LATE and distance

The results discussed in this paper deserve one important caveat, related to this issue of distance discussed above. The impact of housing estimated by instrumental variables should be interpreted as a local average treatment effect (LATE) (Angrist and Imbens, 1994). The instruments identify the effect of treatment on those household that received housing because of their proximity to housing. These are known as "compliers" in the framework of Angrist and Imbens (1994): in the potential outcomes framework they receive government housing if and only if a housing project is constructed nearby.

The average treatment effect (ATE) is identified by the LATE under the assumption of homogenous treatment effects (Heckman, 1990). If households that were selected to receive housing because of their proximity to housing respond differently to treatment than those who receive housing for different reasons, then this assumption is violated.

This group of compliers is likely to differ from other housing recipients in at least one way. They are less likely to have moved a significant distance from their original place of living when receiving housing, precisely because the housing to which they were assigned was close by to their original place of living. Not all individuals were assigned housing because of their proximity to housing. A significant number of individuals were able to access housing far away from their place of living. Among households with predicted probability of receiving housing $\widehat{G}$ above the median, 47% received government housing. Among those below the median of value of $\widehat{G}$, 24.3% received housing.

In this way, the estimated impact on distance moved may not fully capture the way in which households were relocated further away from the city. If individuals who moved further away from the city centre were likely to suffer worse outcomes because of their distance from jobs and communities, this could lead to the LATE estimator being an over-estimate of the average treatment effect. This might explain some of the difference between the OLS fixed-effects estimates and the IV results presented in this paper.

2.7 ROBUSTNESS CHECKS

In this section I show that the main results on household income, female labour supply and household wage earnings, are robust to a variety of additional robustness checks.

I focus on the validity of the instrumental variables strategy. The validity of the instrumental variables results would be undermined by challenges to the exclusion restriction: that is that the location of households in relation to new housing projects is not correlated with the outcomes of interest except through the channel of government housing.

This assumption could be violated if project placement was driven by community organizing. I have tried to argue that housing location decisions were more commonly made at levels above that of the communities. Secondly, I have used the data on planned and cancelled projects to show that it is the completion of projects, and not just the choice of location, that is driving the results. Trimming had no effect on the results.

Here I check for robustness by checking that the results are not driven by ambitious households moving closer to housing projects in order to gain access to housing. Secondly, to mitigate measurement error in the timing of housing improvements, I estimate a two period first difference model with just the first and last periods. To ensure that the results are not being driven by politicians targeting new housing projects to high growth areas or areas where other urban initiatives were being rolled out, I restrict the analysis to specific areas of the city, and to specific ethnic groups. Finally, I check for impacts of housing on sources of income other than wage earnings.

2.7.1 *Opportunistic Relocation*

The exclusion restriction would be violated if households were able to move home in order to access housing. Given that the allocation procedures are well known, highly motivated and mobile households could be able to move closer to projects that were going ahead before allocation took place, in order to get a house. This would mean that more motivated households would also be more likely to be living closer to housing projects.

In practice, this sort of activity is unlikely to have occurred. Most households that receive government housing have either been on the waiting lists in the local

Table 2.14: Replication of the key results without households treated before 2005

	(1) total income	(2) total salary	(3) female salaries	(4) male salaries	(5) female % employed	(6) male % employed
Panel A: Fixed Effects Regressions (Multiple Outcomes)						
house	0.268***	0.321***	0.341***	0.202*	0.0718*	0.0335
	(0.101)	(0.103)	(0.125)	(0.120)	(0.0379)	(0.0627)
Observations	1,900	1,467	966	918	1,844	1,407
R-squared	0.256	0.127	0.172	0.114	0.036	0.045
Number of hhs	498	489	415	385	494	454
Av Group Size	3.815	3	2.328	2.384	3.733	3.099
Panel A: Instrumental Variables Regressions (Multiple Outcomes)						
house	1.401**	1.982**	1.699	-0.154	0.472*	0.298
	(0.698)	(1.004)	(1.127)	(0.978)	(0.260)	(0.310)
Observations	1,900	1,429	858	828	1,835	1,349
R-squared	0.158	-0.120	0.016	0.097	-0.046	0.014
Number of hhs	498	451	307	295	484	396
Av Group Size	3.815	3.169	2.795	2.807	3.791	3.407
Weak IV F	27.57	19.95	13.75	6.726	31.46	23.61

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable in Columns (1)-(4) are the log of total income, salaries of household members, salaries of female household members and salaries of male household members, respectively. Dependent variable "% employed" gives the proportion of female and male household members currently employed, in columns (5) and (6) respectively.

areas in which they receive housing for many years before they are given this housing. It is unlikely that households would be able jump the queue after moving to an area so recently. Also all instruments have used distance from housing in the *first* wave of the data. Thus if households moved during the period 2002 to 2009 in order to access new housing projects, this change in proximity would not be reflected in the instruments.

Still, I want to rule out the possibility that the effects were driven by households who had just moved into new areas in 2002 (the first wave). I drop all households from my sample there were already treated by wave 2 (2005). This means that everyone who was treated before 2005 is dropped from the sample. I then replicate the main results presented in the paper, with the instrumental variables and full set of controls, with only households that were treated after 2006. I find that almost all of the results presented above remain robust to this check.[56] These results are presented in Table 2.14, Panel A shows the fixed effects, Panel B the IV estimates. Thus if the results are driven by successful households relocation decisions, it would have to have been that they moved to their 2002 location in order to pursue a house that they would only get after 2006. While housing projects were often subject to delays, it would be unusual if they took longer than 5 years to build.

2.7.2 *A two period Diff-in-Diff*

One concern for my identification strategy is that the instruments used are good at predicting whether or not a household will be treated in any period, but aren't as accurate at predicting *when* a household will be treated. This is usually because of inaccuracy of data on when housing projects were completed.[57] Thus, using changes in the predicted probability of treatment in the interim periods might be misleading and reduce the efficiency of estimates. To control for this I restrict the sample to just the first and last periods, and compare the changes in key dependent variables using predicted probabilities of ever being treated in any wave. These results are presented in table 2.15 and consistent with the results presented thus far, if a little larger after these issues of measurement error are dealt with.

56 Although standard errors are larger since so many households were treated between 2002 and 2005.

57 Sometimes the data indicates the completion date of the project, but not the date of when households in the area were actually able to move in.

Table 2.15: Replication of the key results with a 2-Period first difference estimator

	(1) total income	(2) total salary	(3) female salaries	(4) male salaries	(5) female % employed	(6) male % employed
Panel A: Fixed Effects OLS Regressions (Multiple Outcomes)						
house	0.285***	0.265**	0.218	0.285	0.0738*	-0.0380
	(0.0919)	(0.133)	(0.187)	(0.182)	(0.0433)	(0.0664)
Observations	1,393	837	404	376	1,360	847
R-squared	0.455	0.266	0.473	0.209	0.050	0.038
Number of hhs	699	419	202	188	680	425
Panel B: Instrumental Variables Regressions (Multiple Outcomes)						
house	0.526	0.981**	1.109*	0.413	0.252*	-0.113
	(0.336)	(0.485)	(0.600)	(0.468)	(0.140)	(0.167)
Observations	1,388	836	404	376	1,360	850
R-squared	0.448	0.194	0.361	0.206	0.023	0.032
Number of hhs	694	418	202	188	680	425
Av Group Size	2	2	2	2	2	2
Weak IV F	89.97	46.66	22.80	30.57	93.86	47.68

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable in Columns (1)-(4) are the log of total income, salaries of household members, salaries of female household members and salaries of male household members, respectively. Dependent variable "% employed" gives the proportion of female and male household members currently employed, in columns (5) and (6) respectively.

2.7.3 *Impacts within subpopulations*

In this section I show that the main results are present *within* areas and subcommunities, rather than driven by differences in treatment between these subpopulations. This reinforces the evidence that the effects aren't driven by geographic selection of project sites. It could be that projects were targeted to certain parts of the city- parts of the city that had different trajectories over time. In particular housing could have been targeted towards entire townships, racial groups, or areas a certain distance from the city, that were exhibiting other changes at the time. Indeed it does appear that, on average, housing was built in areas further away from the city, in black areas, and in poorer areas.

Table 2.16: Average household characteristics by major townships

	Gugulethu N=222	Khayelitsha N=608	Other N=247
Closest Project Distance	0.413	0.144	2.560
Distance to the City	18.57	29.75	21.32
Treated	30%	41%	19.4%
Coloured	7.7%	3.6%	48.8%
Log Income	7.245	7.303	7.598

I divide my sample into three: households living in and around Gugulethu, those living in Khayelitsha, and those living outside of these two townships.[58] Gugulethu and Khayelitsha are the two largest townships, or groupings of townships in the city and are where most of the construction of RDP housing has taken place. Settlements outside of these two townships are relatively neglected. The mean household characteristics for the different areas are presented in Table 2.16. Clearly other informal settlements have less housing construction, are less likely to get houses, and are more likely to be coloured families. Khayelitsha is further from the center, while Gugulethu is relatively close.

Table 2.17 shows the results of regressions of log income and employment on treatment, where the sample is restricted to various subgroups. For compactness, I report the treatment effects coefficient for different subgroups in the same columns (both FE and IV). I look in turn at the three different area in turn, just the poorest half of the sample and just the black sample. Sample sizes are small in some of the specifications, but the coefficients remain similar in magnitude to the original results for both FE and IV results for the impacts on household

58 Refer to the map of Cape Town in the Data section. I define Gugulethu rather more broadly than its strict geographical boundaries; including the townships of Weltevreden Valley, Nyanga, Manenberg and Crossroads Informal Settlement.

Table 2.17: Replication of impacts on Log Income within communities and sub-samples

	(1) FE	(2) IV	(3) FE	(4) IV	(5) FE	(6) IV
	Khayelitsha		Gugulethu		Other	
house	0.262***	1.243*	-0.0150	0.247	0.191***	0.564*
	(0.0844)	(0.733)	(0.0920)	(0.595)	(0.0604)	(0.314)
Obs	1,471	1,470	730	726	3,582	3,561
R^2	0.276	0.168	0.364	0.357	0.290	0.275
	Out of Project		Poor		Black	
house	0.268**	0.988*	0.116*	0.839	0.178***	0.914**
	(0.102)	(0.574)	(0.0696)	(0.562)	(0.0641)	(0.432)
Obs	1,389	1,377	1,733	1,731	3,015	2,998
R^2	0.328	0.294	0.504	0.451	0.264	0.203
HH Chars	Yes	Yes	Yes	Yes	Yes	Yes
Weak IV F		19.61		13.07		95.24

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Dependent variable is log of total household income.

income. This suggests that the results are not driven by difference *across* communities, although I find no evidence of the effect in Gugulethu.[59]

In addition I have a great number of households that were living in areas that *became* housing projects. These areas were being upgraded *in situ* (their entire settlement was replaced with new housing). My use of instruments is less valid for these communities, who may have lobbied for their particular area to be upgraded. In this way it could be that my instrumental variable results are picking up the effect of households that were treated because their area was being upgraded *in situ*. As a result I drop these households from the sample, leaving only households that were outside housing developments, and therefore

59 This could be cause households in Gugulethu were the household most likely to be moved further away from the city centre, since it is a relatively dense urban area where there was little space for new housing construction.

could not have been ensured access to housing. I find the results are present and just as strong for this group (in panel "Out of Project" in 2.17).

2.7.4 *Other Channels*

Government housing has an impact on household incomes through the channel of increased household earnings from wage employment. The results seem to be driven by increased probability of female labour force participation and earnings, as well as increased earnings for men who work, although this result is not robust to the use of instrumental variables.

I show that sources of income other than labour earnings are not driving the results.[60] In Table B.3 I show that receiving government housing was not correlated with getting access to other forms of government grants or welfare. Nor is housing correlated with increases in receipts of remittances or other forms of financial support from other family members or friends.

2.8 CONCLUSION

This paper provides new evidence on the relationship between household living conditions and labour supply. I find that government housing has a large and significant impact on household income, and that this effect is driven by increases in earnings from wage labour for household members. This finding is robust to my instrumental variables estimation, which uses proximity from housing projects to predict selection into housing. It is also robust to the use of cancelled projects to control for non-random location choice for housing projects.

I show that the female household members are more likely to be employed in wage paying labour after receiving government housing. I cannot conclusively ascribe the impact on labour to any one particular mechanism. I use evidence

60 We might worry that increased incomes could be driving the results *directly* by increasing reported household incomes, or indirectly if increased welfare facilities job search or increased mobility (Ardington et al., 2009; Franklin, 2015).

from time use surveys and qualitative work on housing conditions to describe the demands on daily life faced by South African households. These include a lack of to water from some distance away, a lack of security from crime, regular fires and floods which cause damage to houses, and the use of inefficient and time consuming cooking and heating technologies.

I provide evidence that government housing alleviates those constraints by improving access to running water, increasing the probability of ownership of labour saving technologies, reducing the use of dangerous fuels for lighting and heating, and improving feelings of security in the home.[61] I also find that housing increases feelings of security in the home which might make it easier for members to leave the household when they otherwise would have stayed to look after the home. I argue that these impacts are driving the impacts on female labour supply.

This finding is consistent with the view that living conditions can have an impact on the ability of females to work (Dinkelman, 2011; Field, 2007). It bolsters the case that informal settlements can act as a poverty trap to those living in them because of this restriction on labour supply (Marx et al., 2013).

Government housing in South Africa has had a transformative effect on South Africa's urban landscape, with over 25% of the total housing stock estimated to have been built by the government in the last 20 years. This perceived success has been a cornerstone of the Government's electoral platform. Governments elsewhere in the developing world seem to be increasingly enthusiastic about large scale housing projects of this kind.

Yet projects on this scale are not easy to evaluate. Randomization of such projects is unlikely to be practically or politically feasible. This paper provides an example of how projects like this can be evaluated ex post with my unique

61 Data limitations prevent me from investigating the health effects of housing, or looking in more detail at the impact of housing on time use patterns.

identification strategy of using the combined effect of proximity to urban services as instruments for selection into programmes.

This paper highlights the need to take account the effect of urban and housing policy on the labour outcomes of recipients. Ongoing research is required to fully understand the role of labour, place of living, neighbourhoods, housing design and finance in determining the efficacy of urban policy.

BIBLIOGRAPHY

Angrist, J. and Imbens, G. W. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: an empiricist's companion*. Princeton Univ Press.

Ardington, C., Case, A., and Hosegood, V. (2009). Labor supply responses to large social transfers: Longitudinal evidence from South Africa. *American economic journal. Applied economics*, 1(1):22–48.

Attanasio, O. and Vera-Hernandez, M. (2004). Medium- and long run effects of nutrition and child care: evaluation of a community nursery programme in rural Colombia.

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in china. *National Bureau of Economic Research*, pages 1–22.

Banerjee, A., Galiani, S., and Levinsohn, J. (2007). Why has unemployment risen in the new South Africa. *IPC Work Paper Series Number 35*.

Barnhardt, S., Field, E., and Pande, R. (2014). Moving to opportunity or isolation? Network effects of a slum relocation program in India. *mimeo*.

Baum, C. F. and Schaffer, M. E. (2007). Enhanced routines for instrumental variables / generalized method of moments estimation and testing. *Stata Journal*, 7(4):465–506.

Baum, C. F. and Stillman, S. (1999). {DMEXOGXT}: Stata module to test consistency of {OLS} vs {XT}-{IV} estimates. *Statistical Software Components*.

Baum-Snow, N. and Ferreira, F. (2014). Causal inference in urban and regional economics. Technical report, National Bureau of Economic Research.

Becker, G. (1965). A theory of the allocation of time. *The Economic Journal*, 75(299):493–517.

Berniell, M. and Sánchez-Páramo, C. (2012). Overview of time use data used for the analysis of gender differences in time use patterns. *Background paper for the World Development Report*.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5):1671–1748.

Budlender, D., Chobokoane, N., and Mpetsheni, Y. (2001). *A survey of time use: How South African women and men spend their time*. Statistics South Africa.

Bundy, C. (2014). *Short-changed: South Africa since Apartheid*. Jacana.

Case, A. and Deaton, A. (1998). Large cash transfers to the elderly in South Africa. *The Economic Journal*, 108(450):1330–1361.

Cattaneo, M., Galiani, S., Gertler, P. J., Martinez, S., and Titiunik, R. (2009). Housing, health, and happiness. *American Economic Journal: Economic Policy*, 1(1):75–105.

Centre on Housing Rights and Evictions (2009). N2 Gateway Project: Housing rights violations as 'development' in South Africa.

Chance, K. (2008). Housing and evictions at the N2 Gateway Project in Delft. *a report for Abahlali baseMjondolo, available online at: [http://www. abahlali. org/]*.

Charlton, S. and Kihato, C. (2006). Reaching the poor? An analysis of the influences on the evolution of South Africa's housing programme. *Democracy and Delivery: Urban policy in South Africa*, pages 252–282.

City of Cape Town (2005). Study on the social profile of residents of three selection informal settlements in Cape Town.

Collier, P. and Venables, A. J. (2013). Housing and Urbanization in Africa: unleashing a formal market process. *CSAE Working Paper WPS /2013/01*, 44:0–18.

Davidson, R. and MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press, Oxford, UK.

Davis, M. (2007). *Planet of Slums*. Verso Books.

Department of Human Settlements (2009). National Housing Code.

Devoto, F., Duflo, E., Dupas, P., Pariente, W., and Pons, V. (2011). Happiness on tap: piped water adoption in urban Morocco. Technical report, National Bureau of Economic Research.

Dinkelman, T. (2011). The Effects of Rural Electrification on Employment : New Evidence from South Africa valuable. *American Economic Review*, 101(7):3078–3108.

Duflo, E. (2003). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *The World Bank Economic Review*, 17(1):1–25.

Duflo, E. (2012). Women's empowerment and economic eevelopment. *Journal of Economic Literature*, 50(4):1051–1079.

Duflo, E., Galiani, S., and Mobarak, A. M. (2012). Improving access to urban services for the poor: open issues and a framework for a future research agenda. *J-PAL Urban Services Review Paper. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab. http://www. povertyactionlab. org/publication/improving-access-urban-services-poor*.

Duflo, E. and Pande, R. (2007). Dams. *Quarterly Journal of Economics*, 122(2):601–646.

Field, E. (2005). Property rights and investment in urban slums. *Journal of the European Economic Association*, 3(2-3):279–290.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in Peru. *Quarterly Journal of Economics*, 122(4):1561–1602.

Field, E. and Kremer, M. (2006). *Impact Evaluation for Slum Upgrading Interventions*. World Bank, Poverty Reduction and Economic Management, Thematic Group on Poverty Analysis, Monitoring and Impact Evaluation.

Franklin, S. (2015). Location, search costs and youth unemployment: A randomized trial of transport subsidies in Ethiopia. *Working Paper*.

Galiani, S., Gertler, P. J., Cooper, R., Martinez, S., Ross, A., and Undurraga, R. (2014). Shelter from the storm: Upgrading housing infrastructure in Latin American slums. *IDB Working Paper*, IDB-WP-528.

Galiani, S. and Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. Technical report, Business School Working Paper, Universidad Torcuato Di Tella, Buenos Aires.

Gilbert, A. G. (2004). Helping the poor through housing subsidies: Lessons from Chile, Columbia and South Africa. *Habitat International*, 28:13–40.

Gilbert, A. G. (2014). Free housing for the poor: An effective way to address poverty? *Habitat International*, 41:253–261.

Greenwood, J., Seshadri, A., and Yorukoglu, M. (2005). Engines of liberation. *The Review of Economic Studies*, 72(1):109–133.

Gronau, R. (1977). Leisure, Home Production, and Work–the Theory of the Allocation of Time Revisited. *Journal of Political Economy*, 85(6):1099.

Heckman, J. (1990). Varieties of selection bias. *American Economic Review*, 80(2):313–318.

Huchzermeyer, M. (2003). A legacy of control? The capital subsidy for housing, and informal settlement intervention in South Africa. *International Journal of Urban and Regional Research*, 27(3):591–612.

Huchzermeyer, M. (2006). Challenges facing people-driven development in the context of a strong, delivery-oriented state: Joe Slovo Village, Port Elizabeth. In *Urban Forum*, volume 17, pages 25–53. Springer.

Huchzermeyer, M. and Karam, A. (2006). *Informal settlements: a perpetual challenge?* Juta and Company Ltd.

Jacobs, J. M. (1961). *The Death and Life of Great American Cities*. Vintage books.

Jalan, J. and Ravallion, M. (2003). Does piped water reduce diarrhea for children in rural India? *Journal of econometrics*.

Keare, D. H. and Parris, S. (1982). Evaluation of Shelter Programs for the Urban Poor Evaluation of Shelter Programs for the Urban Poor Principal Findings.

Klonner, S. and Nolen, P. (2010). Cell Phones and Rural Labor Markets: Evidence from South Africa. Proceedings of the german development economics conference, hannover 2010.

Lall, S. V., Lundberg, M. K. a., and Shalizi, Z. (2008). Implications of alternate policies on welfare of slum dwellers: Evidence from Pune, India. *Journal of Urban Economics*, 63:56–73.

Lam, D., Seekings, J., and Sparks, M. (2006). The Cape Area Panel Study: Overview and Technical Documentation for Waves 1-2-3.

Lipman, A. (1998). Apartheid ends, but they're still put in little boxes, little boxes all the same.

Marx, B., Stoker, T., and Suri, T. (2013). The economics of slums in the developing world. *Journal of Economic Perspectives*, 27(4):187–210.

