CSAE Working Paper WPS/2012-21

A Model of Comparative Advantage with Matching in the Urban Tanzanian Labour Market

Andrew Kerr*

December 3, 2012

Abstract

In this paper I build an equilibrium search model of the urban Tanzanian labour market that explains the choice between wage and self-employment and the variation in earnings across and within these sectors. Self-employment is very common in urban Tanzania and survey data show both that there are large overlaps in the distribution of earnings in private wage employment and self-employment and that there is little movement between wage and self-employment. This suggests that self-employment represents a worthwhile alternative to wage employment in small, low-productivity firms for the majority of urban Tanzanians, in contrast to the traditional view of African labour markets in which wage employment in small firms and self-employment are lumped together as the informal sector.

1 Introduction

African labour market analysis has traditionally emphasised the existence of large wage differentials between the formal and informal sectors, following Harris and Todaro (1970) and Fields (1975). In this paper I suggest an alternative approach to modeling the urban Tanzanian labour market, focusing instead on the distinction between wage and self-employment and the role of comparative advantage and individual heterogeneity in determining where to work and earnings in each of these sectors.

*DataFirst, University of Cape Town: andrew.kerr@uct.ac.za. I thank my DPhil supervisor Francis Teal, as well as my examiners Margaret Stevens and Andy McKay, for helpful comments. The paper has also benefitted from the comments of participants at the Centre for the Study of African Economies 2011 Conference, Oxford and the Economics Society of South Africa 2011 conference, Stellenbosch.

The awareness of the importance of individual heterogeneity in earnings determination is an important advance in recent labour economics. Mortensen (2003) notes that observable characteristics typically explain only 30% of the variation in earnings, and Heckman (2001), in his Nobel prize lecture, concludes that it has been the collection of micro data which has lead economists to appreciate the "diversity and heterogeneity of economic life." This understanding has also influenced thinking about labour markets in developing countries, with apparent segmentation being argued away by appeals to unobserved heterogeneity (Badaoui et al. (2008)). In a similar vein Falco et al. (2011, pg S24) conclude from their study of earnings data in the urban labour markets of Ghana and Tanzania that "by far the most important factor determining earnings is unobserved market ability." In this paper I attempt to provide a model that is consistent with this conclusion, as well as with further analysis of Tanzanian survey data, and that takes account of the obvious differences between wage and self-employment.

The equilibrium search theory literature suggests that a focus on the differences between wage and self-employment may be important for understanding the workings of the urban labour market in low income countries, like Tanzania. Both ex-ante wage posting and ex-post wage bargaining models suggest workers may earn less than their marginal product due to the monopsony power of firms. Self-employment, which is very common in urban Tanzania, does not, however, involve interaction with a firm and thus the self-employed can capture all the returns to their labour and the return to any capital they have accumulated. This is an important distinction which has not been made in models of African labour markets to date.

This distinction is also important empirically. In section 2 below I show that Tanzanian survey data show that there are limited transitions between wage employment in small firms and self-employment, which have traditionally been lumped together as the informal sector. Tanzania also has large amounts of self-employment and only a small formal sector comprised mainly of public employment and large firm employment. This is further evidence for differentiating, instead, between wage and self-employment. This is in line with Maloney's (2004) conclusion, who suggests that self-employment in the informal sector represents a viable alternative to salaried work either in the formal or informal sector, as a result of the low productivity of firms in many developing countries.

The framework I use to model the Tanzanian labour market and make a distinction between wage and self-employment is equilibrium search theory (Mortensen and Pissarides 1999). Recent research has used the equilibrium search framework to explain features of developing country labour markets. Satchi and Temple (2009) develop a macro matching model of the Mexican labour market that allows for the presence of frictions in the formal sector and assumes their absence in the informal sector, which helps to generate a realistically sized informal sector. Because theirs is a macro model, Satchi and Temple (2009) focus on average wage differentials between sectors, as well as the relative size of the sectors, but cannot explain the distribution of earnings, one of the main goals of my paper. Albrecht et al. (2009) extend the canonical Mortensen and Pissarides (1994) search model to build a two sector model of a Latin American labour market, with both formal and informal sectors. Earnings and ability vary in the formal sector but are constant in informal employment.

Search theory has not only been used by those emphasising heterogeneity in earnings, however. Fields (2004) sugggests that the segmentation literature may have even influenced search theory, since wage posting models can generate earnings differentials across ex ante identical individuals. Ulyssea (2010) provides a two sector matching model of the Brazilian labour market which generates a formal sector premium despite identical firms and workers. In this paper, however, I take as my starting point the importance of individual heterogeneity, the small formal sector in Tanzania and the substantive differences between wage and self-employment.

I thus use the equilibrium search framework and the theoretical and empirical insights discussed above to develop a two-sector matching model of the urban labour market in Tanzania in which individuals choose to work either for a firm or for themselves. Embedding this choice within the search theory framework is an important strength of the model as this generates earnings heterogeneity in both wage and self-employment, despite an assumption that productivity varies only in self-employment.

In the model the productivity of some individuals in wage employment is low relative to their self-employment productivity. Variation in self-employment productivity generates an opportunity cost of wage employment that is increasing in self-employment productivity, thus incorporating insights about comparative advantage in different sectors that date back to Roy (1951). I use the Albrecht et al. (2009) model of a Latin American labour market as the basis for my model. However I improve this model in a number of ways. I alter the model to better fit the variation in earnings and the differences in earnings determination across wage and self-employment. My model generates earnings variation in both wage and self-employment, despite having variation in ability in only one sector. I also provide a complete equilibrium analysis of my (slightly simpler) model, which was absent in the Albrecht et al. (2009) paper.

Section 2 explores the Tanzanian data and how these data speak to different models of the Tanzanian labour market. In section 3 I develop a model which captures some of the features presented in the empirical exploration and in section 4 I solve this model. Section 5 explores how the numerical solution to the model matches the empirical evidence and section 6 concludes.

2 Evidence from the Tanzanian Labour Market

I use two sources of data to explore the urban Tanzanian labour market and to discriminate between competing understandings of this market. The first is the nationally representative 2006 Integrated Labour Force Survey (ILFS), collected by the Tanzanian National Bureau of Statistics (NBS), and the second is the 2004 and 2005 rounds of the Tanzanian Urban Panel Survey (TUPS), collected by the Centre for the Study of African Economies (CSAE) at the University of Oxford[1].

Table 1 shows that self-employment is the most common type of employment in urban Tanzania. Over 80 percent of the self-employed are own account workers with the remainder mostly employing a small number of employees. Wage employment accounts for around a third of urban employment, split roughly equally between smaller firms employing less than 10 workers, larger firms and public sector employment. It is clear that obtaining public sector employment requires significant investments in human capital- with nearly two thirds of public employees having completed O-levels. Urban unemployment rates are relatively high as measured in the ILFS, with around 17% of those between the age of 16 and 64 unemployed in urban Tanzania, compared to 4.5% for the country as a whole.

Table 1 also shows that median earnings in the ILFS are lowest for those in small firms, followed by those in self-employment with no employees and then those in large firms[2]. Public sector workers earn the most, but this is partly due to their high levels of education [3]. Median earnings are also much higher for entrepreneurs with employees, but these are still mostly informal businesses, with one third of these employers having written accounts and only 15 percent having accessed credit in the last year. These rankings are preserved in regressions with detailed controls for human capital, shown in Table 2, using the TUPS data.

[1]For further information on the TUPS see Rankin et al. (2010)

[2]The NBS data release of the 2006 ILFS incorrectly coded the direct measure of net income so I use gross income minus expenses to calculate net income.

[3]Wage employees were asked about their gross income in the ILFS. It is likely that public sector workers paid tax but that small firm employees and the self-employed did not.