McKenzie, D. and Seynabou Sakho, Y. (2010). Does it pay firms to register for taxes? The impact of formality on firm profitability. *Journal of Development Economics*, 91(1):15–24.

Murtazashvili, I. and Wooldridge, J. M. (2008). Fixed effects instrumental variables estimation in correlated random coefficient panel data models. *Journal of Econometrics*, 142(1):539–552.

Newey, W. K. (1990). Semiparametric efficiency bounds. *Journal of applied econometrics*, 5(2):99–135.

Olsen, E. O. and Zabel, J. E. (2014). United States housing policy. In *HandBook of Regional & Urban Economics*.

Payne, G., Durand-Lasserve, A., Rakodi, C., Marx, C., Rubin, M., and Ndiaye, S. (2008). Social and economic impacts of land titling programmes in urban and peri-urban areas: International experience and case studies of Senegal and South Africa. *Oslo and Stockholm: SIDA and Norwegian Ministry of Foreign Affairs*.

Pharoah, R. (2012). Fire Risk in Informal Settlements in Cape Town, South Africa. *Disaster risk reduction: cases from urban Africa*, page 105.

Pitt, M. M., Rosenzweig, M. R., and Hassan, N. (2006). Sharing the Burden of Disease: Gender, the Household Division of Labor and the Health Effects of Indoor Air Pollution in Bangladesh and India. In *Health San Francisco*, volume 202.

Ravallion, M. and Wodon, Q. (2000). Does child labour displace schooling? Evidence on behavioural responses to an enrollment subsidy. *The Economic Journal*, 110(462):158–175.

Republic of South Africa (1994). New Housing Policy and Strategy for South Africa. Technical report.

Rodriques, E., Gie, J., and Haskins, C. (2006). Informal Dwelling Count for Cape Town (1993- 2005).

Rospabe, S. and Selod, H. (2006). Does city structure cause unemployment? The case of Cape Town. In Bhorat, H. and Kanbur, R., editors, *Poverty and Policy in Post-Apartheid South Africa*, pages 262–287. HSRC press.

Royston, L. (1998). South Africa: the struggle for access to the city in the Witwatersrand Region. *Evictions and the right to housing: experience from Canada, Chile, the Dominican Republic, South Africa, and South Korea*.

Royston, L. (2002). Security of urban tenure in South Africa: Overview of policy and practice. *Holding their Ground. Secure Land Tenure for the Urban Poor in Developing Countries*, pages 165–181.

Sachs, A. (2003). The Judicial Enforcement of Socio-Economic Rights: The Grootboom Case. *Current Legal Problems*, 56(1):579–601.

Schwebel, D. C., Swart, D., Hui, S. K. A., Simpson, J., and Hobeb, P. (2009). Paraffin-related injury in low-income South African communities: Knowledge, practice and perceived risk. *Bulletin of the World Health Organization*, 87(9):700–706.

Seekings, J., Jooste, T., Muyeba, S., Coqui, M., and Russell, M. (2010). The Social Consequences of 'Mixed' Neighbourhoods. *CSIR Working Paper*.

South Africa (2013). SA govt passes 3-million housing mark.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression, Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, pages 80–108.

Strauss, J. (1986). Does Better Nutrition Raise Farm Productivity? *Journal of Political Economy*, 94(2):297.

Strauss, J. and Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):766–817.

Tissington, K. (2011). A Resource Guide to Housing in South Africa 1994-2010: Legislation, Policy, Programs and Practice.

Tissington, K., Munshi, N., Mirugi-Mukundi, G., and Durojaye, E. (2013). Jumping the Queue: Waiting lists and other myths: Perceptions and practice around housing demand and allocation in South Africa. Technical report, Socio-economic Rights Institute of South Africa.

Tomlinson, M. R. (1998). South Africa's New Housing Policy: An Assessment of the First Two Years, 1994-96. *International Journal of Urban and Regional Research*, 22(1):137–146.

Tshangana, M. (2009). Allocation of subsidized housing opportunities to households in the Western Cape by municipalities. International Conference, Exhibition and Housing Awards, Cape Town 2009.

UN Habitat (2003). The Challenge of Slums: Global report on human settlements. *Nairobi, Kenya, UN: Habitat*.

UN Habitat (2010). *State of the World's Cities 2010/11: Bridging the Urban Divide*. Earthscan.

Vorster, J. H. and Tolken, J. E. (2008). Western Cape Occupancy Study. Technical report, Department of Sociology and Social Anthropology, Stellenbosch University.

Werlin, H. (1999). The Slum Upgrading Myth. *Urban Studies*, 36(9):1523–1534.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Zwane, A. P. and Kremer, M. (2007). What works in fighting diarrheal diseases in developing countries? A critical review. *World Bank Research Observer*, 22(1):1–24.

B

APPENDICES FOR CHAPTER 2

B.1 ADDITIONAL INFORMATION ON FIRST STAGE ESTIMATES

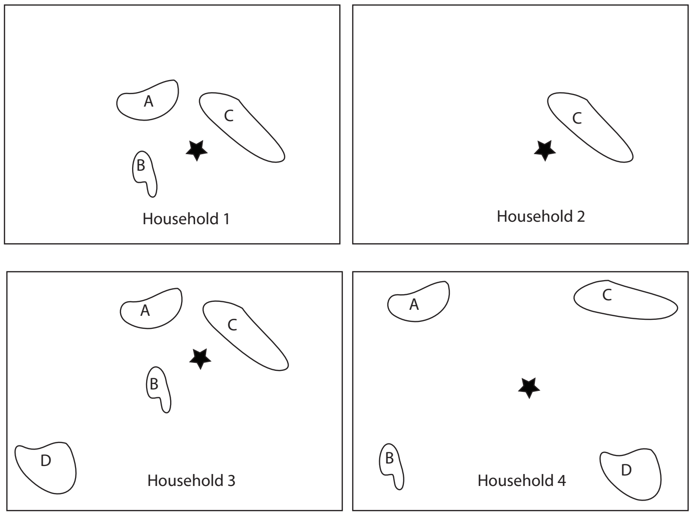
This section covers some additional information related to the estimation of the predicted values for the probability of receiving housing.

I do three things: I explain the rationale for the use of a non-linear predictor of getting housing, as opposed to other linear estimators, or estimators using a single index of housing proximity. Secondly, I explain the time-varying multinominal version of the estimator outlined, which explains how I use this framework to predict time varying treatment effects. Thirdly, I present Monte Carlo estimates that show the reliability of the method to recover the true parameters of the data generating process it describes.

B.1.1 *Why a non-linear estimator*

I estimate the probability of selection into treatment as the joint probability of being selected by a set of neighboring housing projects. Using only the closest housing project as an instrument would completely miss the effect of living in an area with a great number of projects, relative to a house who just has one, very small project nearby. Additionally, the marginal impact of additional housing projects should diminish as more are built: Consider the possible geographical scenarios depicted in figure B.1 below. We want household 1 to be more likely to be treated than 2 of course, but not three times as likely. After all, we might believe the causal effect of a project as close as C to be a 50% chance of treatment for household 2 which would lead to considerable estimation issues for Househould 1 in a linear model. Alternatively, summing distances to all projects and estimating a single coefficient would severely penalize household 3 to the benefit of household 1 in the diagram, as 3's extra project D would increase the sum of distances. However, placing a different coefficient on the distance from each project, would also be misleading- a new project's influence should be diminish-

Figure B.1: Hypothetical housing proximity scenarios



ing in the probability that a household as already been selected. For households like 3, project D is unlikely to influence is probability of treatment at the margin but for households like 4- D could make all the difference.

Trying to incorporate all of these sorts of concerns would involve a great deal of linear restrictions. Instead I estimate a non-linear model that more closely resembles the real world allocation process, and avoids the pitfalls of linear models.

I make two main assumptions in the specification of our function form: firstly, that a household's probability of selection into a individual housing project is a function of a linear combination of a number of proximity instruments, as well as household covariates. I also assume that each project allocates houses independently from the other projects, but that they all do so with the same catchment area; so that parameter on "distance from a project" is constant across households, projects and time.

B.1.2 *Time specific predictions*

The problem is complicated further when we consider that we have panel data, and need to estimate a probability of being treated for each time period $P(T_{it} =$

1) in order to generate a time varying probability of treatment. This problem is analogous to finding the joint probability across programs, except that in this case we find the probability across time. Just like one cannot be treated by two projects at one time, one cannot be treated in more than one time period, and once you are treated you stay treated for all ensuing periods. Thus the probability of not being treated at all at time t is simply the product of the probability of not being treated by any projects built in each time period $A_t \in A$, up to and including t. This can be simply rewritten as the sum of of probabilities across all projects built in the past. In this way, estimation exploits my data on the timing of housing construction to more accurately predict when households were treated.

$$
\begin{aligned}
P(T_{it} = 1) &= 1 - \prod_{j=1}^{t} \prod_{a}^{A_j} P(T_{ia} = 0) \\
&= 1 - \prod_{a}^{A_t, A_{t-1}, \ldots} P(T_{ia} = 0) \\
&= 1 - \prod_{a}^{A_t, A_{t-1}, \ldots} \Lambda(-x_i \beta - dis_{ia} \rho)
\end{aligned}
\tag{2.13}
$$

In this functional form (2.13), the probability for a household at each time is an independent event. The probability of having the treatment must be monotonically increasing with time (for each household) as more projects are built. This is the first of two specifications I use to predict treatment.

A more efficient way to estimate the probability of treatment, which takes into account the panel nature of the data, is to construct a single likelihood function for each household, which predicts *when* the household was treated. Each household is assigned a value for TD, which takes on 0 if never treated or $t = 1, 2, \ldots, Y$ if it received the treatment in period t. These are mutually exclusive outcomes. We can then estimate a model of **multinomial form**, in this case for a Y period model. At each time period we continue to use the form (2.8) in the calculations.

$$
\begin{aligned}
P(TD_i = 0) &= \prod_{t=1}^{4} \prod_{a}^{A_t} P(T_{ia} = 0)) \\
P(TD_i = 1) &= 1 - \prod_{a}^{A_1} P(T_{ia} = 0) \\
P(TD_i = 2) &= (1 - \prod_{a}^{A_2} P(T_{ia} = 0)) \prod_{a}^{A_1} P(T_{ia} = 0) \\
&\dots \\
P(TD_i = Y) &= (1 - \prod_{a}^{A_Y} P(T_{ia} = 0)) \prod_{t=1}^{Y-1} \prod_{a}^{A_t} P(T_{ia} = 0)
\end{aligned} \tag{2.14}
$$

Notice how probability of treatment is now conditioned on not having been treated earlier. For instance, dummy variable $TD_i = 2$ (got treated in period 2) is the probability of not being selected in period 1 times the probability of being selected in period 2. These dummy variables and their predicted probabilities must, by definition, must sum to 1.

This dummy indicates the the household actually got the treatment in that period. This is different to the outcome of interest, which is the probability of *having the treatment* at a given time period. This can be backed out by simply summing the predicted dummy variables for all time periods up to the present. The probability of treatment at a given time period then simplifies to the same expression given by (2.13), although the estimation procedure differs. For instance for could calculate the probability of being treated at time 2 using this framework and get the expression give by (2.13) for $t = 2$:

$$
\begin{aligned}
P(T_{i2} = 1) &= P(TD_i = 1) + P(TD_i = 2) \\
&= (1 - \prod_{a}^{A_1} P(T_{ia} = 0)) + (1 - \prod_{a}^{A_2} P(T_{ia} = 0)) \prod_{a}^{A_1} P(T_{ia} = 0) \\
&= 1 - \prod_{a}^{A_1} P(T_{ia} = 0) \prod_{a}^{A_2} P(T_{ia} = 0) \\
&= 1 - \prod_{t}^{2} \prod_{a}^{A_t} P(T_{ia} = 0)
\end{aligned}
$$

To summarize, I have two non-linear specifications for the probability of being treated at a given time, equation (2.13) and equation (2.14). In both models, probability of being treated in a certain time period depends on the projects built up until that point. While we would expect (2.14) to be the better estimator in the presence of unobserved individual heterogeneity, in the presence of measurement error, we may get less efficient results because it requires the precise time period in which the household was treated. The difference between these two types of estimators and their bias in the presence of unobservables, is explored in the Monte Carlo section below.

Marginal Effects

But first it is useful to have some marginal effects interpretation. This is slightly more complicated than a standard logit framework, but has an intuitive interpretation. We write down the probability of being treated at a particular point using the expression (2.8) and use the properties of the logistic function, to take the derivative with respect to a particular project b.[1]

$$P(T_i = 1) = 1 - \prod_{a}^{A} \Lambda(-x_i\beta - dis_{ia}\rho) \tag{2.15}$$

$$= 1 - \prod_{a}^{A} \frac{1}{1 + \exp(x_i\beta + dis_{ia}\rho)} \tag{2.16}$$

$$\frac{\partial P(T_i = 1)}{\partial dis_{ib}} = - \prod_{a \neq b}^{A} \frac{1}{1 + \exp(x_i\beta + dis_{ia}\rho)} \cdot \frac{-\exp(x_i\beta + dis_{ib}\rho)}{(1 + \exp(x_i\beta + dis_{ib}\rho))^2} \cdot \rho$$

$$= \prod_{a}^{A} \frac{1}{1 + \exp(x_i\beta + dis_{ia}\rho)} \cdot \frac{\exp(x_i\beta + dis_{ib}\rho)}{1 + \exp(x_i\beta + dis_{ib}\rho)} \cdot \rho$$

$$= P(T_i = 0) \cdot P(T_{ib} = 1) \cdot \rho$$

1 Of course, taking a partial derivative invokes the ceteris paribus assumption. Strictly speaking, this is not plausible in my case. Up until now, I have been discussing a set of distances to projects for each household, these projects will be common to a number of households. So it is hard to imagine the distance from a household to a project changing without it effecting the distance for other households, which in turn would influence the probability of a household being treated

In this framework the marginal effects of distance to particular project depend on a household's current probability of being treated (negatively) and on the probability of being by the treated by the project in question (positively). This is consistent with the idea that a new construction has a relatively bigger effect for a household with few existing projects nearby, and that the probability of being treated drops off faster the further away a particular project gets.

Using the results from the estimation of the first stage in Section 2.4.3 and the coefficients in Column (4): the coefficient on distance is 0.672, on distance squared it is 0.0079. Imagine a household close to two projects, with characteristics such that the household has a predicted probability of being selected of 10% for both of the projects. Then imagine that one project (b) was originally located 1km away but is relocated slightly further away the household. Then the probability of that household being treated would fall by over 4% for each kilometer that it moved:

$$\begin{aligned} \frac{\partial P(T_i = 1)}{\partial \mathrm{dis}_{ib}} &= P(T_i = 0) \cdot P(T_{ib} = 1) \cdot (-0.672 + 0.0158 \times \mathrm{dis}_{ib}) \\ &= (0.9 \times 0.9) \times 0.1 \times (-0.672 + 0.0158 \cdot 1) = \qquad -4.16\% \end{aligned}$$

Maximum Likelihood Estimation and Monte Carlo Simulations

In this section I estimate the parameters of the models specified in (2.13) and (2.14) using simulated data. Estimation of these models has to be performed using maximum likelihood estimation. I use the Stata **ml** code in order to maximize the log likelihood functions derived from the predicted probability of treatment given by each model, using the Newton-Raphson method.

To perform Monte Carlo tests, I simulate a dataset of $N = 1000$ observations with 3 time periods each. Then for each time household-time observation I generate 5 random project distances (to simulate the construction of houses nearby that household). Each household has a randomly generated household effect x_i that is constant across time and assumed to be unobserved. In addition, each time period has a random effect on the probability of treatment, common to everyone. Then, for each project at each time a latent variable is generated as a function of time effects, the household fixed effects and, of course, the distance from the household to the project. I use a linearly added logistic error term. If this latent variable is greater than zero we consider a household to be "treated" by that project. A household is treated at that time if it is treated by any *one* of the projects, and it remains treated for the ensuing periods.

$$y^{\star}_{iat} = \alpha + \sigma_x x_i + \sigma_\lambda \lambda_t + \rho dis_{iat} + \epsilon_{it}$$

$$y_{iat} = \mathbf{1}[y^{\star}_{iat} > 0]$$

$$y_{it} = \mathbf{1}[\sum_{a}^{A_t, A_{t-1}\dots} y_{iat} > 0]$$

Where the household characteristics and distance variables are generated in the following way:

$$x_i \sim N[0,1], \lambda_t \sim N[0,1]$$

$$dis_{iat} \sim U[1,10]$$

$$\epsilon_{it} = \frac{\exp(\eta_{it})}{1+\exp(\eta_{it})}, \eta_{it} \sim U[0,1]$$

Having generated simulated values, I recover the parameter of interest, which is ρ, using the models specified. I estimate three different specifications. The first (L_{nt}) uses the functional form (2.13) but without any attempts to control for time trends. The second (L_t) also uses (2.13), but controls for time by specifying time dummies λ_t in the latent $y^\star$ form. The third (MNL) is the estimation of (2.14). I then perform Monte Carlo simulations with 1000 repetitions, for each model, while varying the magnitude of variance of the unobserved effects. The results of these simulations, with different "true" values of ρ, are given in table B.1. The

Table B.1: Results of Monte Carlo Simulations: Estimated value of ρ with different unobserved fixed and time effects

	$\rho = -1$			$\rho = -0.5$		
	L_{nt}	L_t	MNL	L_{nt}	L_t	MNL
$\sigma_x^2 = 0, \sigma_\lambda^2 = 0$	-1.013 (0.148)	-1.013 (0.148)	-1.019 (0.137)	-0.507 (0.078)	-0.507 (0.078)	-0.513 (0.067)
$\sigma_x^2 = 1, \sigma_\lambda^2 = 0$	-0.914 (0.125)	-0.915 (0.125)	-0.950 (0.109)	-0.433 (0.067)	-0.433 (0.067)	-0.464 (0.059)
$\sigma_x^2 = 0, \sigma_\lambda^2 = 1$	-0.997 (0.144)	-1.003 (0.143)	-1.007 (0.143)	-0.494 (0.076)	-0.500 (0.075)	-0.503 (0.076)
$\sigma_x^2 = 3, \sigma_\lambda^2 = 0$	-0.742 (0.099)	-0.744 (0.099)	-0.821 (0.089)	-0.330 (0.061)	-0.332 (0.061)	-0.383 (0.055)
$\sigma_x^2 = 0, \sigma_\lambda^2 = 3$	-0.836 (0.244)	-0.904 (0.175)	-0.928 (0.240)	-0.416 (0.138)	-0.481 (0.082)	-0.453 (0.138)

(N=1000, t=3, R=1000)

model performs very well without any fixed effects or time trends, as expected. The introduction of fixed effect biases the estimates towards zero. The bias can be quite considerable when these fixed effects are relatively large, as the example with $\sigma_x^2 = 3$ indicates. The effects are less severe with the introduction of unobserved time effects, but still biased towards zero. The MNL estimator performs better when there fixed effects. Importantly, the inclusion of time controls in the L model does a very good job of recovering the parameters correctly.

Table B.2: Household characteristics in first & last waves, by treatment

	Wave 1 (2002)			**Wave 4 (2009)**		
	Control	Treatment	Diff	Control	Treatment	Diff
From EC	0.712	0.827	0.114***			
			(0.0277)			
Backyard	0.122	0.0742	-0.0478*			
			(0.0198)			
Coloured	0.187	0.0797	-0.107***			
			(0.0228)			
Black	0.790	0.915	0.125***			
			(0.0238)			
Migrant	0.658	0.657	-0.00119			
			(0.0306)			
Shack	1	1	0	0.624	0.188	-0.436***
			(0)			(0.0328)
City distance	22.68	25.12	2.434***	22.51	25.19	2.681***
			(0.394)			(0.436)
Years Ed. head	11.04	11.38	0.341	12.22	12.46	0.238
			(0.214)			(0.239)
Num. Rooms	3.123	3.259	0.136	3.427	3.828	0.401**
			(0.0975)			(0.122)
HH Size	5.189	5.478	0.289	5.549	6.065	0.516*
			(0.149)			(0.221)
Female Head	0.488	0.563	0.0751*	0.533	0.597	0.0649
			(0.0321)			(0.0361)
Age Head	41.66	42.51	0.854	43.95	46.08	2.129*
			(0.758)			(0.947)
Young adult employed	0.112	0.0769	-0.0353	0.465	0.460	-0.00472
			(0.0193)			(0.0360)
% Females Employed	0.442	0.447	0.00439	0.367	0.403	0.0353
			(0.0289)			(0.0277)
Head Employed	0.691	0.632	-0.0596*	0.618	0.562	-0.0554
			(0.0302)			(0.0354)
Health Score	3.858	3.879	0.0214	3.790	4.022	0.233*
			(0.0865)			(0.0903)
Piped Water	0.130	0.107	-0.0233	0.348	0.534	0.185***
			(0.0211)			(0.0351)
Earnings pc	874.9	620.21	-254.75	330.95	400.01	69.06
			(254.74)			(46.177)
Log Income	7.436	7.218	-0.217***	8.194	8.272	0.0784
			(0.0606)			(0.0635)
Obs	713	364		626	344	

Figure B.2: Map of Cape Town and its housing projects in 2009

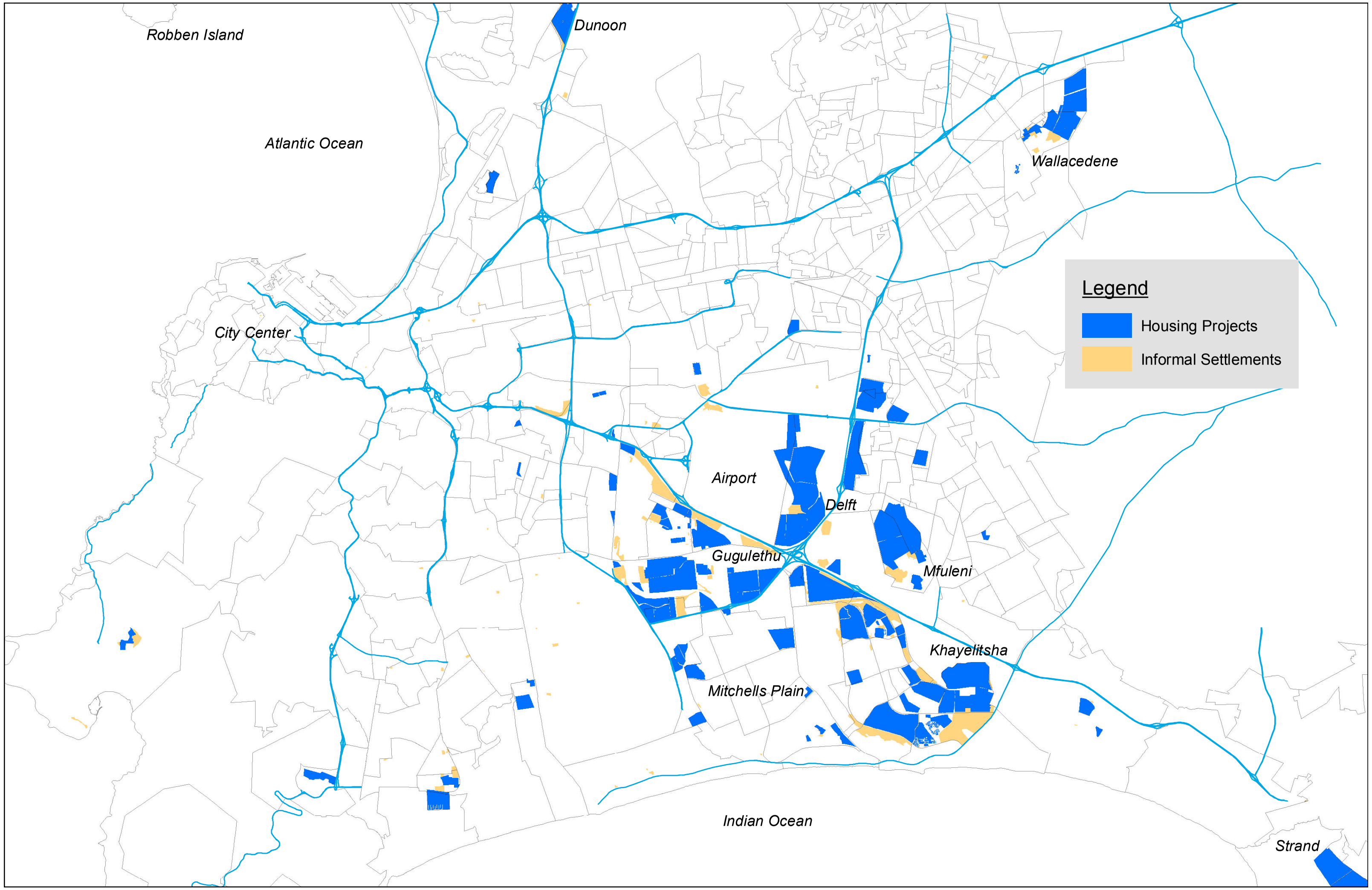


Table B.3: Impact of treatment on household grants and remittances

	(1)	(2)	(3)	(4)
	FE	FE	FE	FE
	hhgrants	hhgrants	totalsend	totalrec
house	-0.00495	-0.0245	-484.3	-80.95
	(0.0323)	(0.0303)	(629.7)	(197.3)
Observations	3,731	3,723	1,854	1,854
R-squared	0.009	0.140	0.006	0.029
Number of hhs	1,077	1,077	1,062	1,062
HH Chars	No	Yes	Yes	Yes
Av Group	3.464	3.457	1.746	1.746

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. All regressions include controls for time-varying household characteristics. Totalsend and totalrec refer to the total amount of remittances sent and received. hhgrants is dummy for whether households received any one of the household grants such as disability benefits, or the childcare grant.

Table B.4: Example of first stage from 2SLS with single fitted instrument

Variable	Coefficient		(Std. Err.)
g-hat	0.885	**	(0.088)
femalehead	-0.070	**	(0.021)
hhsize	-0.009		(0.006)
sexratio	0.097	***	(0.051)
youngratio	0.008		(0.046)
femadultcount	0.002		(0.015)
time2	0.024		(0.020)
time3	0.024		(0.025)
time4	0.037		(0.034)
citydis	0.008	*	(0.004)
maxhhed	0.001		(0.003)
hhmaxage	0.002	**	(0.001)
hhgrants	-0.027	***	(0.016)
Intercept	-0.267	*	(0.122)
N	3711		
R^2	0.318		
$F_{(12,158)}$	28.104		

Notes: These are results from the first stage of the household fixed effects regressions used throughout this paper. These results are basic OLS regression of the dummy variable for having government housing on time varying household characteristics, as well as g-hat, the predicted probability of receiving housing from the maximum likelihood estimator of the probability of getting housing. Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: Impacts on young adult labour supply at the extensive margin

	(1) FE num employed	(2) FE % employed	(3) FD % employed	(4) IV num employed	(5) IV % employed
subhere	0.0338	0.0338	0.0346	0.250**	0.245**
	(0.0311)	(0.0311)	(0.0446)	(0.109)	(0.112)
Observations	2,648	2,648	1,324	2,623	2,622
R-squared	0.151	0.151	0.307	0.135	0.137
Number of personid	662	662	662	662	662
HH Chars	No	No	No	Yes	Yes
Av Group	4	4	2	3.962	3.961
Weak IV F	.	.	.	97.08	87.81

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. Dependent variable "num employed" is the number of young adults employed in the household (regressions include controls for the number of young adults in the household). % employed is the proportion of young adults of household members who are employed.

Table B.6: Replication of key IV results with basic (FE) 2SLS

	(1) SLSa log income	(2) SLSa head employed	(3) SLSa YA work	(4) SLSb log income	(5) SLSb head employed	(6) SLSb YA work
house	1.099***	0.452**	0.166	1.024***	0.430***	0.206**
	(0.367)	(0.202)	(0.145)	(0.183)	(0.0725)	(0.0867)
Instruments	9	9	9	14	14	14
HH Chars	Yes	Yes	Yes	Yes	Yes	Yes
Obs	3,348	3,487	3,486	3,348	3,487	3,486
R^2	0.198	-0.073	0.157	0.211	-0.065	0.151
GMM2S	Yes	Yes	Yes	Yes	Yes	Yes
Av Group	3.528	3.651	3.650	3.528	3.651	3.650
WeakIVF	4.756	5.282	5.281	23.21	21.23	21.22
Stock Yogo Critical Values						
5% Max. Bias	20.74	20.74	20.74	21.23	21.23	21.23
20% Max. Bias	6.61	6.61	6.61	6.42	6.42	6.42

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. Excluding Instruments in all regressions: projdis#1-5, rank3#1-3, inproject.

Table B.7: Main FE results only among individuals in clusters where many households were treated

	(1) FE lginc	(2) FE lginc	(3) FE lginc	(4) FE ln_hhtotsal	(5) FE ln_hhtotsal	(6) FE ln_hhtotsal
house	0.226***	0.150**	0.144**	0.193***	0.181**	0.176**
	(0.0609)	(0.0588)	(0.0582)	(0.0656)	(0.0722)	(0.0719)
femalehead		-0.297***	-0.239***		-0.221***	-0.203***
		(0.0566)	(0.0596)		(0.0730)	(0.0765)
hhsize		0.133***	0.143***		0.0441***	0.0418***
		(0.0102)	(0.0110)		(0.0124)	(0.0136)
citydis		-0.0118	-0.00929		-0.0216*	-0.0197
		(0.0110)	(0.0108)		(0.0130)	(0.0130)
yamom			-0.0145			0.0753
			(0.0682)			(0.0900)
maxhhed			0.0315***			0.0396***
			(0.00718)			(0.00911)
sexratio			-0.125			-0.184
			(0.142)			(0.184)
youngratio			-0.623***			-0.273*
			(0.112)			(0.143)
Observations	2,388	2,397	2,393	1,895	1,895	1,892
R-squared	0.216	0.301	0.325	0.129	0.147	0.165
Number of hhs	747	758	758	718	718	718
Av Group Size	3.197	3.162	3.157	2.639	2.639	2.635

Notes: Clustered standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1. house=1 if household reported getting a subsidized house at any point in the past. Sample restricted to households in EAs in which more than 20% of hhs received housing by 2009.

3

ECONOMIC SHOCKS AND LABOUR MARKET FLEXIBILITY

Financial support from the CSAE and Oxford Economic Papers Fund to purchase the data is gratefully acknowledged. We thank seminar participants in Oxford for comments and input. In particular we thank Marcel Fafchamps, Stefano Cario and Clement Imbert for useful suggestions.

Abstract:

The paper explores how labor markets adjust to large but temporary economic shocks in a context where those shocks are common. We leverage geo-referenced data on the path of typhoons in the Philippines to estimate the impacts of typhoons occurrence on employment and wage income. Using a balanced panel of about 1,100 municipalities over 26 quarters, we find short-run negative effects of the shocks on average weekly income but no effects on employment levels. We then take advantage of a large repeated cross-section of working age individuals and an individual-level panel to establish clear evidence of downward flexibility of both hours worked and hourly wages. We can rule out that results are driven by either changes in labour supply, sample composition or by individuals switching to low-paying jobs. These results hold in formal, wage-paying employing jobs, suggesting significant downward flexibility built into long-term employment agreements that insure workers from layoffs during times of economic shock.

3.1 INTRODUCTION

How do labor markets adjust to large economic shocks? A long literature has looked at the response of wages and employment to productivity shocks. A common feature of the response is that both the increase in unemployment and its duration are greater than one would expect if labor markets cleared (DeLong and Summers, 1986). This has often been blamed on nominal wage rigidities, with wages failing to adjust downward after shocks. The rigidities can have large negative welfare consequence, especially in developing countries where social safety nets are less likely to be in place.

Testing for the existence of downward nominal wage rigidity is challenging. Few studies have been able to account for issues related to aggregation bias due to changes in the composition of job types or the workforce which might accompany shocks (Keane et al., 1988; Bils, 1985). Relatively few studies use plausibly exogenous labour demand shocks (for which there is sufficient variation over time and space). Few researchers have access to high frequency data to track the effects of shocks through time. Finally, evidence from developing country is limited, especially outside of agricultural labor markets.

We overcome these challenges by leveraging a unique series of nationally representative labor force surveys in the Philippines, covering more than 3.4 million individuals in 1,100 municipalities over 26 quarters between 2003 and 2009. A number of individuals were interviewed more than once and we build a panel of 1.8 million individuals over the period. We combine this data with geo-referenced data on the path and strength of typhoons over the same period. Controlling for time and municipality fixed effects, we take advantage of the arguably exogenous nature of typhoon occurrence to estimate how labor markets adjust to large, but temporary, labor demand shock.

First, we use the municipal-level data to estimate the overall impacts of large storms on economic activity. We find that large storms do not affect employment rates but lead to a seven percent reduction in per capita wage income. The impact on incomes appears to be driven by both a reduction in the average number of hours worked and a reduction in average hourly wage. Those impacts are short-lived as the estimated effects are no longer significant after one quarter. Overall, this first set of results confirm that large storms act as short-lived labor demand shocks.

Second, we take advantage of the individual-level data to establish that nominal wages exhibit downward flexibility when storms hit. Since we are interested in total wages that firms pay workers, our preferred measure is weekly wage income, as this is the highest level of aggregation we can use. We find large and significant negative impacts on average weekly wages, while confirming that there are no impacts on employment rates. The impacts on weekly wages can be decomposed into reductions in the number of hours that workers work per week, and by reductions in average hourly wages. The effect is driven by reductions in days, and hours per day. The adjustment in hours per worker are not due to some workers taking zero hours of work, or temporary lay-offs (Feldstein, 1976). We find no evidence of labour market failures: labour markets seem to clear in times of shock, with no impact on rates of employment, unemployment, labour force participation or demand for additional labour hours.

Third, we confirm the existence of downward nominal wage flexibility by showing that shocks do not appear to systematically affect the composition of individuals in the sample, the composition of individuals that are employed and the composition of individuals who report a wage. We further restrict our analysis to the panel of individuals that we observe in employment in at least two periods. We find that, even on this restricted subsample, individuals are no less likely to be employed but that, conditional on working, wages are lower during quarters when storms hit. The results on wages are robust to further restricting the panel to individuals that are employed in similar jobs and on similar contracts across the sample period. Those results allow us to rule out that the evidence for downward nominal wage flexibility is driven by changes in sample composition, or in the composition of job types or in employment contracts of the labour force. It is important to note that here we restrict the sample to individuals who did not migrate as a result of the shock.

Fourth, we explore mechanisms behind the effects. We show that the results are strongest, and exhibit the clearest evidence of downward flexibility in non-agricultural private sector wage paying jobs where the jobs are permanent. The results don't seem to be driven by jobs that are governed by spot markets which we interpret as wage flexibility within jobs with longer term relationships which are likely to be governed by implicit contracts.

We develop an implicit contract model (Azariadis, 1975; Baily, 1974; Rosen, 1985; Miyazaki and Neary, 1985) to explain our results. Firms and workers engage in risk-sharing in the event of large demand shocks. Workers accept cuts in total wages when shocks hit, while firms insure them against the risk of layoffs, which would leave them with no income in a time of great need. Optimal contracts can adjust to large shocks through lay-offs or work-sharing (reductions in hours per worker). Firms will rely more on the latter when workers are more risk averse or face worse outside options, or when labour is relatively divisible (Mortensen, 1978).[1] Under certain conditions, even when labour is relatively indivisible, no lay-offs occur. We argue that it is plausible that these conditions could hold in the context of the Philippine labour market.

Finally, we find some evidence that managers experience a sizeable increase in their weekly wages that appears to be driven by a large increase in the numbers

1 In other words the marginal productivity of the number of hours used in production is equal to the marginal productivity of increasing the number of employees. When labour is indivisible, it is more costly for firms to cut hours than it is for them to lay-off workers.

of hours worked. This suggests a skill bias of large economic shocks. We speculate that these results are driven by the need for managerial oversight during times of crisis, as firms shift priorities away from usual business to recovering assets, dealing with storm damage, and otherwise adjusting to shocks.

Our results have a number of implications for the literature. First, we contribute to a growing literature on the impacts of large natural disasters, particularly those driven by climate change and weather (Dell et al., 2014). Our results suggest that large storms have large impacts on total output in the short run. Our estimate of total aggregate income loss is 7% for municipalities that are effected. Yet, contrary to a large literature that shows persistent distortions (Kaur, 2014) or even bounce back-better recovery after large shocks (Skidmore and Toya, 2002; Gignoux and Menéndez, 2014) we find little evidence of persistence of these effects. This is perhaps a result of adaptive mechanisms developed in the labour market, which we discuss in detail in this paper.[2]

Second, we contribute to the literature on the identification of wage flexibility during economic shocks.[3] We provide some of the only evidence for developing countries, outside of a rural context, on the question of downward wage flexibility. Importantly, we overcome many of the econometric challenges associated with identifying flexibility or rigidity of wages. Abraham and Haltiwanger (1995) argue that most of the evidence on the relatively dampened responses of wages to economic shocks is simply due to aggregation bias, possibly because lower paid workers are more likely to loose their jobs. Following Kaur (2014) we use arguably exogenous shocks to identify changes in labour demand. Panel

2 Our findings do not estimate the impact of storms on growth trajectories or other long term outcomes, because of our use of municipal fixed effects, time fixed effects, and quarterly data. Our results without municipal fixed effects show suggest that municipalities are hit regularly, look a lot poorer than areas that are not hit (although these findings are not necessarily causal). Therefore our findings do not conflict with the growing body of evidence showing that natural disasters have long term consequences for economic growth and household well-being (Anttila-Hughes and Hsiang, 2013; Hsiang and Jina, 2014).

3 See Holzer and Montgomery (1993) and Abraham and Haltiwanger (1995) for detailed overviews of this literature.

data allows us to guard against sample composition changes leading to bias in our results. We find no impact on employment, either for aggregate levels, or in sectoral composition of jobs. This is an improvement over the existing literature which often had to rely on large shifts in the employment rate to identify these shocks.

Third, we contribute to a literature on the effects of implicit contracts on labour market adjustments.[4] Our evidence suggests that contracts do not perfectly insure workers against shocks, yet risk sharing does not lead to lay-offs.[5] We provide new theoretically motivated evidence on the relative role of work-sharing and lay-offs in labour adjustments in long-term labour relationships. Both the theoretical and empirical literature has focused on long term labour contracts as a source of inflexibility in labour markets (Azariadis and Stiglitz, 1983; Holmstrom, 1983; Shimer, 2005; Beaudry and Dinardo, 1991; Hall and Milgrom, 2008). Yet we find evidence that downward wage flexibility is strongest among individuals in long-term, formal sector wage-paying jobs. This suggests that long term relationships can allow for more flexibility, rather than less. We find little evidence that workers are laid off from these jobs during shocks.

We argue that the conditions of labour markets when typhoons hit are conducive to an outcome where implicit contracts ensure full employment. Instead wages and hours of employment at the intensive margin are reduced.[6] If workers face poor outside options, and indemnity insurance for unemployment is not possible or does not exist, workers strongly prefer (average) wage cuts to unemployment risk. If labour is relatively indivisible (the marginal returns to hours

4 We think this is a setting where risk sharing contracts are particularly feasible: since storms are easily observed and verified by both sides of the market, they should be easy to implicitly contract upon. This could happen through long established norms about employers and employees expect each other to react when storms hit.

5 Models of implicit contracts with full insurance predict that workers' wages do not fall when productivity or demand shocks hit. We have further evidence that self-employed individuals in our sample are likely to go out of business when storms hit. This could be partly because these smaller business are a less able to share risk with employees, as larger firms are able to do.

6 This is contrary to the evidence from OECD countries where the most fluctuations in total hours worked are accounted for by changes in rates of employment (Rogerson and Shimer, 2011).

worked per worker are relatively low at equilibrium weekly hours) the optimal contract involves reductions in working hours rather than lay-offs.[7]

In longer term wage relationships, workers take cuts in their total wages, perhaps in return for other benefits, such as higher wages in normal periods. Hourly wages fall in this case. In labour relationships characterized by shorter contracts, hours adjust downwards more, perhaps reflecting that workers need to be compensated for lower wages, so that the hourly wage rate does not move by much. These patterns are confirmed in the hourly wage data. Further, we find that workers are less likely to be paid when storms hit, but more likely to be paid in the quarters after storms hit. This is consistent with a story where, in firms where wages are paid irregularly, workers in long term contracts accept not being paid when firms are in hard times, in return for more regular pay later on.

We argue that our results cannot be driven primarily by shifts in labour supply. Labour supply elasticity does play a role in the optimal implicit contracts. The no lay-offs contracts we observe are consistent with a setting where labour supply is inelastic at the extensive margin, because workers desperately need their pay-checks in the absence of unemployment insurance or good outside options. However, elasticity at the intensive margin is relatively high, at least for workers already working long hours. The destruction caused by storms to homes and farms requires time to rebuild (Anttila-Hughes and Hsiang, 2013) and reduces income from non-wage sources. Therefore, we speculate that workers may simultaneously have greater need for income and for time off work when storms hit. We do not find any evidence that labour supply increases when storms hit, as

7 Rosen (1985) points out that this outcome is not efficient in terms of production, relative to a case where workers can be fully insured by firms or have unemployment benefits. Employment is socially excessive (in the sense that it's marginal product is exceeded by the outside earnings options of workers) because of risk averse workers.

has been found for farming households that use wage labour markets as a way to smooth income in bad times (Jayachandran, 2006; Kochar, 1999).[8]

The remainder of the paper is organized as follows. Section 2 discusses the context and data. Section 3 establishes that shocks have large but temporary negative effects on labor markets. Section 4 presents evidence of substantial downward nominal wage flexibility. Section 5 explores mechanisms. In Section 6 we develop a model to explain why wages and hours fall, but employment does not. Section 7 concludes.

3.2 CONTEXT AND DATA

In this Section we describe the context and argue that the Philippines are an ideal setting for our analysis. Typhoons are a regular occurrence in the Philippines and generate large welfare costs (Anttila-Hughes and Hsiang, 2013; Bankoff, 2002; Ugaz and Zanolini, 2011).

3.2.1 *Typhoons in the Philippines*

We leverage data from the Japan Meteorological Agency Tropical Cyclone Database to generate quarter*municipality specific measures of storm exposure. The database provides information on each tropical storm passing through the North-West Pacific Ocean from 2000 to 2010.[9] The data takes the form of geo-referenced observations at six hour intervals of the storms lifespan, including pressure readings, and maximum windspeeds for the storm at each point.