Differences in median earnings tell only part of story however. Figure 1 shows the real earnings densities in employment for the same categories as in Table 1. There is clearly much more variance in earnings in self-employment compared to wage employment[4]. There are also large overlaps between what are generally considered formal and informal sector jobs. Around a third of employees in firms that have more than 10 employees earn less than the median in self-employment for those without employees. The differences in median earnings suggest that a focus on average earnings differentials is important but the large variation within sectors, particularly within self-employment, suggests a model of earnings determination should also account for this heterogeneity in earnings.

Table 3 uses longitudinal data from the 2004 and 2005 rounds of the Tanzanian Urban Panel Survey to explore employment transitions. There is very little movement between self-employment and wage employment[5], though there is movement within these categories between self-employment with, and without, employees, as well as some between small and large firm wage employment. This suggests that comparative advantage in wage and self-employment may play an important role in decisions about whether to work for one's self or a firm.

The data I have presented from the urban Tanzanian labour market suggests there is substantial variation in earnings in self-employment, with earnings for many self-employed exceeding that in wage employment in small firms. There is also very little movement in either direction between wage employment (in small or larger private firms) and self-employment. This evidence, together with the obvious differences in earnings determination between wage and self-employment, suggests that a fruitful way of modeling the labour market may be to assume that individuals in the private sector choose between wage and self-employment based on their comparative advantage.

3 A Model of the Tanzanian Labour Market

3.1 Model Outline

As a result of the evidence presented above, as well as intrinsic differences in earnings determination for wage and self-employment, I build a two-sector

[4]This may partly be explained by measurement error, since the self-employed are asked about and must therefore calculate their profits, whereas wage employees simply report their earnings, as explained above.

[5]This is also the case in the data presented by Ulyssea (2010) for the Brazilian labour market.

model of comparative advantage in the Tanzanian labour market that also captures the difference in the way earnings are determined in wage and self-employment[6]. The model generates outcomes that are broadly consistent with the empirical evidence from Tanzania.

I use a matching framework in a Roy-type model (Roy 1951) of comparative advantage, in which individuals choose either wage or self-employment opportunities based on their expected earnings in each sector, building on the model of Albrecht et al. (2009). Frictions limit the rate at which workers and firms meet each other and the rate at which the self-employed find self-employment opportunities. Individuals earn their marginal product in self-employment, though wage employees do not as they are assumed to bargain with the firms they meet.

As a result of the variation in earnings within sectors that I highlighted above, I assume that individuals vary in self-employment productivity but that there is constant productivity across all firms, as well as across all workers in wage employment. Despite there being variation only in self-employment productivity, my model also generates earnings variation in the bargained wage, as a result of the variation in workers' outside options due to their differences in self-employment productivity. This highlights the importance of the distinction between self and wage employment- it is this distinction within the search framework that buys variation in earnings in both sectors, despite homegenous productivity in wage employment.

As a result of the differences in self-employment productivity, individuals with low ability take only wage employment, some individuals take both, and some take only self-employment. All individuals experience unemployment when their jobs are destroyed, and they then go back to searching for either wage opportunities or self-employment opportunities or both[7].

[6]The empirical analysis above suggested that earnings in the public sector provides a large part of the incentive to acquire advanced secondary or tertiary education. I do not include it in my model, however, and focus instead on the choice between private wage employment and self-employment.

[7]Some readers may find it odd that I include a job destruction rate and frictions in self-employment but I would argue both that this is plausible and that it makes sense of the empirical analysis above. An individual selling goods from an illegally sited kiosk may find it destroyed by government officials. In both cases the individual concerned would have to look for a new opportunity, in the sense of finding a site to sell goods at. This would take time and could be modeled by assuming there are frictions in self-employment. Frictions may well be lower in self-employment, as argued in Zenou (2008) and I allow for this possibility, exploring the effects of varying the level of frictions in section 5.

3.2 Individual Behaviour

In the model self-employment productivity is distributed over a population of individuals whose (constant) size is normalised to 1. Employment for workers and the filled jobs of firms are thought of as assets that bring an instantaneous return, in the form of wage or self-employment income for workers and profits for firms. The capital value of these assets may change in the future as a result of shocks to the match which result in it being destroyed. For workers, the new value of the employment asset is assumed to be the value of unemployment that results from the job ending. For the firm, the new value is the (negative) value of holding a vacancy open.

An individual with productivity y in self-employment earns y but faces a probability q_s that this self-employment opportunity will be destroyed and that she will enter unemployment. The asset equation for the self-employed is thus

$$rV_s(y) = y + q_s(V_u(y) - V_s(y)) \tag{1}$$

where V_u is the value of unemployment, V_s is the value of self-employment and r is the interest rate.

Unemployed workers search for employment opportunities and whilst doing this receive unemployment income ky with $0 < k < 1$. Unemployment income is a function of individuals' self-employment ability, as in Wong's (2003) model of the US labour market[8]:

$$rV_u(y) = ky + \alpha\max[(V_s(y) - V_u(y)), 0] + m(\theta)\max[(V_e(y) - V_u(y)), 0]. \tag{2}$$

α is the probability that an unemployed worker will find a self-employment opportunity. $m(\theta)$ is the rate at which wage employment opportunities arrive and is a function of θ, a measure of labour market tightness. $\theta = v/u$, where v is the number of vacancies and u is the number of unemployed persons. When the labour market is tight v is low relative to u, θ is thus low and workers will have a harder time finding firms to match with, implying that the matching rate will be lower. The matching rate in self-employment is assumed to be exogenous. When an individual finds a self-employment opportunity or a wage employment opportunity this results in a change in the capital value of the unemployment asset that an unemployed person is thought of as holding.

I noted above that variation in self-employment productivity creates an opportunity cost of taking wage employment that is increasing in self-employment

[8]Assortative mating by income, education or assets is a possible explanation for unemployment earnings depending on productivity, with a spouse providing income in a period of unemployment, but neither the household nor education are taken into account explicitly in this model. Fafchamps and Quisumbing (2005) provide evidence that sorting on human capital occurs in the Ethiopian marriage market.

productivity, similar to the model of Albrecht et al. (2009). Thus workers with low self-employment productivity choose employment in the wage sector and reject potential self-employment options, because they would earn much less in self-employment and are better off searching for wage opportunities. Similarly, workers with high self-employment productivity reject wage employment offers.

The asset equation for wage employees is

$$rV_e(y) = w(y) + q_e(V_u(y) - V_e(y)), \tag{3}$$

where the wage w depends on y and V_e is the value of wage employment. As in self-employment, wage employment is assumed to end with an exogenous probability (q_e in wage employment) in which case the worker returns to unemployment.

3.3 Firms

Firms are assumed to have homogeneous productivity y_w and to employ only one worker. They match with workers at the rate $m(\theta)/\theta$. When a match occurs the firm produces a continuous stream of output y_w and pays a wage $w(y)$ to the worker it matches with. The wage paid depends on worker ability because ability affects the worker's outside option. Filled jobs can also be thought of as assets that firms hold, meaning that the asset equation for a filled job is

$$r\prod_e(y) = y_w - w(y) + q_e(\prod_v - \prod_e(y)). \tag{4}$$

Matches are assumed to end with exogenous probability q_e, which brings about a change in the capital value of the filled job asset for the firm. Firms whose matches end are assumed to return to search for workers who are also searching for employment. While this search occurs, firms are assumed to incur a cost of holding a job open, meaning that the asset equation for a vacant job is

$$r\prod_v = -c + (m(\theta)/\theta)E[\max[(\prod_e(y) - \prod_v, 0]]. \tag{5}$$

In equation (5), c is the cost of holding a vacant job open per unit of time and $m(\theta)/\theta$ is the rate at which firms match with workers. Following much of the literature, it is also assumed that firms are risk neutral.