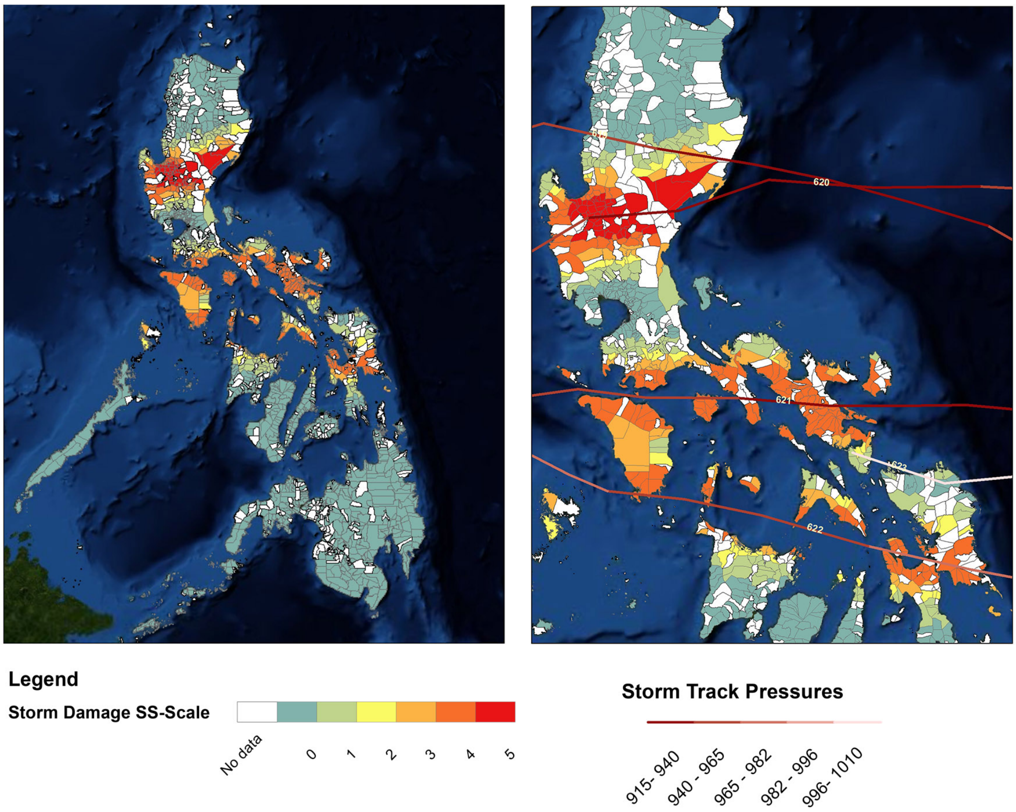
The process involves three main steps. First, for each storm, we compute the maximum wind speed that affected the municipality. To do so, we start by generating best fit lines through the six-hourly observations to mimic the storm

8 This difference is likely explained by (i) the nature of shocks in our sample, which are not only agricultural and thus effect labour demand in the wage sector, and (ii) the fact that typhoons cause the kind of catastrophic damage that requires homes to be rebuilt.

9 These data can be accessed online at http://www.jma.go.jp/en/typh/ Last accessed on December 1, 2012.

path. Then for each municipality we calculate the distance to every storm in the dataset, recover the storm track point to which it is closest, and the corresponding storm pressure (in hPa) at the moment at which the storm passed over the municipality. We apply a model of wind-speed decay from the center of the storm to estimate wind speeds for each municipality-storm combination (Holland, 1980).[10] The model uses distance from the eye of storm and the pressure at the eye to calculate a wind speed at any point.

Figure 3.1: Storm damage by municipality (Sept-Dec 2006)



Second, using the time-storm storm data data, we assign the wind speed readings during a storm to one of the three month periods preceding each of the 26 rounds of employment data described below. For instance, we have surveys from October for every year from 2003 to 2009. If a storm passed during the months

10 Many wind speeds generated in this way are negligibly small and can be safely dropped because the storm passed too far from the municipality to register an impact. We ignored all storms not registering on the Saffir-Simpson scale; that is those not reaching wind-speeds above 60 knots.

of August, September, or October, it would be assigned to the time period in the LFS data corresponding to the survey taken in October that year.

Table 3.1: Average municipality storm measures across all quarters (2003-2009)

Measure	Obs	Mean	Std. Dev.	Max
Max Windspeed	21064	13.146	29.819	142.96
Standardized Storm Measure	21064	.0320	.111	1
Any wind detected	21064	18.57%	38.88%	
Storm Registered on SS-Scale	21064	10.60%	30.78%	
SS class-0	21064	7.97%	27.08%	
SS class-1	21064	4.60%	20.97%	
SS class-2	21064	2.15%	14.50%	
SS class-3	21064	2.02%	14.0%	
SS class-4	21064	1.68%	12.87%	
SS class-5	21064	.012%	3.57%	
Big Storms (SS class-4&5)	21064	1.81%	13.34%	
Small Storms (SS class-1, 2&3)	21064	8.7%	28.31%	

Third, we aggregate the measures across the three month time periods. For each municipality and for each three month time period we take the maximum typhoon wind that the municipality was exposed to. These wind data can then be used to generate various measures of storm intensity by time period according to the Saffir-Simpson classification. This scale classifies Hurricane Wind Speeds into 5 categories, by the types of damage they will cause. Our main regressions will distinguish between Category 1-3 and Category 4-5 storms. Both of the two top category storms are said to cause Catastrophic Damage.[11] According to NOAA, it is expected that after a category 5 storm 'a high percentage of framed homes will be destroyed, with total roof failure and wall collapse. Fallen trees and power poles will isolate residential areas. Power outages will last for weeks to possibly months. Most of the area will be uninhabitable for weeks or months.'

Table 3.1 gives some indication of the damage caused by the storms in our sample using this system, looking at averages across all municipalities and all time

11 The latest version of Saffir-Simpson hurricane classifications is outlined by the National Oceanic and Atmospheric Administration's (NOAA) National Hurricane Center, available online at http://www.nhc.noaa.gov/aboutsshws.php

periods. The biggest wind speed experienced was 143 knots. On average 18.57 percent of the municipality-quarter observations are affected by a tropical storm, but 39.2 percent of those are too small to be classified on the Saffir-Simpson Scale. Across the country, 23 of the 26 quarters for which we have employment data experienced storms. Twenty of those registered storms on the Saffir-Simpson Scale, and 8 of those quarters were classified as Catastrophically Damaging (scale 4 & 5). Just less than two percent our municipal*quarter observations experienced very large storms (Saffir-Simpson classification 4 or 5).

The most active Typhoon season over the sample period was the period September-December 2006. The variation in damage across municipality storm experiences at this point in time are revealing. This is not the season of the largest storm, but damage is fairly widespread due to a number of storms: Eighteen percent of municipalities experience Catastophic Damage and 30 percent have some experience of typhoons. The geographical variability is plotted on the maps below with the municipalities coloured according to the Saffir-Simpson score of the biggest storm passing through during the quarter (blue is no damage, darker red the most catastrophic storms). Figure 3.1 (left panel) gives some impression of the damage across the entire country. Few storms passed through the South, although there are storms that did in other years. The right panel in Figure 3.1 gives a more localized picture, with the actual plots of the five typhoons that passed through that area in this time period with darker red storms having lower pressure cells. Storm Chebi (620) clearly registers the greatest damage as it passed through the centre of Luzon, while Storm Durian (621) reached the southern shores of Luzon.[12] Storm Durian is reported to have killed an estimated 720 people in the Philippines.

12 Our data contains names from the Japan Meteorological Agency Tropical Cyclone Database. The Philippine Atmospheric, Geophysical and Astronomical Services Administration names for the storms.

3.2.2 *Employment data*

We use data from Labor Force Surveys (LFS) collected by the National Statistics Office (NSO) of the Philippines. The surveys are implemented four times a year (January, April, July and October) and we have access to all 26 surveys in the period July 2003 - October 2009.[13] Data from the surveys are used to compute official employment statistics. We only use working-age individuals (above 15 year old) and are left with 3.4 million observations.

We use the dataset in three ways. First, we aggregate the individual-level data build a balanced panel of about 1,140 cities and municipalities across the 26 quarters. Second, we use the repeated cross-section of individuals. Third, we extract a panel of individuals from the cross-section. Indeed, the NSO used the same sampling frame over the period and to minimise sampling error across years, common samples were used in consecutive years. As a result, a number of households were interviewed more than once. We have access to the household IDs allowing to track households through time. We then use information on gender, age and education level to build a panel of individuals.

A person is considered employed if s/he reported at work for at least an hour during the week prior to the survey. In addition, information is collected on the total number of hours worked during the past week, the sector of employment and the daily wage. As discussed in Labonne (2014), the definition of the economically active population changed in April 2005 and it is not possible to compute employment rate as a share of the economically active population consistently across survey waves. The information required to adjust past series is not available. However, the definition of employment has not changed and we compute the employment ratio as a share of the working age population rather as a share of the economically active population.

13 More information on the survey design is available at: http://www.census.gov.ph/data/technotes/notelfs_new.html visited on March 26, 2012.

Table 3.2: Descriptive statistics: Individual data

Variable	Full Sample		Panel	
	Mean	Std. Dev.	Mean	Std. Dev.
Income per capita (PHP)	358.1	(768.9)	353.1	(762.0)
Average Wage (PHP)	1319	(954.1)	1318	(946.5)
Hours per worker	40.80	(19.4)	40.10	(19.2)
Employed	58.10%	(49.3)	61.1%	(48.7)
Unemployed	5.60 %	(23.0)	5.10%	(21.9)
No schooling	2.30%	(14.8)	2.30%	(15.1)
Some primary	14.4%	(35.1)	15.4%	(36.1)
Primary graduate	14.9%	(35.6)	15.8 %	(36.5)
Some secondary	17.3%	(37.8)	16.1%	(36.7)
Secondary graduate	24.2%	(42.8)	23.9%	(42.7)
Some college	26.9%	(44.4)	26.4%	(44.1)
Female	50%	(50)	50%	(50)
Age	35.80	(16.3)	37.40	(15.9)
Composition of jobs				
Wage employment	51.50%	(50.0)	48.70%	(50.0)
Agriculture	35%	(47.7)	37.70%	(48.5)
Key Job Types				
Own farm	26.3%	(44.0)	28.7%	(45.3)
Wage farm	8.60%	(28.1)	8.90%	(28.5)
Self employed	22.2%	(41.6)	22.6%	(41.8)
Government	7.60%	(26.5)	8.10%	(27.2)
Private permanent	26.2%	(44.0)	23.6%	(42.4)
Private temporary	9.0%	(28.6)	8.10%	(27.3)
	(N=3,402,456)		(N=1,000,687)	

For the purposes of analysis in this paper, and to isolate differential treatment effects by employment type, we use a basic typology of the jobs available in the Philippine labour market. These are:

Own farm: If these jobs are paid, which they rarely are, they are paid on a daily, commission, or pakyaw basis.[14] This work is mostly subsistence agriculture as classified as self employment or unpaid family work. Wages are rarely

14 According to the Republic of the Philippines Government Procurement Policy Board, "Pakyaw refers to a system of hiring a labor group for the performance of a specific work and/or service incidental to the implementation of an infrastructure project by administration whereby tools and materials are furnished by the implementing agency. For the specific work/service output, a lump-sum payment is made either through the group leader or divided among the pakyaw workers and disbursed using a payroll system" (GPPB Resolution No. 18-2006, December 6, 206).

observed for these jobs, and so these workers do not influence the estimates on aggregate wages.

Wage farm: This is wage employment on a farm other the household's own. These jobs are usually paid on a daily basis.

Self employment: These are mostly very small retail or small scale construction enterprises. This category excludes those who define themselves as self-employed agriculturists. Wages were rarely observed for this category. These workers, too, do not influence or analysis of aggregate wages.

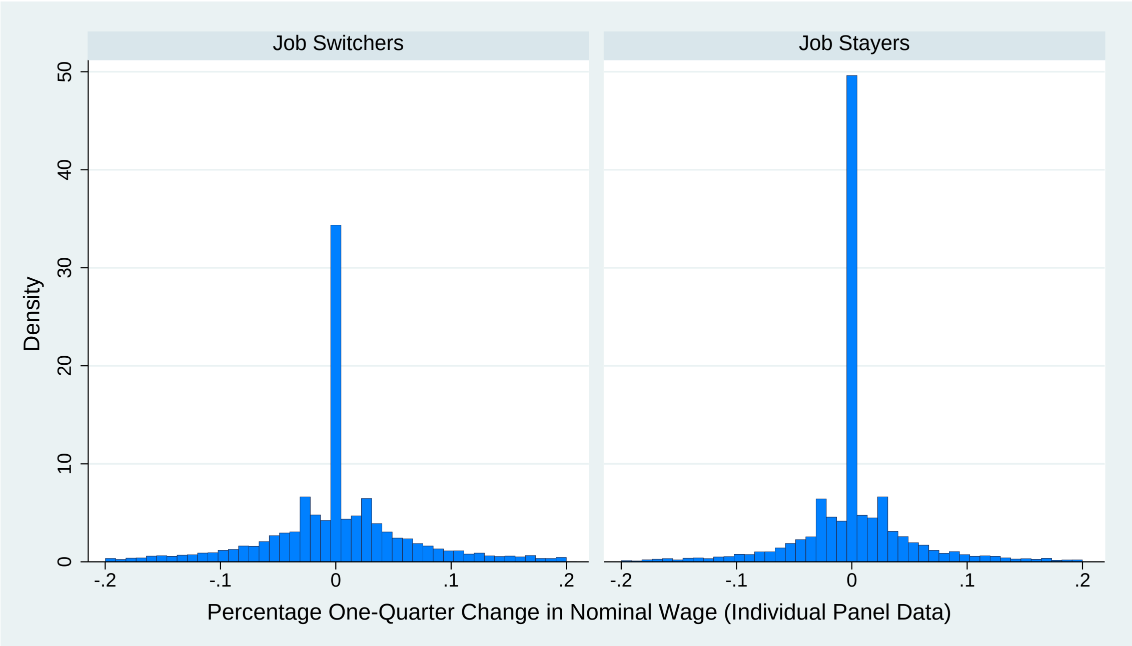
Government work: Formal wage work in the public sector, usually paid monthly. Most of these jobs are permanent.

Permanent private sector wage employment: These are jobs where the respondent considers the job permanent. Wages are usually paid on a monthly basis. Daily wages are also common. These are jobs that are most likely to based on longer term relationships and contracts, and are the focus of much of the analysis of the paper. These are most often formal sector jobs.

Temporary private sector wage employment: These are job at private establishments where workers identified the jobs as "short term". This includes casual labour, seasonal work, and short term contracts. The most common mode of payment was a daily wage, although piece-rate and pakyaw payments are more common than for permanent jobs.

Table 3.2 shows the composition of these different jobs in the full individual sample and in the panel. Roughly a third of employed individuals are self-employed (if own-farm workers are included as self-employed), a little more than a third are employed by private employers. The public sector makes up about eight percent of employment. The rest is made up of unpaid family work, which is mostly in agriculture, and domestic work. About half of self-employment is in agriculture, mostly labour on the households own farm with produce sold for income. The data we use does not measure income from self-employment, or shadow wages from home production. Most of the income data comes from those individuals earning wages in the private or public sector.

Figure 3.2: Percentage in wage changes for individuals who switch jobs and those that stay in the same jobs



The individual panel data shows considerable variability individual *nominal* wages. In Figure 3.2 we plot the distribution of quarter on quarter percentage wage changes for wage earning individuals in all periods (not just when storms hit). We compare wage changes for those who stay in jobs with identical employment characteristics (occupation, pay-type, pay regularity, sector) versus individuals whose job characteristics change in any way. Not surprisingly, wages are more variable when workers change jobs, but in most quarters wages do not change. However, individuals staying at the same jobs seem to exhibit downward flexility in nominal wages. Large drops in the nominal wage are common.

3.3 AGGREGATE EFFECTS

In this Section we establish that typhoons act as a strong but temporary labor demand shock. In the next Section, we provide evidence on the channels through which the adjustment takes place.

3.3.1 *Short-term effects*

We start by estimating equations of the form:

$$Y_{mpt} = \alpha S_{mpt} + \beta X_{mpt} + u_{mp} + \nu_t + w_{mpt} \tag{3.1}$$

Where Y_{mpt} is the outcome of interest in municipality m in province p at time t, S_{mpt} is a vector of variables capturing whether municipality m has been hit by a typhoon in the previous quarter, X_{mpt} is a vector of municipal characteristics that vary across time, u_{mp} is a municipality-specific unobservable, ν_t is a time-specific unobservable and, w_{mpt} is the usual idiosyncratic term. Given that we expect standard errors to be correlated for municipalities in the same provinces, standard errors are clustered at the provincial-level.[15]

Results, available in Panel A of Table 3.3, indicate that municipalities hit by a strong typhoon do not experience a change in their employment rate in the quarter following the shock. That is, labor markets do not appear to adjust along the extensive margin. Those results are robust to adding municipal fixed effects (Column 2) and a number of quarter-specific measures of sample composition at the municipal-level: education, gender and age (Column 3). We obtain similar results if we exclude municipalities from the southern island of Mindanao (Column 4). Typhoon incidence increases with latitude in the Philippines and, historically, Mindanao is very rarely hit by typhoons. No municipality in Min-

15 The sample includes data from more than 80 provinces so were are not concerned about bias in our standard errors as a result of having too few clusters (Cameron et al., 2008).

Table 3.3: Aggregate-level results

	(1)	(2)	(3)	(4)
Panel A: Impact on Employment Rate per Adult				
Big Storm	0.014	-0.005	-0.005	-0.007*
	(0.015)	(0.004)	(0.004)	(0.004)
Small Storm	-0.011	-0.001	-0.001	0.000
	(0.007)	(0.002)	(0.002)	(0.002)
Observations	29,560	29,560	29,560	21,064
R-squared	0.005	0.011	0.017	0.021
Mean Dep. Var	0.600	0.600	0.600	0.600
Panel B: Impact on Log Income per Adult				
Big Storm	-0.332***	-0.065***	-0.072***	-0.078***
	(0.091)	(0.022)	(0.023)	(0.024)
Small Storm	0.175***	-0.004	-0.004	-0.012
	(0.065)	(0.009)	(0.009)	(0.009)
Observations	28,608	28,608	28,608	20,808
R-squared	0.015	0.051	0.061	0.073
Mean Dep. Var	5.300	5.300	5.300	5.400
Mun FE	No	Yes	Yes	Yes
Agg Contr	No	No	Yes	Yes
Mindanao Incl.	Yes	Yes	Yes	No

Notes: Results from weighted municipal*quarter regressions. The dependent variable is the employment rate in the municipality (Panel A) and the average wage in the municipality (Panel B). Regressions control for time fixed effects (Column 1-4), municipal fixed effects (Column 2-4), as well as the share of the working age population in each education category, the share of women in the working age population, the number of men, the number of women, the number men age 15-30 and the number of women age 15-30 (Column 3-4). In Column 4, the sample is restricted to municipalities outside of Mindanao. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 3.4: Decomposing the aggregate-level effects

	(1) inc/ adult	(2) wage/ week	(3) wage/ hour	(4) hours/ earner	(5) earners/ job	(6) job/ adult
Big Storm	-0.078***	-0.035**	-0.020*	-0.015*	-0.032	-0.011
	(0.024)	(0.014)	(0.010)	(0.009)	(0.023)	(0.007)
Small Storm	-0.012	-0.013**	-0.012**	-0.002	0.002	-0.001
	(0.009)	(0.007)	(0.005)	(0.004)	(0.008)	(0.004)
Denominator	Adults	Earners	Earned Hours	Earners	Jobs	Adults
Observations	20,808	20,808	20,808	20,808	20,808	20,808
R-squared	0.073	0.131	0.146	0.068	0.024	0.016

Results from weighted municipal*quarter regressions. The dependent variable is the average income from employment per adult (Column 1), the average income from employment for employed individuals (Column 2), the average hourly wage for employed individuals (Column 3), the average number of hours worked for employed individuals (Column 4), the proportion of individuals who had jobs who reported a salary (Column 5), the proportion of adults who had jobs (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as the share of the working age population in each education category, the share of women in the working age population, the number of men, the number of women, the number men age 15-30 and the number of women age 15-30. The sample is restricted to municipalities outside of Mindanao. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

danao was hit by either a small or a large typhoon over the sample period and since employment patterns might be different there, we prefer excluding those observations from the sample as they do not contribute to the estimation of α.

Once we focus on income from employment, we find that municipalities experience a large decline in average income in the quarter following the shocks (Panel B of Table 3.3). The point estimates reported in Column 1 are very large, of the order of 32 percent but, once we control for municipal fixed effects (Column 2) the point estimate drop to a still economically significant 6.5 percent. This suggest that municipalities that tend to be hit by strong typhoons tend to be disadvantaged, which is consistent with findings by Hsiang and Jina (2014). Once we control for time-varying municipal controls and exclude municipalities from Mindanao the point estimates increase slightly and are still statistically different from zero at the one percent level.

We now decompose the effects on average income and estimate equation (3.1) for a number of other outcomes of interest using the specification with municipal fixed effects, time dummies and quarter-specific municipal controls on the non-Mindanao sample (Table 3.4).[16] We find that the overall effect comes from a two percent decline in hourly wage and a 1.5 percent decline in hours worked. To put it differently, at the aggregate-level, labor markets adjust by lowering hourly wages and reducing the number of hours worked.