3.4 Wage Determination

The match between a firm and a worker generates a gain for both parties, who are assumed to bargain over how the match surplus is distributed between them. A worker's outside option is the value of unemployment and a firm's is the value of keeping a vacancy open. Thus the rents are $(V_e(y) - V_u(y))$ for a worker and $\prod_e - \prod_v$ for a firm, whilst the surplus is $S = \prod_e - \prod_v -(V_e(y) - V_u(y))$.

I assume that with free entry firms compete away expected profits, so that $\prod_v = 0$ and hence that the firm's outside option has an expected value of zero in equilibrium. Thus $S = \prod_e -(V_e(y) - V_u(y))$. Following the matching literature I assume that the bargaining outcome is the maximum of the Nash criterion (Nash (1953), Binmore et al. (1986)). Thus the negotiated wage, $w(y)$, maximises

$$[V_e(y) - V_u(y)]^{\gamma}[\prod_e(y)]^{(1-\gamma)}. \tag{6}$$

The first order conditions from this problem give $\prod_e(y) = (1-\gamma)S$ and $V_e(y) - V_u(y) = \gamma S$, the surplus sharing rule. Thus firms and workers share the surplus according to their bargaining power, which is determined outside the model. Substituting the value of $\prod_e$ from equation (4) it is possible to show that

$$w(y) = \gamma y_w + (1-\gamma) r V_u(y)). \tag{7}$$

Equation (7) shows a dependence of the negotiated wage w on y because workers' outside options, $V_u(y)$), vary with self-employment ability y, despite identical productivity for all employees. There are two sources of this variation. The first is that unemployment income depends on self-employment ability y. The second is that even if unemployment income was constant across workers, workers have different productivities in self-employment.

4 Model Equilibria

The model setup and the variation in self-employment ability implies an opportunity cost of taking wage employment that is increasing in self-employment ability. The hypothesised equilibrium of the model is thus one in which there are two cutoff values in ability, y^* and y^{**} with $y^* < y^{**}$. These correspond respectively to the self-employment productivity at which unemployment is no longer preferred to self-employment, and at which wage employment is no longer preferred to unemployment [9].

[9]There is also the possibility of a one cutoff equilibrium, in which all individuals take wage employment and some high productivity individuals also take self-employment, as

Below I show that in the model I have set out some individuals never take wage employment, preferring to stay unemployed and search for self-employment. I assume that these individuals do not congest the matching process between firms and workers looking for wage employment opportunities. In this model a steady state equilibrium is one in which the expected value of maintaining a vacancy is zero, matches are consummated only if it is in the interests of the firm and individual to do so, the steady state employment flows conditions hold, so that the unemployment rate is not changing, and the search incentives of individuals are actually consistent with the proposed equilibrium (Albrecht et al. 2009). A value of θ together with cutoffs y^* and y^{**} and an unemployment rate u at which the above conditions are satisfied is an equilibrium of the model.

4.1 Verifying the Proposed Two Cutoff Equilibrium

In this section I verify that the search incentives of individuals are indeed consistent with the proposed equilibrium, checking that individuals actually act in the manner in which it was hypothesised they acted. The model is based on Albrecht et al. (2009) but this paper provides no verification that the equilibrium proposed in the paper is indeed an equilibrium[10]. In the model I have outlined verification requires showing that y^* is indeed a point below which $V_u(y) > V_s(y)$, ie unemployment is preferred to self-employment, and y^{**} is indeed a point above which $V_u(y) > V_e(y)$, ie unemployment is preferred to wage employment.

In the proposed equilibrium, by assumption y^* is the point at which $V_u(y) = V_s(y)$ and y^{**} is the point at which $V_u(y) = V_e(y)$. To verify that individual search incentives are consistent with the hypothesised equilibrium it is thus sufficient to show that $V_u(y)$, $V_e(y)$ and $V_s(y)$ are increasing in y for all y between zero and 1 and then to show that $V_s'(y^*) > V_u'(y^*)$ and that $V_u'(y^{**}) > V_e'(y^{**})$. It must be true that $y^* < y^{**}$.