Table 3.5: Aggregate-level results - Persistence

	(1) inc/ adult	(2) wage/ week	(3) wage/ hour	(4) hours/ earner	(5) earners/ job	(6) job/ adult
Big Storm						
current	-0.079***	-0.036**	-0.023**	-0.014	-0.029	-0.013**
	(0.026)	(0.015)	(0.011)	(0.010)	(0.025)	(0.006)
lag 1	-0.030	-0.017	-0.005	-0.011	-0.006	-0.007
	(0.026)	(0.015)	(0.014)	(0.011)	(0.027)	(0.006)
lag 2	0.036	0.017	-0.002	0.019*	0.026	-0.008
	(0.026)	(0.013)	(0.011)	(0.011)	(0.022)	(0.006)
lag 3	-0.036	-0.007	-0.007	-0.001	-0.012	-0.016**
	(0.022)	(0.012)	(0.013)	(0.011)	(0.022)	(0.007)
Small Storm (lags estimated but not displayed)						
current	-0.014	-0.014**	-0.013***	-0.001	0.001	-0.001
	(0.009)	(0.006)	(0.005)	(0.004)	(0.007)	(0.004)
Observations	20,579	20,579	20,579	20,579	20,602	20,835
R-squared	0.074	0.131	0.144	0.068	0.025	0.017

Notes: Results from weighted municipal*quarter regressions.The dependent variable is the average income from employment per adult (Column 1), the average income from employment for employed individuals (Column 2), the average hourly wage for employed individuals (Column 3), the average number of hours worked for employed individuals (Column 4), the proportion of individuals who had jobs who reported a salary (Column 5), the proportion of adults who had jobs (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as the share of the working age population in each education category, the share of women in the working age population, the number of men, the number of women, the number men age 15-30 and the number of women age 15-30. The sample is restricted to municipalities outside of Mindanao. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

16 Importantly, the results are robust to using alternative measures of storm strength (Tables C.1 and C.2).

3.3.2 *Persistence*

A potential concern with our results is that they only focus on short-term impacts of the storm and might fail to capture more relevant, longer-term impacts. We now estimate equation (3.1) including lagged values of the shock variables. Results displayed in Table 3.5 confirm our modelling choice. Storms do not appear to affect labor markets after one quarter. For example, when focusing on our main measures of economic activity, the point estimate of the shock measure lagged once is 60 percent lower than it is for the current version of the shock and it is no longer statistically significant. There is a similar pattern for other outcomes of interest: it is more than 50 percent lower for average wage and it is almost 80 percent lower for average hourly wage. We are not always able to reject the null that the estimated effects of the current value and the first lag are equal but once we look at the second and third lags the results confirm that the impacts on labor markets of storms are short-lived. From now on we focus on the current impacts of storms.

3.4 DOWNWARD NOMINAL WAGE FLEXIBILITY

Having established that large typhoons lead to a large aggregate decline in income from employment but no effects on employment levels, we now explore how firms and their workers adjust to these impacts. We use the full set of individual-level labour force observations and find results that are consistent with the results in the aggregate data. Average wages decrease after typhoons hit due to a combination of decline in hours worked per week and hourly wages. Consistent with our previous results, the effects on unemployment are very small, and we will show that the effects that we do find are driven entirely by impacts on self-employment.[17] Employment in wage labour is not effected

17 A finding in line with previous studies on the effects of typhoons in the Philippines (Anttila-Hughes and Hsiang, 2013).

It is important to note that, as we are interested in total wages, our preferred measure is weekly wage income as this is the highest level of aggregation we can use. We can decompose it into number of hours worked and hourly wage. Further, to understand how the adjustments take place we also look at the number of days worked and number of hours per days worked.

3.4.1 *Individual Decomposition*

Consistent with the aggregate results discussed in the previous Section, we estimate individual-level equations of the form:

$$Y_{imt} = \alpha S_{mt} + \beta X_{imt} + u_m + v_t + w_{imt} \tag{3.2}$$

Where Y_{imt} is the outcome of interest for individual i in municipality m i at time t, S_{mt} is a vector of variables capturing whether municipality m has been hit by a typhoon in the previous quarter, X_{imt} is a vector of individual characteristics, u_m is a municipality-specific unobservable, v_t is a time-specific unobservable and, w_{int} is the usual idiosyncratic term. Standard errors account for potential clustering of the errors at the municipal-level. As above, we first estimate equation (3.2) without any controls, then add time dummies, municipal fixed effects and individual controls (education, age, age squared and gender).

Individual-level results, available in Table 3.6 are consistent with the municipal-level results discussed above.[18] The probability of being employed isn't affected

18 We note discrepancies between the aggregate and individual data in the effects estimated thus far. The total effect on total wages per person at the municipal level is 7% (using the log of total wages). This is effect represents our estimate of the total average percentage change in labour earnings due to storms. It includes the effects of storms on average wages, employment, and missing incomes. By comparison, the effect on average wages in the aggregate data is 3.5%, while estimated effect on average wages in the individual data is 2.7%. This discrepancy seems to be driven by the use of the log of aggregate wages. If poorer municipalities are hit harder by storms (in relative terms) then the impact on the log of the average wage will be different from the average impact on the log of individual wages. We fully reconcile these results by looking at the impact of storms on the main variables in levels, in Table C.8. This also allows us to look at the impact of the storms on income per adult for the individual data. In this table we find that the results are almost identical between the two datasets. When expressed as percentage

Table 3.6: Individual-level results: Impacts on wages and employment

	(1)	(2)	(3)	(4)
Panel A: Impact on Employment per Adult				
Big Storm	0.014*	-0.005	-0.005	-0.007*
	(0.008)	(0.004)	(0.004)	(0.004)
Small Storm	-0.012***	-0.001	-0.001	0.000
	(0.004)	(0.002)	(0.002)	(0.002)
Observations	3,402,456	3,402,456	3,402,456	2,464,172
R-squared	0.000	0.023	0.228	0.219
Mean Dep. Var	0.600	0.600	0.600	0.600
Panel B: Impact on Log of Weekly Wages				
Big Storm	-0.246***	-0.022*	-0.024**	-0.027**
	(0.044)	(0.013)	(0.011)	(0.011)
Small Storm	0.105***	-0.005	-0.007	-0.010**
	(0.019)	(0.006)	(0.005)	(0.005)
Observations	860,809	860,809	860,809	660,650
R-squared	0.012	0.216	0.444	0.446
Mean Dep. Var	6.900	6.900	6.900	7.000
Mun FE	No	Yes	Yes	Yes
Ind Contr	No	No	Yes	Yes
Mindanao Incl.	Yes	Yes	Yes	No

Notes: Results from weighted individual regressions. The dependent variable is a dummy equal to one if the individual is employed (Panel A) and log of wages for employed individuals (Panel B). Regressions control for time fixed effects (Column 1-4), municipal fixed effects (Column 2-4), as well as the respondent's age, age square, education levels and gender (Column 3-4). In Column 4, the sample is restricted to municipalities outside of Mindanao. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

by typhoons but average wage for individuals employed is 2.7 percent lower in the post-storm quarter. As before, the effects are no longer present after one quarter (Table C.3).

We can decompose the effect on average income (Table 3.7). In the quarter after the storm, individuals report working 1.8 percent fewer hours (Column 2). The

of the mean depedent variable, we find that storms have an impact of income per adult of 3%. This shows that the results are driven by the use of logarithms of aggregate data rather than inconsistencies in our application of sample weights or definition of variables.

Table 3.7: Individual-level results: decomposition

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Impact on Intensive Margins (Earnings and Hours)						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm	-0.027**	-0.018**	-0.016*	-0.011	-0.015**	-0.002
	(0.011)	(0.009)	(0.009)	(0.008)	(0.007)	(0.004)
Small Storm	-0.010**	-0.008**	-0.003	-0.007*	-0.001	-0.002
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	660,650	1,430,357	660,650	660,650	660,650	660,650
R-squared	0.446	0.128	0.094	0.417	0.093	0.039
Panel B: Impact on Extensive Margins						
	employed	job	wage missing	wage observed	zero hours	lost job quarter
Big Storm	-0.007*	-0.006	0.006	-0.006*	0.001	0.001
	(0.004)	(0.004)	(0.006)	(0.004)	(0.001)	(0.002)
Small Storm	0.000	0.000	-0.001	0.000	0.000	-0.002**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.001)
Sample	All	All	Earners	All	All	All
Observations	2,464,172	2,464,172	1,430,353	2,464,172	2,464,172	2,464,172
R-squared	0.219	0.228	0.188	0.097	0.015	0.021
Mean Dep. Var	0.573	0.581	0.507	0.286	0.009	0.030

Notes: Results from weighted individual regressions. In Panel A, the dependent variable is the log weekly wage for employed individuals (Column 1), number of hours worked for employed individuals (Column 2), number of hours worked for employed individuals earning a wage (Column 3), hourly wage for employed individuals (Column 4), number of days worked for employed individuals earning a wage (Column 5), number of hours worked per day for employed individuals earning a wage (Column 6). In Panel B, the dependent variables are a series of dummies equal to one if: the individual is employed (Column 1), the individual has a job (Column 2), the individual is employed but their wage is not observed (Column 3), the individual reports a wage regardless of employment status (Column 4), the individual reports having a job but working zero hours in the last 7 days (Column 5), the individual reports not having a job now, but having worked in the last 3 months (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

point estimate on the hourly wages is negative and of the same order of magnitude as before. The individual-level are noisier than the aggregate measures used previously and we are unable to reject the null of no effect, however.

Results discussed so far suggest that nominal wages exhibit significant downward flexibility when a typhoon hits. However, a well-established literature on the cyclicality of wages suggests that aggregation or selection effects that could bias results either in favour or against findingwage flexibility (Keane et al., 1988). For instance, if high wage employees are more likely to be laid off during times of labour demand shocks, results could be biased in favour of finding declines in real wages. If these high productivity workers were replaced with lower productivity workers, not only would this increase the extent of the bias, it would also suggest no impact on overall employment. Conversely, if low wage workers are laid off, this could lead to bias against finding a result of downward wage adjustment.

In the remainder of this Section, we take each of those threats in turn and show that are results are robust. First, shocks could generate sectoral reallocation which could lead to a reduction in average wage without any wage adjustments for workers who do not switch jobs. Second, shocks could affect sample composition. We start by showing that average observable characteristics of individuals in the sample, employed individuals and individuals earning a wage are not affected by storm. We then use a panel of individuals - hereby keeping sample composition constant - and find that results hold in this subsample. Finally, firms may have substituted highly paid skilled jobs for lower skilled jobs to reduce their total wage costs. We restrict the sample to individuals who are employed in similar jobs during the period and find that the results hold on this subsample.

3.4.2 *Are the results driven by sectoral reallocation?*

Economic shocks like those caused by large natural disasters can have a large impacts on the composition of employment in affected areas and can change the sectoral composition of economic activities (Moretti, 2010; Kirchberger, 2014). If the storms studied in this paper caused sectoral shifts toward lower paying industries and jobs, this could be driving the effects on average wages. While this appears unlikely since the effects discussed so far are short-lived, here we show that the overall composition of jobs did not change in the full individual sample.

Panel A of Table 3.8 shows the impacts of storms on the probability of a working individual being employed at a particular type of job. Only one category of work is affected by storms; individuals are less likely to be engaged in self-employment in weeks when storms hit. In fact the main impact of storms on employment found in Table 3.7 is driven almost entirely by the impact on reduced self-employment. The impacts on self-employment are not likely to be driving our results on real wages. Self-employed wages are not observed in this data: 99% of all self-employed individuals have their wages reported as missing, and the data allows no way to impute income from self-employment. This is an unfortunate limitation of the data. Anttila-Hughes and Hsiang (2013) use household data to show that storms have large impacts on household earnings, much of the effect comes through self-employment but the focus of this paper is on wage employment. Wages are very rarely missing in private sector jobs (only about 9% of the time). As we show in sectoral analysis in Section 5 the impacts on income are driven by wage changes in the private sector.

Panel B of Table 3.8 shows that the composition of jobs across wage paying forms of employment (temporary or permanent private sector work, government work, and wage paying farm labour) are unaffected by storms. Panel C of Table 3.8 reproduces the analysis on the sample of individuals earning a wage. We

Table 3.8: Individual-level results: Employment in different types of jobs

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Effect (Unconditional on having a job)						
	selfemp	permpriv	temppriv	ownfarm	wagefarm	gov
Big Storm	-0.005**	-0.001	-0.001	0.001	0.001	-0.002
	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)	(0.001)
Small Storm	0.000	-0.002	0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	2,464,172	2,464,172	2,464,172	2,464,172	2,464,172	2,464,172
R-squared	0.056	0.092	0.028	0.247	0.115	0.073
Mean Dep. Var	0.131	0.169	0.057	0.127	0.046	0.043
Panel B: Composition Effect (Conditional on having a job)						
	selfemp	permpriv	temppriv	ownfarm	wagefarm	gov
Big Storm	-0.006	0.002	-0.001	0.004	0.000	-0.002
	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)	(0.002)
Small Storm	-0.001	-0.002	0.002	0.001	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Observations	1,453,619	1,453,619	1,453,619	1,453,619	1,453,619	1,453,619
R-squared	0.084	0.170	0.065	0.315	0.160	0.113
Mean Dep. Var	0.226	0.291	0.097	0.217	0.078	0.079
Panel C: Composition Effect (Conditional on earning a wage)						
	selfemp	permpriv	temppriv	ownfarm	wagefarm	gov
Big Storm	0.001	-0.000	0.005	-0.001	0.004	-0.009**
	(0.001)	(0.009)	(0.007)	(0.001)	(0.006)	(0.004)
Small Storm	-0.000	-0.008*	0.006	0.001	0.002	-0.001
	(0.001)	(0.004)	(0.004)	(0.000)	(0.003)	(0.002)
Observations	669,711	669,711	669,711	669,711	669,711	669,711
R-squared	0.005	0.145	0.073	0.023	0.366	0.210
Mean Dep. Var	.005	.54	.183	.001	.132	.127

Notes: Results from weighted individual regressions. The dependent variable is a dummy equal to one if the individual is: self-employed (Column 1), has a permanent job in the private sector (Column 2), has a temporary job in the private sector (Column 3), works on the family farm (Column 4), works for a wage on someone's else farm (Column 5), is employed in the public sector (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

find that individuals earning a wage are very slightly less likely to work in the public sector. Overall, we interpret this set of results as indicating that the decline in nominal wages observed in the quarter after storms hit are not driven by sectoral reallocation. It is important to note now that, once we focus on the panel of individuals that we observe more than once in the data, there is no evidence that storms affect the sectoral composition of jobs on this subsample (Table C.4).

3.4.3 *Are the results driven by changes in sample composition?*

Typhoons could affect sample composition which, depending on which individuals drop out of the sample, could generate the observed pattern of wage reduction. This could happen, for instance, if certain types of individuals left the local area to look for work in areas less effected by storms. We estimate equation (3.2) regressing the individual-level characteristics on which we have data (education, age and gender) on the full set of municipal and time fixed effects and the storm dummies. We estimate each of those equations on the full sample, on the sample of employed individuals and on the sample of individuals earning a wage. Results, available in Table 3.9, do not suggest that the timing of typhoon occurrence affects sample composition. Among the 24 tests carried out (gender, age and 6 education categories on the 3 samples), we only reject the null three times. Employed individuals are slightly older and slightly less likely to have graduated from high school in the quarters where storms hit.

3.4.4 *Panel Decomposition*

We take advantage of the availability of panel data and show that the results are similar for this sub-sample. By construction, this set of analyses keeps the sample constant. Importantly, on average, individuals observed more than once do not appear to systematically differ from the rest of the sample (Table 3.2). This

Table 3.9: Individual results: Impacts on composition of the sample

	(1) female	(2) age	(3) edu 1	(4) edu 2	(5) edu 3	(6) edu 4	(7) edu 5	(8) edu 6
Panel A: Impact on the Characteriristic (Composition) of the Full Sample								
Big Storm	0.001	0.102	-0.001	-0.001	0.003	0.004	-0.003	-0.003
	(0.002)	(0.113)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Small Storm	0.001	0.084*	0.001	0.000	0.001	-0.003**	0.001	0.001
	(0.001)	(0.049)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	2,464,172	2,464,172	2,464,172	2,464,172	2,464,172	2,464,172	2,464,172	2,464,172
R-squared	0.002	0.010	0.023	0.080	0.038	0.008	0.032	0.072
Mean Dep. Var	0.510	36.070	0.010	0.130	0.150	0.160	0.260	0.280
Panel B: Impact on the Characteriristic (Composition) of the Individuals Employed								
Big Storm	0.002	0.150	-0.000	-0.002	0.006	0.005	-0.009**	-0.001
	(0.003)	(0.125)	(0.001)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)
Small Storm	0.004**	0.031	0.000	-0.000	0.001	-0.003**	0.001	0.002
	(0.002)	(0.054)	(0.000)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Observations	1,453,619	1,453,619	1,453,619	1,453,619	1,453,619	1,453,619	1,453,619	1,453,619
R-squared	0.013	0.016	0.041	0.106	0.048	0.010	0.043	0.091
Mean Dep. Var	0.39	37.66	0.01	0.15	0.17	0.13	0.27	0.28
Panel C: Impact on the Characteriristic (Composition) of the Individuals Earning a Wage								
Big Storm	0.009	0.431**	0.000	-0.002	0.007	0.008	-0.013**	0.000
	(0.006)	(0.178)	(0.001)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
Small Storm	0.006**	0.091	-0.000	-0.001	-0.000	-0.003	0.002	0.002
	(0.003)	(0.076)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Observations	669,711	669,711	669,711	669,711	669,711	669,711	669,711	669,711
R-squared	0.017	0.015	0.024	0.094	0.046	0.012	0.035	0.075
Mean Dep. Var	0.51	36.07	0.01	0.13	0.15	0.16	0.26	0.28

Notes: Results from weighted individual regressions. The sample is restricted to individual employed (Panel B) and individuals observed earning a wage (Panel C) . The dependent variable is a dummy variable equal to one if the respondent is female (Column 1), respondent age (Column 2), a dummy variable if the respondent did not complete any grade (Column 3), attended, but did not graduate from, primary school (Column 4), graduated from primary school but did not attend high school (Column 5), attended, but did not graduate from, high school (Column 6) graduated from high school but did not attend college (Column 7), attended College (Column 8). Regressions control for municipal fixed effects, region-specified time fixed effects. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

mitigates concerns about the representativeness of the panel data. We estimate equation (3.2) on the panel described in Section 3.2. Panel A of Table 3.10 shows the main results for the individual panel sample.[19] Wages fall by 3.5 percent, a slightly larger estimate than the results from the individual data.[20] Again the results seem to be driven by combination of a drop in hours per worker, and a fall in the hourly wage: here the impact on the hourly wages is larger and significant.

A related concern is that individuals who stayed in the panel might have switched to different job types. As above this would generate our results without any worker experiencing a drop in hours or in income within the same job. We estimate equation (3.2) but further restrict the sample to individuals who stay in similar job types throughout the sample period.[21] Results, available, Panel B of Table 3.10, confirm that even in this restricted sample workers experience a drop on their wage income in the short term that is driven by both a decline in hours and a decline in hourly wages.

A final concern is that individuals who did not move and stayed in similar job types might have renegotiated their contracts with, for example, workers switching from permanent to temporary contracts. To address those concerns, we restrict the sample to individuals who stay in similar job types and similar contract types and estimate equation (3.2) on this sub-sample. This individuals also remain on the same payment schedule, whether it be monthly payments, daily payments, or pay on commission. Again results available in Panel B of Table C.9 confirm our earlier results.

19 Given that the outcomes we're interested in are persistent and subject to measurement error, we do not estimate an individual fixed-effects model although the main results are robust to the use of individual fixed-effects in these regressions (see Table C.5 in the appendix).

20 Importantly, as in the full sample, there is no evidence that the probability of being employed is affected by the timing of typhoons (Column 1 of Table C.6).

21 The data does not allow us to distinguish between workers who have switched jobs and those who have remained in the same job since the last quarter.

Table 3.10: Panel-level results: decomposition

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Impact on Earnings and Hours (All Employees)						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm	-0.034***	-0.025**	-0.024**	-0.017*	-0.014*	-0.011**
	(0.012)	(0.010)	(0.010)	(0.009)	(0.008)	(0.005)
Small Storm	-0.010	-0.011**	-0.003	-0.010*	0.000	-0.003
	(0.006)	(0.005)	(0.004)	(0.005)	(0.003)	(0.002)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	267,038	699,704	277,932	267,038	277,928	277,928
R-squared	0.465	0.131	0.107	0.439	0.100	0.052
Panel B: Impact on Earnings and Hours (Same Job Type)						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm	-0.028*	-0.020*	-0.009	-0.018*	-0.010	0.001
	(0.014)	(0.011)	(0.011)	(0.011)	(0.009)	(0.005)
Small Storm	-0.003	-0.007	0.007	-0.010*	0.002	0.004
	(0.007)	(0.005)	(0.004)	(0.006)	(0.004)	(0.002)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	194,717	502,444	195,728	194,717	195,726	195,726
R-squared	0.491	0.146	0.124	0.462	0.121	0.054
Mun Fe	No	No	No	No	No	No

Notes: Results from weighted individual regressions. In Panel A, the dependent variable is the log weekly wage for employed individuals (Column 1), number of hours worked for employed individuals (Column 2), number of hours worked for employed individuals earning a wage (Column 3), hourly wage for employed individuals (Column 4), number of days worked for employed individuals earning a wage (Column 5), number of hours worked per day for employed individuals earning a wage (Column 6). In Panel B, the dependent variables are a series of dummies equal to one if: the individual is employed (Column 1), the individual has a job (Column 2), the individual is employed but their wage is not observed (Column 3), the individual reports a wage regardless of employment status (Column 4), the individual reports having a job but working zero hours in the last 7 days (Column 5), the individual reports not having a job now, but having worked in the last 3 months (Column 6). Regressions control for time fixed effects as well as municipal fixed effects (Panel A) and individual fixed effects (Panel B). In Panel A, regression control for the respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Before exploring the mechanisms behind the estimated effects, we further clarify why panel data are especially useful in our context. First, while Keane et al. (1988) have suggested that the use of panel estimators do not fully deal with the problem of selection bias we argue that their concerns are less likely to hold in our case. Their argument is that if high-skilled individuals in the panel are less likely to be employed in quarters of storms this could lead to the impression that wages are flexible downwards. However this problem arises from a setting where changes in unemployment are used as the dependent variable; by definition these estimators look at situations where there is a lot of movement out of the labour force. However, this is unlikely to explain our results as we found no evidence that storms affect the probability of being employed and of being engaged in different types of wage paying work conditional on being employed (Table C.4). Furthermore we restrict our sample of panel observations to individuals that we observe working in at least two periods. The vast majority of individuals are observed in the panel only twice. By looking at the sample of those individuals who were earning in both of those periods of the panel, we clearly document changes in their wages between the two periods.