To begin I show that $V_u(y)$, $V_e(y)$ and $V_s(y)$ are increasing in y for $y < y^*$. Equations (3) and (7) give the asset equation for wage employment. Below y^* individuals only take wage employment in the hypothesised equilibrium.

well as a zero cutoff equilibrium in which all individuals take only wage employment. Intuitively, these types of equilibria will occur when the firm productivity is "high enough" to make self-employment unattractive. I do not investigate these equilibria any further in the rest of the paper and confine my attention to the two cutoff equilibrium, which is both more interesting and more realistic in the case of the Tanzanian labour market. By assuming that $0 < k < 1$ in the unemployment asset equation I have ruled out an equilibrium in which no one works.

[10] I am grateful to Margaret Stevens for pointing this out to me.

Thus the unemployment asset equation, equation (2), becomes

$$rV_u(y) = ky + m(\theta)(V_e(y) - V_u(y)). \tag{8}$$

Solving the unemployment and wage employment asset equations it is possible to show that

$$rV_u(y) = \frac{m(\theta)\gamma y_w + ky(r + q_e)}{r + q_e + m(\theta)\gamma}. \tag{9}$$

It is clear from equation (9) that $V_u(y)$ is increasing in y and therefore, using equation (8), that $V_e(y)$ is increasing in y and, using equation (1), that $V_s(y)$ is increasing in y. This is all for $y < y^*$ in the hypothesised equilibrium.

Having shown that $V_u(y)$, $V_e(y)$ and $V_s(y)$ are increasing in y for $y < y^*$ I now show that $V_s'(y^*) > V_u'(y^*)$. At y^* V_u is defined by equation (9) and V_s can be found using equations (1) and (9). Thus

$$V_u'(y) = \frac{k(r + q_e)}{r(r + q_e + m(\theta)\gamma)} \tag{10}$$

and hence

$$V_s'(y) = \frac{1 + q_s V_u'(y)}{r + q_s}. \tag{11}$$

Using equation (10), the fact that $k < 1$ and that $m(\theta)\gamma > 0$, it is clear that $V_u'(y)r < 1$ for all y and hence $V_s'(y^*) > V_u'(y^*)$.

I now show that V_s, V_e and V_u are increasing in y for $y > y^{**}$ and that $V_u'(y^{**}) > V_e'(y^{**})$. Above y^{**} the unemployment asset equation, equation (2), and the self-employment asset equation, equation (1), can be solved for $V_u(y)$:

$$rV_u(y) = \frac{ky(r + q_s) + \alpha y}{r + q_s + \alpha}. \tag{12}$$

$V_u(y)$ is clearly increasing in y above y^{**} in the hypothesised equilibrium. Equations (12) and (1) imply that $V_s(y)$ is also increasing in y above y^{**} in the hypothesised equilibrium.

I must now show that $V_e(y)$ is increasing in y above y^{**} and that $V_u'(y^{**}) > V_e'(y^{**})$. Equations (3), (7) and (12) can be solved for $V_e(y)$:

$$V_e(y) = \frac{\gamma y_w}{r + q_e} + (\frac{1 - \gamma}{r + q_e} + \frac{q_e}{r(r + q_e)})(\frac{yk(r + q_s) + \alpha y}{r + q_s + \alpha}). \tag{13}$$

Thus $V_e(y)$ is increasing in y above y^{**} and, since $0 < \gamma < 1$ this implies $V_u'(y^{**}) > V_e'(y^{**})$.

For $y^* < y < y^{**}$ the unemployment asset equation becomes

$$rV_u(y) = ky + \alpha(V_s(y) - V_u(y)) + m(\theta)(V_e(y) - V_u(y)), \qquad (14)$$

since individuals with these levels of productivity take both wage and self-employment in the hypothesised equilibrium.