Third, the use of panel data is used to overcome the possibility of sample selection changes induced by storms, which might lead to our results being driven by churn in the labour market rather than actual falls in particular workers' wages. While we cannot fully rule out the possibility that panel sample switched to lower paid jobs within the same sector and with the same terms of employment, it seems unlikely that the large treatment effects found here are consistent with levels of churn of this kind. Especially since, as we show in Section 3.5.2 the results are particularly strong for longer term jobs where churn is likely to be lower and they are only detectable in the short-run.

Fourth, the panel data helps with us to deal with concerns related to aggregation bias due to migration, since we observe reductions in wages for individuals who have not migrated.

3.5 MECHANISMS

We now explore mechanisms through which nominal wages adjust downwards after a typhoons hit. Results discussed so far are consistent with the adjustments one would predict from a spot market for labour where wages adjust to a market clearing level, unemployment doesn't increase but the total number of hours worked falls.[22] Are these results driven by adjustment in wages for labour spot markets rather than long term relationships? Indeed if the results are concentrated among jobs governed by spot markets the results could be driven purely by lower supply of labour hours resulting in lower weekly and and daily wage rates.

We start by showing that typhoons occurrence does not appear to affect measures of labour supply. We argue that, and present evidence for, flexibility arises in established contractual employment relationship, with strong effects observed for individuals employed on permanent contract in the private sector which we interpret as flexibility in implicit contracts. Finally, we explore heterogeneity in the estimated effects.

3.5.1 *Market Clearing Labour supply*

We have argued that the results presented here are driven by reductions in labour demand, resulting in movements down the labour supply curve. Furthermore, downward wage flexibility seems to allow labour markets to clear: although average hours per worker falls in response to storms, there is no impact on

22 Interestingly Kaur (2014) finds that spot markets for agriculture in labour markets actually exhibit considerable wage rigidity.

unemployment and, we would expect, no impact on unemployment unless there were coinciding changes in labour supply.

Table 3.11: Individual-level and panel-level results: Labour supply

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full Individual Dataset						
	inlf	any search	unemployed	lookw	want more	search addw
Big Storm	-0.005	0.002	0.004	-0.002	0.002	0.001
	(0.004)	(0.003)	(0.004)	(0.003)	(0.009)	(0.006)
Small Storm	0.002	-0.004**	0.003*	-0.001	-0.007	-0.005*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)
Observations	2,464,172	2,464,172	1,588,750	1,010,552	1,430,353	1,098,598
R-squared	0.233	0.043	0.060	0.063	0.114	0.104
Mean Dep. Var	0.640	0.071	0.106	0.066	0.184	0.093
Panel B: Panel Dataset						
	inlf	any search	unemployed	lookw	want more	search addw
Big Storm	-0.003	-0.001	-0.005	-0.003	-0.008	0.005
	(0.004)	(0.004)	(0.004)	(0.003)	(0.010)	(0.008)
Small Storm	-0.001	-0.004*	0.001	-0.002	0.007	0.007**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)
Observations	1,294,842	1,294,842	1,294,842	399,704	699,704	455,862
R-squared	0.001	0.002	0.002	0.001	0.005	0.016
Mean Dep. Var	0.665	0.070	0.603	0.047	1.808	1.900

Notes: Results from weighted individual regressions. The dependent variable is a dummy equal to one if the individual is:in the labor force (Column 1) report having searched for work in the past week, regardless of labour force status (Column 2), not working, conditional on being in the labour force (Column 3), looking for work, conditional on being in the labour force and not working (Column 4), wanting more work, conditional on already having a job (Column 5), reported looking for additional work, conditional already having a job (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

We rule out that our results are driven by changes in labour supply. This is important as Jayachandran (2006) finds that large agricultural productivity shocks causes shifts in labour supply away from farm work *towards* wage labour, which in turn accounts for large reductions in wages. Similarly, Kochar (1999) shows that hours worked increase in rural areas as rural households attempt to smooth consumption during shocks.

In Panel A of Table 3.11 we show that storms have no impact on various measures of labour supply. Respondents are no less likely to report being in the labour force (Column 1), no more likely to be searching for work (whether employed or not), and no more likely to be looking for work while unemployed. Also there is no increase in the probability of an individual who has employment either wanting more work (Column 5) or having searched for additional work (Column 6). This provides strong evidence that large labour demand shocks do not result in wage rationing: labour markets seem to clear in the wake of large shocks.

In Panel B of Table 3.11 we confirm that this holds for the sample that stayed in the individual panel, with the coefficients following much the same pattern as in the individual data. This result is important: the analysis of wages in the panel data has focused upon individuals who were observed in employment and with wages for at least two periods.

Labour markets seem to clear both at the extensive margin (no rise in unemployment) and at the intensive margin (no rise in under-employment as measured by a demand for additional hours of work). First, at the extensive margin, this could be accounted for by movements along a very inelastic short-run labour supply curve. Wage flexibility in labour contracts allows wages to fall, and workers to keep their jobs. Workers are still happy to work at these reduced weekly wages, partly because they cannot afford to loose their entire wage incomes when storms hit. Secondly, workers may anticipate that the reduction in salaries will be short-lived (as our results on persistence suggest they usually are) and thus be willing to take lower salaries in the current quarter to ensure that they keep their jobs once wages rise again.

Second, at the intensive margin, we find some evidence that hourly wages adjust downwards, but that the main results are driven by reductions in hours worked. This suggests movements along an individual labour supply curve that appears to be highly *elastic*. We cannot rule out that these results are also driven in part by shifts in the labour supply curve at the intensive margin. Storms do substantial damage to homes, farms and home enterprises- workers might be willing to substitute labour away from work at wage paying work to spend more time dealing with such problems at home. This substitution is mutually beneficial if firms simultaneously have less demand for worker hours when storms hit, which appears to be the case in the data.

3.5.2 *Wage employment in the private sector*

We provide evidence consistent with the argument that downward wage flexibility is driven by wage flexibility within wage employment contracts, rather in forms of labour for which wages are determined by spot contracts. First, we estimate equation (3.2) but interact the storms variable (and all other control variables) with a dummy equal to one for individuals in wage employment in the private sector (either on permanent or temporary contracts). Results are available in Panel A of Table 3.12. Interestingly, the base effect suggests that there is no impact of storms outside of the private sector but the interaction term indicates that weekly wages in the private sector decrease by 4.7 percent in the post-storm quarter. While workers outside of the private sector experience a reduction in the number of hours worked, private sector workers experience a reduction in their hourly wage.

In addition, we restrict the sample to workers in wage employment in the private sector and compare the effects for individuals employed on temporary contracts and individuals employed on permanent contracts. Overall, we are unable to reject that the effects on weekly wage are similar but the adjustment margins differ greatly (Panel A of Table 3.12). Indeed, while individuals on temporary contract reduce the number of hours worked (mostly through a reduction in the

Table 3.12: Individual-level results: A closer look at the private sector

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Decomposition of Impacts among Private Sector Wage Employment and Other Jobs						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm	0.002	-0.031***	-0.021	0.020	-0.031**	0.010*
	(0.019)	(0.011)	(0.014)	(0.013)	(0.013)	(0.006)
Small Storm	-0.017*	-0.003	-0.003	-0.013*	-0.002	-0.001
	(0.009)	(0.006)	(0.006)	(0.008)	(0.005)	(0.003)
Big Storm * priv	-0.049**	0.055***	0.010	-0.056***	0.028**	-0.019***
	(0.024)	(0.017)	(0.016)	(0.019)	(0.014)	(0.007)
Small Storm * priv	0.014	-0.017*	-0.001	0.013	0.001	-0.002
	(0.011)	(0.009)	(0.007)	(0.010)	(0.006)	(0.004)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	660,650	1,430,357	660,650	669,711	660,650	660,650
R-squared	0.469	0.156	0.124	0.441	0.119	0.051
Panel B: Decomposition of Impacts among Permanent and Temporary Private Sector Wage Jobs						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm * permanent	-0.024**	0.003	0.003	-0.027**	0.003	-0.001
	(0.012)	(0.010)	(0.010)	(0.012)	(0.007)	(0.005)
Small Storm * permanent	-0.009	0.000	0.003	-0.012**	0.002	0.001
	(0.006)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)
Big Storm * temporary	-0.037	-0.064***	-0.057***	0.019	-0.044***	-0.013
	(0.023)	(0.019)	(0.020)	(0.017)	(0.014)	(0.010)
Small Storm * temporary	0.005	-0.012	-0.010	0.014	-0.003	-0.007
	(0.011)	(0.010)	(0.010)	(0.010)	(0.007)	(0.006)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	465,245	510,571	465,245	465,245	465,245	465,245
R-squared	0.418	0.088	0.089	0.395	0.081	0.045
Equality F-stat	0.261	9.617	6.986	5.343	8.613	1.221
Equality p-val	0.610	0.002	0.008	0.021	0.003	0.269

Notes: Results from weighted individual regressions. In Panel A, the dependent variable is the log weekly wage for employed individuals (Column 1), number of hours worked for employed individuals (Column 2), number of hours worked for employed individuals earning a wage (Column 3), hourly wage for employed individuals (Column 4), number of days worked for employed individuals earning a wage (Column 5), number of hours worked per day for employed individuals earning a wage (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as respondent's age, age square, education levels and gender. In Panel A regressions include a private sector dummy. In Panel B regressions include a permanent contract dummy. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

number of days worked), individuals on permanent contracts do not adjust their hours but experience a 2.7 percent reduction in their hourly wage.

The evidence suggests that the results are significantly different between temporary and permanent jobs. Most striking is that permanent jobs exhibit considerable downward flexibility in *hourly* wages. There is relatively little adjustment in hours worked per paid worker (Column 3). The weekly wage adjustment for temporary jobs is not significantly different to that in permanent jobs, but the results seems to be driven by a fall in the number of hours worked rather than the hourly wage.

The evidence here seems to suggest that, even long- term permanent contracts agreements exhibit high levels of flexibility. Those findings are consistent with implicit contracts that allow state-contingent wages. Conversely, results for temporary forms of employment are consistent with the behaviour of a spot-market, with highly elastic labour supply: workers reduce the number of days worked. No layoffs occur for either type of job.

Another mechanism by which labour markets adjust to storms deserve mention. While we are unable to detect any impacts of storms on the probability of having a wage paying job and on having worked positive hours at such a job, we establish that storms cause an increase in the probability that an individual reported working at wage paying job, but not receiving a wage in the particular period (Panel B of Table 3.7). We find similar effects once we focus on the panel of individuals observed more than once (Table C.6). The effects on this outcome are small - they do not seem to be playing an important role in driving the main results in the paper- but they suggest an interesting risk-sharing channel through which employers adjust to large shocks. Workers were not laid off, and did not stop working at their jobs but they appear not to receive their wages when storms hit, possibly because firms are unable to cover the costs of their salaries. Does this mean that employers compensated their workers in the later periods, after the major effects of the storms have passed? Results reported in Table C.3 indicate that wages are significantly *less* likely to be missing two quarters after storms hit. We interpret this as suggestive evidence that employers are more likely to pay workers in periods after the storms to compensate for the weeks in which they were less likely to pay.

3.5.3 *Heterogeneity*

We now explore heterogeneity in the estimated effects. We focus on two main dimensions: the level of urbanisation and the type of occupation. The evidence suggests that urban and rural areas are equally affected by strong storms. We further establish that managers tend to increase their earnings during storms due to an increase in the number of hours worked.

Urban-rural Heterogeneity

The extent of wage flexibility might differ between rural and urban areas. In rural settings, we might expect that outside options might be more sensitive to storms: labour markets are likely to be thinner so that workers are less likely to find alternative work options in other jobs, and rural households rely far more on subsistence agriculture to supplement incomes and insure against risk of layoff. Subsistence agriculture is very likely to be adversely effected by storms, which might limit lower workers outside options, and limit labour supply flexibility and lead to stronger downward adjustment of wages (Jayachandran, 2006). Therefore wages in labour contracts might be more likely to adjust downwards during shocks. By contrast it may be that smaller communities and more traditional behavioural norms in rural areas regulate labour markets and ensure that wages cannot fall due after shocks (Kaur, 2014).

We estimate estimate equation (3.2) but interact the storms variables with a city dummy (Table C.7). We find no significant heterogeneity between the rural and urban areas.[23] All of the effect comes through the storm variable, the interaction term is not significant.[24] One additional important result emerges. Until now we have seen little impact of small storms on labour outcomes. This is perhaps because the damage caused by these storms, while often severe for small scale farmers and individual households, is not enough to significantly disrupt the formal sector. However Table C.7 suggests that for rural areas, small storms do have an impact. The size of the effects is small relative to larger storms, but statistically significant. But contrast, the sign on the interaction of small storm and city in Column 1 is significant, in the opposite direction, suggesting the impact of being hit by a small storm is completely mitigated in urban areas.

Skill Bias

A long literature looks at the impacts of large shocks on the relative composition and earnings within local labour markets (Moretti, 2010). Kirchberger (2014) shows that damage caused by earthquakes leads to persistent increases in wage premia in the construction sector when reconstruction occurs. Keane and Prasad (1996) show that large spikes in the price of oil leads to a rise in the relative wage of more skilled workers, although wages decline for all workers.

We estimate estimate equation (3.2) on the sample of private sector workers and distinguish between individuals employed as managers and individuals employed in other occupations (Table 3.13). The negative coefficient on average wages for non-manager workers estimated here is consistent with the main results. However, we find that managers see large rises in their wages. This impact is significantly different from the impact on non-managers. Interestingly, this effect is not driven by an increase in the hourly wage of these workers. The increases in their wages are driven by large increases in the number of hours worked by managers. They work both longer days and more days. We specu-

23 This finding is robust to using municipal level urbanization rates.

24 Although it is positive, suggesting that if anything, impacts are slightly bigger in rural areas.

late that these results are driven by the need for managerial oversight during times of crisis, as firms shift priorities away from usual business to recovering assets, dealing with storm damage, and otherwise adjusting to shocks. Firms may arrange with managers to work additional, or overtime, hours during times of crisis to manage the fall-out from storms.

Table 3.13: Individuals-level results: Heterogenous treatment effects by managerial and non-managerial private sector jobs)

	(1) wage/ week	(2) hours/ worker	(3) hours/ earner	(4) wage/ hour	(5) days/ earner	(6) hours/ day
Big Storm * non manag	-0.035***	-0.035***	-0.021**	-0.017*	-0.019**	-0.003
	(0.013)	(0.010)	(0.010)	(0.009)	(0.009)	(0.005)
Small Storm * non manag	-0.011*	-0.011**	-0.005	-0.006	-0.002	-0.003
	(0.006)	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)
Big Storm * manag	0.199**	0.141***	0.176***	0.008	0.092***	0.081***
	(0.085)	(0.021)	(0.037)	(0.094)	(0.021)	(0.024)
Small Storm * manag	-0.026	0.004	-0.011	-0.017	-0.019	0.008
	(0.033)	(0.012)	(0.020)	(0.032)	(0.014)	(0.012)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	566,279	1,317,287	566,279	575,322	566,279	566,279
R-squared	0.464	0.157	0.101	0.414	0.101	0.045
Equality F-stat	7.148	56.877	25.197	0.067	21.428	11.371
Equality p-val	0.008	0.000	0.000	0.795	0.000	0.001

Notes: Results from weighted individual regressions. Sample is restricted to individuals working in the private sector. In Panel A, the dependent variable is the log weekly wage for employed individuals (Column 1), number of hours worked for employed individuals (Column 2), number of hours worked for employed individuals earning a wage (Column 3), hourly wage for employed individuals (Column 4), number of days worked for employed individuals earning a wage (Column 5), number of hours worked per day for employed individuals earning a wage (Column 6). Regressions control for municipal fixed effects, region-specified time fixed effects as well as respondent's age, age square, education levels and gender. Regression also include a full set of job type dummies. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

3.6 THEORETICAL FRAMEWORK

In this section we develop a model to explain the key findings for the private sector in the paper. We use a model with long term contractual relationships, in which risk sharing occurs between workers and firms, and workers are insured against shocks through work-sharing.[25]

In the absence of downward rigidities wage adjustments moderate the impact of shocks on firm labour demand and allow the market to clear. Our results show a fall in weekly wages across all private sector jobs. However, contracts must determine the trade off between lay-offs and reduction in hours per worker, to the extent that total labour demand does fall during shocks. Similar models have been used to explain stylized facts from the US where labor markets are characterized by high variability of employment and relatively constant hours per worker (Burdett and Mortensen, 1980). Our setting is different as hours appear to be relatively flexible.

We demonstrate conditions for which it is optimal for no lay-offs to occur. Workers are paid less and work fewer hours during periods when storms hit. The model predicts that wages and hours should fall but we do not explicitly model the impact on the hourly wage. In the case where the adjustment occurs mostly through nominal wage adjustments, the hourly wage will fall significantly. This is the result that we find for permanent jobs in the private sector. In the case that the adjustment in hours and total wages are similar, the effect on the hourly wage is ambiguous, which is the finding for temporary jobs in our data.

We use a version of the classic implicit contract models of Baily (1974) and Azariadis (1975). In the standard model risk averse firms and workers contract over total labour demand (employment) and wages for each state of the world. We adapt these models with extensions by Rosen (1985) and Miyazaki and Neary (1985), which focus on the role of lay-offs and hours per worker in optimal contracts by allowing hours per worker to enter the production function separately from the number of employed workers.

25 Many models, including auction markets for daily labour, would show reductions in wages due to labour demand shocks. We need a model to explain why no layoffs occur when firm labour demand is reduced.

Rosen (1985) writes that implicit labour countries should "[specify] precisely the the amount of labour to be utilized and the wages to be paid in each state of nature, that is, conditional on information (random variables) observed by both parties." Importantly, this assumption is realistic in our setting: storms are observed by everyone and can be contracted upon.

3.6.1 *The Model*

A representative firm contracts with a set of n workers. Workers and firms are risk averse. Contracts are perfectly enforceable and contingent on the realized state of the world θ. Low realizations of θ correspond to large negative shocks, driven by typhoons in this paper. In the model, θ represents a shock to firm marginal revenue product. We imagine that storms could impact firm profits by reducing output, for instance if capital is destroyed or the efficiency of labour inputs is disrupted. Alternatively, storms could reduce domestic demand which would lead to lower prices. The model does not distinguish between these channels, both are fully captured by changes in θ.

In the benchmark model, firm production is a function only of a single labour input, usually the number of workers employed by the firm. If n is the number of workers under contract (which is constant in this model) and $p(\theta)$ is the proportion that are hired when the value of θ is realized, then production is given by $\theta f(pn)$. Labour demand is adjusted through changes in p alone for this simple case.

We adapt this benchmark model by allowing hours per worker h to be adjusted, so that firms use total worker-hours given by phn. Labour is not necessarily perfectly divisible, production is given by $f(np, h)$. Firms pay wages only to those workers who they employ, at wage rate w. We simplify the standard model by assuming that firms cannot provide private insurance to laid off workers, so workers earn only an outside option wage when they are laid off.[26] Firm profit is given:

$$\pi = \theta f(pn, h) - wnp \tag{3.3}$$

Firms have utility over profits $v(\pi)$. This assumption is justified by credit and insurance market failures on the part of firms (Rosen, 1985; Blanchflower et al., 1996) which makes them unable to absorb short-term losses associated with the damage caused by storms.

Workers value consumption of wages w and leisure (the complement of hours worked h). So $U_h < 0$, $U_{hh} > 0$ while $U_w > 0, U_{ww} < 0$. If workers are laid off they do not find alternative employment immediately, they earn only income from outside options given here by $\overline{w}$.[27] In this setting this might correspond to going back to work in agriculture. Worker's expected utility, conditional on the realization of the state of the world, is given by:

$$EU(\theta) = pU(w, h) + (1 - p)U(\overline{w}, 0) \tag{3.4}$$

So firms offer contracts that specify wages, hours and the probability of employment for workers, $(w(\theta), h(\theta), p(\theta))$, for each realization of θ. For ease of exposition, from here on we write each endogenous variable without specify-

26 The results are not significantly altered by this assumption (see Miyazaki and Neary (1985) for a similar argument presented here with indemnity pay for laid off workers), but it is likely to be true in our setting, and it makes the exposition considerably simpler. The standard literature has paid attention to whether employment is classified as involuntary or voluntary: that is, whether laid off workers would prefer to be employed. In many of the models employed is voluntary: workers paid less for during paid states of the world but are compensated by the utility of additional leisure, since they work fewer hours. Version of the standard model where firms are either similarly risk averse (Blanchflower et al., 1996) or face credit constraints such that profits are bounded at zero (or some other negative lower bound), produce outcomes where employment is involuntary. We largely ignore this question, and study a scenario where employment is involuntary by construction.

27 This assumption is particularly likely to hold after large shocks when new jobs are unlikely to be available in abundance.

ing it is a function of θ, (w, h, p). Workers face the risk of being laid off with probability $(1 - p)$.

In this model, firms compete for workers, driving up offers made to workers until firms push up against a probability constraint given by:

$$E\nu(\pi) = \bar{\nu} \tag{3.5}$$

Thus the optimal contract problem is solved by the constrained maximization of expected worker utility $Eu(\theta)$ with Lagrange multipliers for (1) firms profit constraint (λ) and (2) the total labour constraint $p \leqslant 1$ (η).[28] This second constraint is important: when it is binding at the optimal contract ($\eta > 0$) firms do not lay off workers.

This optimization problem yields the following F.O.C's for w, h, and p, respectively:

$$U_1'(w, h) = \lambda\nu'(\pi)n \tag{3.6}$$

$$pU_2'(w, h) + \lambda\nu'(\pi)\theta f_2'(pn, h) = 0 \tag{3.7}$$

$$\eta = \lambda\nu'(\pi)[\theta n f_1'(pn, h) - wn] + U(w, h) - U(\overline{w}, 0) \tag{3.8}$$

Equation 3.6 expresses how wages react to economic shocks through risk-sharing between workers and firms in a manner similar to the result in (Blanchflower et al., 1996). When firms are very risk averse, workers accept large falls in wages in exchange for higher wages in normal periods. So the more risk averse are firms, the stronger the downward wage adjustment. However, firms could insure workers against layoffs at the same time, especially if workers are particularly risk averse at low levels of consumption due to subsistence constraints. This would increase the sensitivity of wages to shocks, while employment lev-

28 Expected utility and profit is of course given by integrating over the distribution of realizations for θ. $Eu(\theta) = \int[p(\theta)U(w(\theta), h(\theta)) + (1-\theta)U(\overline{w}, 0)]dG(\theta)$. We do not specify the distribution of shocks $G(\theta)$.

els remain constant. So workers accept a lower probability of unemployment in exchange for lower wages when shocks hit.[29]

Equation 3.6 shows an important insight: when firms are risk neutral ($\nu'(\pi) = 1$) wages respond to shocks to θ only if hours do, and if hours work impact on the marginal utility of consumption (non-separability) so that $U_{wh}! = 0$. In this way workers are paid less when they are working less because the marginal utility of consumption falls when they have more leisure (when $U_{wh} > 0$). The results in this paper show that for permanently employed workers in the private sector, hourly wages fall dramatically without commensurate reductions in hours worked. This suggests that risk sharing is an important part of the results in this paper, since the magnitude of reductions in wages cannot be explained by substitution between consumption and leisure alone.

Layoffs and Work-Sharing

Wage adjustments moderate the impact of shocks on labour demand. However, when labour demand falls, as it does in most of our empirical results, we seek to understand the relationship between changes in hours worked and layoffs. For ease of exposition, but without loss of generality, we put aside the issue of risk sharing from this point on. We assume that $\nu'(\pi) = 1$: firms are risk neutral. We focus instead of the "work-sharing" mechanisms which determine the trade off between hours per worker and employment.[30]

29 Worker risk aversion pushes in the other direction: workers bargain for wages to remain relatively constant (conditional on hours worked remaining constant) in exchange for lower average wages.

30 Wages are still state dependent in this case due to adjustments to hours worked, as the point in the previous paragraph makes clear.

The trade off between hours worked and layoffs is captured by the second and third FOCs. Recall that $U_2'(w,h) < 0$. We re-arrange Equation 3.7 and substitute for λ from Equation 3.6:

$$\begin{aligned} \theta f_2'(pn,h) &= -\frac{pU_2'(w,h)}{\lambda} \\ \theta f_2'(pn,h) &= -\frac{npU_2'(w,h)}{U_1'(w,h)} \end{aligned} \tag{3.9}$$

Do firms adjust down the hours worked per worker h (work-sharing) or reduce employment p (lay-offs) in response to bad realizations of θ? This is determined by the value of η for the optimal contract. Miyazaki and Neary (1985) show that a precondition for layoffs is that $\eta < 0$ when $p = 1$. After all, if the optimal outcome is full employment ($p^* = 1$), then $\eta > 0$. But if layoffs occur, the optimal value for p^* lies on $0 < p < 1$ and $\eta = 0$. This implies that at $p = 1$ then $\eta < 0$. In other words, if firms were "forced" to keep full employment when the optimal solution has $p < 1$, the marginal product of additional employment would be less than the marginal costs (the wage bill and the foregone leisure of those workers) and firms wish to make layoffs.

The expression for 3.8 is a surprisingly tractable. Firstly we re-arrange, and add subtract terms:

$$\begin{aligned} \eta = \lambda n[\theta f_1'(pn,h) - \frac{h\theta f_2'(pn,h)}{pn} - \overline{w}] \\ + U(w,h) - U(\overline{w},0) - (w-\overline{w})\lambda n + \frac{\lambda h\theta f_2'(pn,h)}{p} \end{aligned} \tag{3.10}$$

Then substituting from 3.9 and 3.6:

$$\eta = \lambda n[\theta f_1'(pn, h) - \frac{h\theta f_2'(pn, h)}{pn} - \overline{w}]$$
$$+ U(w, h) - U(\overline{w}, 0) - (w - \overline{w})U_1'(w, h) + hU_2'(w, h) \tag{3.11}$$
$$\eta = \lambda n[\theta f_1'(pn, h) - \frac{h\theta f_2'(pn, h)}{pn} - \overline{w}] + H(w, h) \tag{3.12}$$

The second part of 3.11 we denote with $H(w, h)$. $H(w, h)$ is strictly positive, by the concavity of U. Layoffs occur when $\eta < 0$ at $p = 1$, when Expression 3.12 is negative. Thus a necessary, but not sufficient, condition for layoffs is:

$$n[\theta f_1'(pn, h) - \overline{w}] < h\theta f_2'(pn, h) \tag{3.13}$$

The LHS of expression 3.13 shows the marginal product of employment at the extensive margin, the RHS the marginal product of employment at the intensive margin. If the latter is larger than the former, firms would prefer to lay off workers and increase hours.

So layoffs are more likely when $\overline{w}$ is larger: workers have better outside options and thus are more tolerant of layoffs. This result is similar to Baily (1977) who argues that unemployment insurance can encourage layoffs. Similarly, when workers are less risk averse so that $H(w, h)$ is smaller, layoffs are more likely to occur.

In the case where workers have no outside earnings options the expression reduces to $n\theta f_1'(pn, h) < h\theta f_2'(pn, h)$. So layoffs occur only if the marginal product of increased hours is sufficiently large relative to marginal product of additional labour.

Divisibility of Labour

In the limit case where labour is perfectly divisible, firms production becomes $f(pn, h) = f(pnh)$. Hours per worker and additional workers are perfect substitutes. This production function with divisible labour is used in Stiglitz (1986). In this case $f_1'(pn,h) = f'(.)h$, and $f_2'(pn,h) = f'(.)pn$. Therefore $h\theta f_2'(pn,h) = n\theta f_1'(pn,h)$, so these terms cancel and η becomes, at $p = 1$:

$$\begin{aligned}\eta &= -\lambda n\overline{w} + H(w,h) \\ &= U(w,h) - U(\overline{w},0) - (w)U_1'(w,h) + hU_2'(w,h) \qquad (3.14)\end{aligned}$$

Firms make lay-offs depending on the opportunity cost of employment: the outside wage option. Notice that if $\overline{w} = 0$ layoffs never occur.[31] This logic explains why the case for layoffs depends on the divisibility of labour. Following Rosen (1985), production is written as:

$$f(np,h) = f(np\gamma(h)) \qquad (3.15)$$

where $\gamma(h)$ is often assumed to be ogive shaped: at low numbers of hours per worker returns to hours are small due to fixed costs of worker days. This could be the case if the first few hours of the work day are dedicated to setting up or preparation before productive activities start. Then returns increase rapidly for intermediate values of h and then begin to suffer diminishing marginal returns as workers fatigue during the course of the day.

With this production function, the first order condition for p becomes:

$$\eta = \lambda n[\theta f'(.)\gamma(h) - h\theta f'(.)\gamma'(h) - \overline{w}] + H(w,h) \qquad (3.16)$$

31 If workers are indifferent between employment and unemployment, such that $U(w,h) = U(\overline{w},0)$, then layoffs definitely do occur. This is similar to the result in (Rosen, 1985), where firms provide full insurance to laid off workers, such that they are indifferent between employment and unemployment. In that model, by introducing indivisibility in labour (such that the returns to additional hours per work are decreasing for higher values of h) for sufficiently low θ, firms *only* layoff workers, and hours are constant.

Again with $\overline{w} = 0$, layoffs happen only if:

$$\gamma(h)/h < \gamma'(h) \tag{3.17}$$

This says, of course, that when the marginal returns to hours are higher than the average returns to hours, firms prefer to keep hours constant high and employ fewer (more) workers in response to bad (good) realizations of θ. Given the assumption of the ogive shape of γ, there are many points along $\gamma(h)$ at which this holds. However, beyond a certain point, diminishing marginal returns mean that firms prefer to cut workers' hours rather than lay them off.

The impact of storms on hours is about 3.5%. If average hours are about 48 in a 'normal' period (where $p = 1$), they fall to only about 46.4 hours when shocks hit. Very specific conditions on the the slope of γ would have to prevail to result in a switch of sign of $\gamma(h)/h - \gamma'(h)$ on the range 46.4-48. The second FOC in hours (equation 3.9) with this production function becomes:

$$\theta f'(.)\gamma'(h) = \frac{U_2'(w,h)}{U_1'(w,h)} \tag{3.18}$$

The optimal outcome for h need not be close to an inflection point where $\gamma(h)/h = \gamma'(h)$. Indeed if decreasing returns to hours per worker take a long time to kick in, implying that labour is divisible for reasonably high levels of h, then firms will prefer to reduce hours rather than lay-off workers.

Recall that we are talking about necessary but not sufficient condition for layoffs. With low $\overline{w}$, $H(w,h)$ get very large which makes layoffs less likely, even when labour is relatively indivisible.

3.6.2 *Discussion*

The aim of this framework is not to argue that layoffs do or do not occur in optimal contract models. Indeed without strong assumptions on the functional forms of $U(w,h)$ and $f(np,h)$, these models can say little more than $dp/d\theta \geqslant 0$ and $dh/d\theta \geqslant 0$ (Rosen, 1985). Instead we have made a case for work-sharing as a way of insuring workers against risk (especially when severance pay is not made). The results presented here suggest that there are parameter values under which adjustment in hours can dominate adjustments in employment.

Secondly, we have shown that trade off between work-sharing (reduction in hours) versus is layoffs is determined by three key factors. Firms are more likely to reduce hours and maintain full-employment if 1) workers are more risk averse, 2) workers outside options are worse, and 3) if labour is relatively divisible. These findings are similar to those in Azariadis (1975).

Our empirical results show large adjustment in wages and hours, and few layoffs. We argue that these findings are not surprising in light of the model: workers may well be very risk averse when their entire livelihoods are based on their wage earnings, and outside options may be made considerably worse when storms hit, because of the damage caused to home production and own-farm agriculture. We have no direct evidence on the divisibility of labour, but argue that our results suggest that the firms are relatively willing to reduce workers' hours.

Yet this illuminates an important point. It may be the case that labour is highly indivisible, but that the high risk aversion of workers means firms are cutting hours and wages to protect workers from lay-offs. This would imply inefficient levels of hours relative to the case where workers are fully insured and firms can adjust optimally by reducing the size of their labour forcing but keeping hours high. This again mirrors the argument in Rosen (1985). Markets for either private or public insurance for workers would considerably improve the efficiency of outcomes after storms hit.

The model also illuminates the role of labour supply. The extent of flexility of hours is in part due to workers' preference for leisure (or time off work for home production). In our setting we have argued that workers may have a particularly strong preference for more time off work during times when the storms hit, in order to spend time repairing damages caused by storms.

However, their outside options are still poor, and in fact may be particularly poor in periods where storms hit because of storm destruction of farming or other consumption generating activities at home. This limits labour supply elasticity at the extensive margin. In this way workers are willing to sacrifice hours at the intensive margin, and therefore wages, as governed by the relationship given in Equation 3.6, in order to avoid the chance of being laid off. We have no direct evidence for this phenomenon of increased labour supply elasticity during storms, but this mechanism is consistent with the results of Jayachandran (2006).

This paper has not considered dynamics considerations that could be contributing to our finding of no lay-offs. That is, we have not assumed that firms have a preference to 'hoard' labour, which would be the case if there were adjustment costs associated with hiring or firing labour (Bloom, 2009), or if there were job-specific returns to human capital (Hashimoto, 1981). Adding these elements to the model would strengthen our results, by making firms less willing to lay off workers.

3.7 CONCLUSION

In this paper, taking advantage of a unique individual-level labor force dataset spanning 26 quarters between 2003 and 2009, we explore how labor markets adjust to large economic shocks, namely strong typhoons. Our results suggest that employment levels are unaffected but nominal weekly wages adjust downwards, through a combination of lower hours and lower hourly wage. The effects are driven by individuals employed on permanent contracts in the private sector and dissipate shortly after the storms hit.

The results have implications for our understanding of labor markets in developing countries. First, there is evidence of flexibility in established long-term contractual relationships which is consistent with theories of implicit contracts. Second, the adjustments take place along the intensive rather than extensive margin which we interpret as risk-sharing between the firms and the workers. This built-in insurance mechanisms seem to be indicative of sophisticated informal arrangements for coping with large economic shocks. In contexts where social safety nets might not be adequate, utility loss associated with unemployment are likely large and it appears that considerable risk sharing occurs between firms and workers, as well as among workers in the form of work-sharing. Third, our results are obtained in a context where typhoons are relatively common and so could be thought of as an adaptive response to repeated natural disaster shocks. Fourth, managers increase their working hours to respond to the shocks which is indicative that adequate management is an important component of a firm's ability to deal with the shocks.

BIBLIOGRAPHY

Abraham, K. G. and Haltiwanger, J. C. (1995). Wages and the business cycle. *Journal of Economic Literature*, 33(3):1215–1264.

Anttila-Hughes, J. and Hsiang, S. (2013). Destruction, disinvestment, and death: Economic and human losses following environmental disaster. *UC Berkeley, mimeo*.

Azariadis, C. (1975). Implicit contracts and underemployment equilibria. *The Journal of Political Economy*, pages 1183–1202.

Azariadis, C. and Stiglitz, J. E. (1983). Implicit contracts and fixed price equilibria. *The Quarterly Journal of Economics*, pages 2–22.

Baily, M. N. (1974). Wages and employment under uncertain demand. *Review of Economic Studies*, 41(1):37–50.

Baily, M. N. (1977). On the theory of layoffs and unemployment. *Econometrica: Journal of the Econometric Society*, pages 1043–1063.

Bankoff, G. (2002). *Cultures of Disaster: Society and Natural Hazard in the Philippines*. Routledge.

Beaudry, P. and Dinardo, J. (1991). The effect of implicit contracts on the movement of wages over the business cycle : Evidence from micro data. *The Journal of Political Economy*, 99(4):665–688.

Bils, M. J. (1985). Real wages over the business cycle: Evidence from panel data. *Journal of Political Economy*, 93(4):666.

Blanchflower, D. G., Oswald, A. J., and Sanfey, P. (1996). Wages, profits and rent-sharing.

Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, 77(3):623–685.

Burdett, K. and Mortensen, D. T. (1980). Search, layoffs, and labor market equilibrium. *Journal of Political Economy*, 88(4):652.

Cameron, C., Gelbach, J., and Miller, D. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? The New climate-economy literature. *Journal of Economic Literature*, 52(3):740–798.

DeLong, J. B. and Summers, L. H. (1986). Are business cycles symmetric? *NBER Working Paper*, (w1444).

Feldstein, M. (1976). Temporary layoffs in the theory of unemployment. *Journal of Political Economy*, 84(5):937.

Gignoux, J. and Menéndez, M. (2014). Benefit in the wake of disaster: Long-run effects of earthquakes on welfare in rural Indonesia.

Hall, R. E. and Milgrom, P. R. (2008). The limited influence of unemployment on the wage bargain. *American Economic Review*, 98:1653–1674.

Hashimoto, M. (1981). Firm-specific human capital as a shared investment. *The American Economic Review*, pages 475–482.

Holland, G. J. (1980). An analytic model of the wind and pressure profiles in hurriances. *Monthly Weather Review*, 108:1212–1218.

Holmstrom, B. (1983). Equilibrium long-term labour contracts. *Quarterly Journal of Economics*, 98(1983):23–54.

Holzer, H. J. and Montgomery, E. B. (1993). Asymmetries and rigidities in wage adjustments by firms. *The Review of Economics and Statistics*, 75(3):397–408.

Hsiang, S. and Jina, A. (2014). The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones. *NBER Working Paper 20352*.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Kaur, S. (2014). Nominal wage rigidity in village labor markets. *Working Paper, Columbia University*, (October):1–66.

Keane, M., Moffitt, R., and Runkle, D. (1988). Real wages over the business cycle: estimating the impact of heterogeneity with micro data. *Journal of Political Economy*, 96(6):1232.

Keane, M. P. and Prasad, E. S. (1996). The employment and wage effects of oil price changes: A sectoral analysis. *The Review of Economics and Statistics*, 78(3):389–400.

Kirchberger, M. (2014). Natural disasters and labor markets. *Oxford University, CSAE Working Paper WPS/2014-19*.

Kochar, A. (1999). Smoothing consumption by smoothing income: hours-of-work responses to idiosyncratic agricultural shocks in rural India. *Review of Economics and Statistics*, 81(February):50–61.

Labonne, J. (2014). Local political business cycles. evidence from philippine municipalities. *University of Oxford, mimeo*.

Miyazaki, H. and Neary, H. M. (1985). Work hours and employment in the short run of a labour-managed firm. *The Economic Journal*, 95(380):1035–1048.

Moretti, E. (2010). Local multipliers. *The American Economic Review*, pages 373–377.

Mortensen, D. T. (1978). On the theory of layoffs. Technical report, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Rogerson, R. and Shimer, R. (2011). Search in macroeconomic models of the labor market. *Handbook of Labor Economics*, 4:619–700.

Rosen, S. (1985). Implicit contracts: AsSurvey. *Journal of Economic Literature*, 23(3):1144–1175.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49.

Skidmore, M. and Toya, H. (2002). Do natural disasters promote long-run growth? *Economic Inquiry*, 40(4):664–687.

Stiglitz, J. E. (1986). Theories of wage rigidity. Technical report, National Bureau of Economic Research.

Ugaz, J. and Zanolini, A. (2011). Effects of extreme weather shocks during pregnancy and early life on later health outcomes: the case of philippines typhoons. *University of Chicago Harris School of Public Policy, mimeo*.

APPENDICES FOR CHAPTER 3

Table C.1: Aggregate-level results (income per capita): Alternative storm measures

	(1) inc/ adult	(2) inc/ adult	(3) inc/ adult	(4) inc/ adult
Wind-speed (knots)	-0.00027**			
	(0.000)			
Normalized Wind-speed (0-1)		-0.076***		
		(0.025)		
ss scale 1			-0.007	
			(0.012)	
ss scale 2			-0.009	
			(0.015)	
ss scale 3			-0.025	
			(0.015)	
ss scale 4			-0.076***	
			(0.023)	
ss scale 5			-0.118	
			(0.072)	
Big Storm				-0.078***
				(0.024)
Small Storm				-0.012
				(0.009)
Observations	20,808	20,808	20,808	20,808
R-squared	0.072	0.072	0.073	0.073
Mean Dep. Var	5.400	5.400	5.400	5.400

Notes: Results from weighted municipal*quarter regressions. The dependent variable is the log of total income per capita for the municipality. Regressions control for municipal fixed effects, region-specified time fixed effects) as well as the share of the working age population in each education category, the share of women in the working age population, the number of men, the number of women, the number men age 15-30 and the number of women age 15-30. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.2: Aggregate-level results (employment): Alternative storm measures

	(1) employed	(2) employed	(3) employed	(4) employed
Wind-speed (knots)	0.000			
	(0.000)			
Normalized Wind-speed (0-1)		-0.005		
		(0.005)		
ss scale 1			0.001	
			(0.003)	
ss scale 2			0.006	
			(0.005)	
ss scale 3			-0.007*	
			(0.003)	
ss scale 4			-0.007*	
			(0.004)	
ss scale 5			-0.007	
			(0.007)	
Big Storm				-0.007*
				(0.004)
Small Storm				0.000
				(0.002)
Observations	21,064	21,064	21,064	21,064
R-squared	0.021	0.021	0.021	0.021
Mean Dep. Var	0.600	0.600	0.600	0.600

Notes: Results from weighted municipal*quarter regressions. The dependent variable is the employment rate in the municipality. Regressions control for municipal fixed effects, region-specified time fixed effects) as well as the share of the working age population in each education category, the share of women in the working age population, the number of men, the number of women, the number men age 15-30 and the number of women age 15-30. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.3: Individual-level results: persistence

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Impact of Lagged Storms on Earnings and Hours						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm						
current	-0.022**	-0.010	-0.014	-0.008	-0.014*	0.000
	(0.011)	(0.009)	(0.009)	(0.008)	(0.007)	(0.004)
lag 1	-0.012	0.005	-0.014	0.001	-0.009	-0.005
	(0.014)	(0.010)	(0.011)	(0.009)	(0.009)	(0.005)
lag 2	0.004	0.026***	0.016*	-0.012	0.008	0.007*
	(0.011)	(0.009)	(0.009)	(0.010)	(0.008)	(0.004)
lag 3	-0.018	-0.005	-0.010	-0.008	-0.007	-0.003
	(0.011)	(0.010)	(0.010)	(0.009)	(0.008)	(0.005)
Small Storm (lags estimated but not displayed)						
current	-0.005	-0.003	-0.002	-0.002	0.000	-0.002
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	860,809	2,006,022	860,809	860,809	860,809	860,809
R-squared	0.444	0.130	0.092	0.419	0.090	0.040
Panel B: Impact on Lagged Storms on Employment (Extensive Margins)						
	employed	job	wage missing	wage observed	zero hours	lost job quarter
Big Storm						
current	-0.006	-0.005	0.006	-0.006*	0.001	0.001
	(0.004)	(0.004)	(0.006)	(0.004)	(0.001)	(0.002)
lag 1	-0.003	-0.006*	0.004	-0.003	-0.003***	-0.002
	(0.004)	(0.004)	(0.006)	(0.004)	(0.001)	(0.002)
lag 2	0.000	-0.004	-0.014**	0.006*	-0.004***	-0.001
	(0.004)	(0.004)	(0.006)	(0.004)	(0.001)	(0.002)
lag 3	-0.004	-0.005	0.006	-0.006	-0.001	0.000
	(0.004)	(0.004)	(0.006)	(0.004)	(0.001)	(0.002)
Small Storm (lags estimated but not displayed)						
current	-0.001	-0.001	0.001	-0.001	0.000	-0.002**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.001)
Sample	All	All	Earners	All	All	All
Observations	3,402,456	3,402,456	2,006,018	3,402,456	3,402,456	3,402,456
R-squared	0.228	0.238	0.197	0.105	0.015	0.021
Mean Dep. Var	0.600	0.600	0.500	0.300	0.000	0.000

Notes: Results from weighted individual regressions. In Panel A, the dependent variable is the log weekly wage for employed individuals (Column 1), number of hours worked for employed individuals (Column 2), number of hours worked for employed individuals earning a wage (Column 3), hourly wage for employed individuals (Column 4), number of days worked for employed individuals earning a wage (Column 5), number of hours worked per day for employed individuals earning a wage (Column 6). In Panel B, the dependent variables are a series of dummies equal to one if: the individual is employed (Column 1), the individual has a job (Column 2), the individual is employed but their wage is not observed (Column 3), the individual reports a wage regardless of employment status (Column 4), the individual reports having a job but working zero hours in the last 7 days (Column 5), the individual reports not having a job now, but having worked in the last 3 months (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.4: Panel-level results: Employment in different types of jobs

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total Effect (Unconditional on having a job)						
	selfemp	permpriv	temppriv	ownfarm	wagefarm	gov
Big Storm	-0.003	-0.002	-0.000	0.001	0.001	-0.002
	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)
Small Storm	-0.000	-0.002	0.001	0.000	-0.001	-0.001
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Observations	1,294,842	1,294,842	1,294,842	1,294,842	1,294,842	1,294,842
R-squared	0.017	0.059	0.017	0.196	0.082	0.015
Mean Dep. Var	0.141	0.161	0.054	0.148	0.051	0.048
Panel A: Composition Effect (Conditional on having a job)						
	selfemp	permpriv	temppriv	ownfarm	wagefarm	gov
Big Storm	-0.003	-0.004	-0.000	0.005	0.002	-0.003
	(0.004)	(0.004)	(0.003)	(0.006)	(0.005)	(0.003)
Small Storm	-0.000	-0.003	0.002	0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)
Observations	805,430	805,430	805,430	805,430	805,430	805,430
R-squared	0.040	0.144	0.036	0.263	0.118	0.026
Mean Dep. Var	0.230	0.263	0.089	0.241	0.084	0.078
Panel C: Composition Effect (Conditional on earning a wage)						
	selfemp	permpriv	temppriv	ownfarm	wagefarm	gov
Big Storm	0.001	-0.003	0.004	-0.001	0.007	-0.007
	(0.001)	(0.009)	(0.007)	(0.001)	(0.007)	(0.005)
Small Storm	-0.000	-0.006	0.008**	0.001	-0.001	-0.001
	(0.001)	(0.004)	(0.004)	(0.001)	(0.004)	(0.003)
Observations	396,552	396,552	396,552	396,552	396,552	396,552
R-squared	0.005	0.148	0.039	0.044	0.293	0.066
Mean Dep. Var	0.004	0.502	0.170	0.002	0.160	0.149

Notes: Results from weighted individual regressions. The dependent variable is a dummy equal to one if the individual is: self-employed (Column 1), has a permanent job in the private sector (Column 2), has a temporary job in the private sector (Column 3), works on the family farm (Column 4), works for a wage on someone's else farm (Column 5), is employed in the public sector (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.5: Panel results: Comparison of municipal and individual fixed effects (Decomposition)

	(1)	(2)	(3)	(4)
Panel A: All Employees				
	wage/ week	wage/ week	wage/ week	wage/ week
Big Storm	-0.019**	-0.022**	-0.031**	-0.034***
	(0.009)	(0.010)	(0.012)	(0.012)
Small Storm	-0.008	-0.009*	-0.006	-0.010
	(0.005)	(0.005)	(0.006)	(0.006)
Observations	349,605	267,038	349,605	267,038
R-squared	0.021	0.022	0.460	0.465
FE	Ind	Ind	Muni	Muni
Mindanao	Yes	No	Yes	No
Panel B: All Employees with similar jobs				
	wage/ week	wage/ week	wage/ week	wage/ week
Big Storm	-0.023**	-0.027**	-0.026	-0.031*
	(0.011)	(0.011)	(0.017)	(0.017)
Small Storm	-0.007	-0.011*	0.002	-0.002
	(0.006)	(0.006)	(0.008)	(0.008)
Observations	163,043	125,078	163,043	125,078
R-squared	0.020	0.021	0.519	0.523
FE	Ind	Ind	Muni	Muni
Mindanao	Yes	No	Yes	No

Notes: Results from weighted panel regressions. The dependent variable is the average weekly wage. Regressions control for individual fixed effects, region-specified time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.6: Panel-level results: Employment

	(1) employed	(2) job	(3) wage missing	(4) wage observed	(5) zero hours	(6) lost job quarter
Big Storm	-0.005	-0.004	0.009*	-0.007**	0.003	0.005
	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.005)
Small Storm	0.001	0.001	0.003	0.000	0.001	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Observations	1,294,842	1,294,842	792,550	1,294,842	805,430	489,412
R-squared	0.002	0.002	0.002	0.001	0.001	0.013
Mean Dep. Var	0.603	0.612	0.536	0.283	0.015	0.058

Notes: Results from weighted individual regressions. The dependent variables are a series of dummies equal to one if: the individual is employed (Column 1), the individual has a job (Column 2), the individual is employed but their wage is not observed (Column 3), the individual reports a wage regardless of employment status (Column 4), the individual reports having a job but working zero hours in the last 7 days (Column 5), the individual reports not having a job now, but having worked in the last 3 months (Column 6). Regressions control for time fixed effects as well as municipal fixed effects (Panel A) and individual fixed effects (Panel B). In Panel A, regression control for the respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.7: Aggregate-level decomposition: Heterogeneity for rural-urban areas

	(1) inc/ adult	(2) wage/ week	(3) wage/ hour	(4) hours/ earner	(5) earners/ job	(6) job/ adult
Big Storm	-0.080***	-0.038**	-0.020	-0.018*	-0.033	-0.009
	(0.024)	(0.016)	(0.012)	(0.009)	(0.021)	(0.007)
Big Storm * city	0.012	0.014	-0.002	0.016	0.012	-0.015
	(0.046)	(0.028)	(0.017)	(0.014)	(0.034)	(0.012)
Small Storm	-0.020*	-0.015**	-0.013**	-0.002	-0.002	-0.002
	(0.010)	(0.007)	(0.006)	(0.004)	(0.009)	(0.004)
Small Storm * city	0.026**	0.006	0.005	0.001	0.015	0.005
	(0.013)	(0.008)	(0.007)	(0.005)	(0.010)	(0.007)
Denominator	Adults	Earners	Earned Hours	Earners	Jobs	Adults
Observations	20,808	20,808	20,808	20,808	20,831	21,064
R-squared	0.073	0.131	0.146	0.068	0.024	0.016

Note: esults from weighted municipal*quarter regressions. The dependent variable is the average income from employment per adult (Column 1), the average income from employment for employed individuals (Column 2), the average hourly wage for employed individuals (Column 3), the average number of hours worked for employed individuals (Column 4), the proportion of individuals who had jobs who reported a salary (Column 5), the proportion of adults who had jobs (Column 6). Regressions control for municipal fixed effects, time fixed effects as well as the share of the working age population in each education category, the share of women in the working age population, the number of men, the number of women, the number men age 15-30 and the number of women age 15-30. The sample is restricted to municipalities outside of Mindanao. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.8: Impacts in levels: Comparison between individual and aggregated results

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Main Impacts in Levels for Aggregated Data						
	inc/ adult	wage/ worker	wage/ earner	hours/ adult	hours/ worker	hours/ earner
Big Storm	-12.796***	-13.102*	-20.872	-0.637**	-0.505	-0.586
	(4.396)	(7.114)	(14.499)	(0.241)	(0.313)	(0.361)
Small Storm	2.777	9.236*	3.832	-0.164	-0.182	-0.142
	(2.990)	(5.232)	(7.519)	(0.119)	(0.117)	(0.141)
Observations	21,064	21,064	20,831	21,064	21,064	20,831
R-squared	0.143	0.149	0.169	0.048	0.052	0.074
Mean Dep. Var	383.225	700.562	1,280.171	24.139	42.622	43.190
BStorm as % of Mean	-0.030	-0.024	-0.012	-0.022	-0.014	-0.009
Panel B: Main Impacts in Levels for Individual Data						
	inc/ adult	wage/ worker	wage/ earner	a hours	hours/ worker	hours/ earner
Big Storm	-11.779**	-11.350	-15.498	-0.619***	-0.609**	-0.564*
	(4.599)	(7.721)	(12.913)	(0.204)	(0.269)	(0.295)
Small Storm	4.259	11.619***	9.605	-0.137	-0.175	-0.056
	(2.731)	(4.380)	(7.026)	(0.105)	(0.129)	(0.133)
Observations	2,464,172	1,439,415	669,711	2,464,172	1,453,620	669,711
R-squared	0.061	0.167	0.174	0.013	0.110	0.072
Mean Dep. Var	391.800	680.000	1,370.700	24.100	41.500	44.700
BStorm as % of Mean	-0.030	-0.017	-0.011	-0.026	-0.015	-0.013

Notes: Results from weighted individual regressions. The dependent variables are: the income per adult in the sample. This is the total income divided by the total number of adults (Column 1), the wage per worker- the total wages divided by the total number of workers (Column 2), the wage per worker for whom a wage is observed (Column 3), hours per adult- the total hours worked divided by the number of adults (Column 4), total hours over the number of workers (Column 5) and the hours per worker for whom a wage is observed (Column 6). Regressions control for municipal fixed effects, region-specified time fixed effects as well as respondent's age, age square, education levels and gender. The standard errors (in parentheses) account for potential correlation within province. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table C.9: Panel-level results: Decomposition for workers who stay at similar jobs

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Impact on Earnings and Hours (Same Job Characteristics)						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm	-0.025**	-0.020**	-0.021**	-0.004	-0.014	-0.008*
	(0.012)	(0.009)	(0.010)	(0.008)	(0.009)	(0.004)
Small Storm	-0.011*	-0.007	-0.007	-0.004	-0.006	-0.003
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	157,273	410,445	157,963	157,273	157,962	157,962
R-squared	0.020	0.005	0.011	0.018	0.014	0.001
Panel B: Impact on Earnings and Hours (Same Job Characteristics and Payment Type)						
	wage/ week	hours/ worker	hours/ earner	wage/ hour	days/ earner	hours/ day
Big Storm	-0.027**	-0.021**	-0.021**	-0.006	-0.017*	-0.005
	(0.011)	(0.009)	(0.009)	(0.008)	(0.009)	(0.004)
Small Storm	-0.011*	-0.003	-0.003	-0.008*	-0.004	0.001
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)
Sample	Earners	All	Earners	Earners	Earners	Earners
Observations	125,078	125,098	125,087	125,078	125,087	125,087
R-squared	0.021	0.014	0.014	0.020	0.016	0.001

Notes: Results from weighted individual fixed-effects regressions. Panel A shows results for individuals who are working in at least two periods of the data, for who remain working at jobs of the same job type. Panel B shows results for workers whose stay at jobs that look identical in terms of job type, occupation, type of employer and method of payment. The dependent variable is the log weekly wage for employed individuals (Column 1), number of hours worked for employed individuals (Column 2), number of hours worked for employed individuals earning a wage (Column 3), hourly wage for employed individuals (Column 4), number of days worked for employed individuals earning a wage (Column 5), number of hours worked per day for employed individuals earning a wage (Column 6). Regressions control for time fixed effects and individual fixed effects. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

CONCLUSION

This thesis is concerned with the ways in which labour markets in developing countries operate differently because of the constraints of the physical environments of participants in these markets. Chapter One showed that cash constraints and costs of job search due to where people live can inhibit their ability to search for work. Chapter Two showed that poor housing conditions can limit the ability of individuals to participate in the labour force. Chapter Three showed that natural disasters have large impacts on economic activity, and that the risk of these disasters leads to complex informal risk-sharing relationships in labour relations.

DIRECTIONS FOR FURTHER RESEARCH

The first chapter on job search in Ethiopia offers a number of fruitful avenues for future work. In the short run, I plan to use data from a new expanded and long-term phone call survey currently underway in Addis Ababa, to further test my dynamic job search model. I will be able to estimate and validate this model structurally and compare it to the performance of more standard search-theoretic models. That structural model can then be used to predict the size of the impacts of a new transport subsidy experiment underway in Addis Ababa. The model could be also be used to predict the effect of other job search interventions, such as an unconditional cash grants or unemployment insurance on job search outcomes.[1]

1 An on-going field experiment similar to mine in this paper is ongoing in South Africa (Banerjee and Sequeira, forthcoming) which compares the impact of conditional (transport) and unconditional cash transfers on employment outcomes. It will be interested to see if these results conform with the predictions of the model and my empirical results.

Screening and Signalling Frictions

More research is required to understand the role of information asymmetries in job search and matching in urban labour markets. Discussions with job seekers and firms suggest that it is hard for firms to screen respondents and for applicants to signal their skills to firms. Were the job search subsidies to be rolled out on a larger scale the full effects of the program could be severely limited if the increase in applications adds considerably to noise in the market. Improved information flows may be required to ensure that firms and workers are able to make good matches as the pool of vacancies and candidates grows.

To pursue this line of inquiry, co-authors[2] and I have initiated a trial into the effects of screening on the labour market in Addis Ababa. We have provided CV and job application training to 1300 job seekers to better help them to signal their ability in the job market. In addition respondents received a series of detailed aptitude, cognitive and non-cognitive tests. The trial is complete and the results will be available after the endline in June 2015. We will be able to use our detailed test results to make statements about the skills that are most important for the job market, as well as providing evidence on skills training programmes that would best assist the youth in gaining a foothold in the labour market (Bertrand et al., 2013).

Firms and Match Quality

The evidence from training and certification experiment will provide a test of whether screening can improve the ability of job seekers to find jobs. We will also seek to test whether it impacts the quality of matches. In the standard search theoretic framework (Pissarides, 2000) improved matching expands the size the employment pool (the treatment effects are not displaced). Recent evidence from France suggests that this is not always the case. Crépon et al. (2013) show that

2 This is joint with Girum Abebe, Stefano Cario, Paolo Falco, Marcel Fafchamps, and Simon Quinn

the benefits of scaled up youth job search assistance programs are often at the cost of other job seekers, so that there is no effect in equilibrium.

There are reasons to believe that increasing the pool of job applicants would not lead such displacement effects. For instance, there may be a great many skilled individuals who are genuinely locked out of labour markets by the barriers to entry, such as search costs or certification. There may in fact be many productivity improving matches that aren't happening because of the high costs of making those matches.

I am interested in the effects of an expanded, and better screened, labour pool on firm hiring decisions. I want to test whether improving access to good candidates increases firm hiring, reduces turnover, and increases productivity of matches. To do this coauthors[3] and I have initiated an innovative randomized experiment designed to expose some firms to an improved pool of job applicants. We are running large job fairs in Addis Ababa, where representatives from our survey of 500 large employers in Addis Ababa are randomly invited to job fairs, to which a random group of job seekers are invited. This experiment tests the effect of simply exposing firms to many *average* workers in the labour market at very low costs. Do firms hire from this pool of candidates? Does reducing the costs of meeting candidates allow firms to more carefully select employees from the group that they meet? In a parallel experiment we expose a random subset of firms to job fairs where the only attending job seekers are those that have been through the screening and training curriculum. This will allow us to the test improved signalling in the market on hiring of the firms.[4]

3 This is joint with Girum Abebe, Stefano Cario, Paolo Falco, Marcel Fafchamps, Simon Quinn and Forhad Shilpi

4 The job fairs are complete. Results will be available after the firm endline survey in July 2015.

Urban Transport and Growth

The work on transport subsidies raises broader questions about the impacts of transport costs and connectivity in growing African cities. This year I will be working with the "Urbanisation in Developing Economies Programme" at Oxford University and the London School of Economics. We aim to generate research on the impacts of urban policies, including transport and infrastructure upgrading on the spatial layout of economic activities in African cities. One promising project idea of mine is to evaluate rail and road upgrades in the city of Addis Ababa. A new light railway system has just been completed and is due to be opened later this year.

Al of the data collected for this project is accurately geo-referenced, and baseline samples were taken long before the rail way line was complete. The first endline surveys on both the firm sample and the job-seekers will be conducted shortly after the rail system is complete. We have plans for another round of endlines a year later. This should allow me to estimate the impacts of the rail system on the respondents in those samples. I will extend methods used on my housing paper to use distance from the new railway lines as instruments for the effect of improved transport links. This study will have the added advantage of using detailed urban satellite data to measure concurrent changes in city structure and economic activity, or generate additional instruments for the placement of road investments.

Housing, Neighborhoods and Relocation

My results on the impact of housing in South Africa were strengthened precisely because households were not moved a considerable distance from their original place of living. The literature on place of living and neighborhoods on economic outcomes has a lot of unanswered questions. The impacts of moving to a new area with a different composition of inhabitants and neighborhood characteristics is difficult to identify because so much changes when a household moves (Topa and Zenou, 2014; Chetty and Hendren, 2015).

I am planning to conduct a study of the Ethiopian governments' condominium housing project, whereby households are given flats at highly subsidized rates, in different areas of Addis Ababa.[5] The random assignment of households to housing presents a unique natural experiment to study questions of housing and neighborhood effects.

By carefully evaluating this programme, we hope to contribute to both the literature of the impact of improved housing, distance from jobs, and neighborhood characteristics, while disentangled all three of these channels. The government does not only randomly select individuals to receive housing, they are randomly assigned to housing blocks in different parts of the city. This random assignment of households to housing blocks should induce random variation in the composition of the new neighborhoods living in the new housing blocks. We will gather detailed data on the composition of neighboring households after lottery results are announced.

Firms and Contracts

An interesting result from our paper on Typhoons in the Philippines warrants further attention. We find that the self-employed are likely to be unemployed when storms hit. That is, small firms are likely to go out in times of disaster. Larger firms seem to be able to survive the storms. Very little firm data exists for the Philippines. We are trying to access the Annual Survey of Philippine Business and Industry to see if we can reconcile our labour force data results with firm activities, but also to look at the effects of storms on the concentration of economic activity.

5 This is planned joint with the Alebel Bayrou and Berihu Assefa.

Finally, our work on contracts and labour flexibility in the Philippines raises interesting questions about the nature of contracts and labour relations in developing countries. This is a relatively understudied subject. I am aware no major firm survey from a developing country with detailed information describing the nature of employment contracts, firm recruitment strategies, and detail turnover data. While the work of Bloom and Van Reenen (2010) describes firm management practices in detail, the work on labour relations is less detailed. Our on-going survey of firms in Ethiopia contains extremely detailed questions on hiring and labour, which aims to begin to fill this gap.

BIBLIOGRAPHY

Bertrand, M., Crepon, B., Chuan, A., Haget, R., Mahoney, M., Murphy, D., Naud, A., Powers, S., and Takavarasha, K. (2013). J-PAL Youth Initiative: review paper.

Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *The Journal of Economic Perspectives*, pages 203–224.

Chetty, R. and Hendren, N. (2015). The effects of neighborhoods on children's long-term outcomes: quasi-experimental estimates for the United States. *Unpublished paper*.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Quarterly Journal of Economics*, 128(2):531–580.

Pissarides, C. A. (2000). Equilibrium unemployment theory. *MIT Press Books*.

Topa, G. and Zenou, Y. (2014). Neighborhood versus network effects. *Handbook of Regional and Urban Economics*, 4.