The wage employment asset equation, from equations (3) and (7), together with the self-employment asset equation, equation (1), and the unemployment asset equation, equation (14), imply that

$$rV_u(y) = \frac{ky(r+q_s)(r+q_e) + \alpha y(r+q_e) + m(\theta)\gamma y_w(r+q_s)}{(r+q_s)(r+q_e) + \alpha(r+q_e) + m(\theta)\gamma(r+q_s)}. \qquad (15)$$

This is clearly increasing in y for $y^* < y < y^{**}$. From equations (1), (3) and (7) it is obvious that $V_s(y)$ and $V_e(y)$ are also increasing in y for $y^* < y < y^{**}$.

I have thus shown that individuals' incentives are consistent with the actions they are hypothesised to take in the proposed two cutoff equilibrium. Below I solve the model in the case of the two cutoff equilibrium and then show that this is the only two cutoff equilibrium possible.

4.2 Solving for y^* , y^{**} and θ

Solving the model requires obtaining the cutoff values y^* and y^{**} and solving for the equilibrium value for θ, labour market tightness. This first requires finding an expression for y^* in terms of the exogenous model parameters and θ. In the previous section it was verified that below y^*, $V_u(y) > V_s(y)$ (unemployment is preferred to self-employment). This explains why low ability individuals take only wage employment opportunities. Individuals with productivity above y^*, but below y^{**}, take both wage and self-employment. At y^*, $V_u(y^*) = V_s(y^*)$, which with equation (1) implies that $rV_s(y^*) = y^*$ and thus, using equations (2), (3) and (7), it is possible to express y^* in terms of θ and the exogenous parameters of the model:

$$y^* = \frac{m(\theta)\gamma y_w}{(1-k)(r+q_e) + m(\theta)\gamma}. \qquad (16)$$

In a similar fashion one can also find the value of y^{**}, the level of productivity above which individuals take only self-employment. At y^{**}, $V_u(y^{**}) = V_e(y^{**})$ and hence the unemployment asset equation (2) and the self-employment asset equation (1), can be used to substitute out $V_s(y^{**})$ and solve for $V_u(y^{**})$:

$$rV_u(y^{**}) = \frac{ky^{**}(r+q_s) + \alpha y^{**}}{r+q_s+\alpha}. \qquad (17)$$

By assumption, y^{**} is the value of y such that $V_u(y^{**}) = V_e(y^{**})$, and together with equation (3) and (17), implies that $rV_e(y^{**}) = w(y^{**})$ and hence

$$w(y^{**}) = \frac{ky^{**}(r+q_s) + \alpha y^{**}}{r+q_s+\alpha}. \tag{18}$$

Now the bargained wage equation (7), with $V_u(y^{**}) = V_e(y^{**})$ and $rV_e(y^{**}) = w(y^{**})$, implies that $w(y^{**}) = y_w$. Thus the highest wage firms pay is paid to workers with the highest ability in self-employment who still take wage employment, and this is the wage which leaves the firm indifferent between agreeing to the match and returning to search for another worker, where the expected return must zero in equilibrium as a result of free entry. This logic implies

$$y^{**} = \frac{y_w(\alpha + q_s + r)}{\alpha + k(r+q_s)}. \tag{19}$$

I now have an equation for y^{**} in terms of the exogenous parameters of the model and have expressed y^* in terms of θ.

Thus far in proposing a two cutoff equilibrium and then verifying that it indeed was an equilibrium I assumed that $y^* < y^{**}$. I have not yet shown that this is indeed true in the hypothesised equilibrium, which I am now able to do based on the solutions for y^* and y^{**} found above. Using equation (19), if $k = 1$ then $y^{**} = y_w$ but $k < 1$. Thus $y^{**} > y_w$. Similarly, using equation (16), if $k = 1$ then $y^* = y_w$ but $k < 1$ so $y^* < y_w$. Thus $y^* < y^{**}$.

I have thus shown that the proposed equilibrium does imply that $y^* < y^{**}$ and found an equation linking y^* and θ (16). I now require a second equation to be able to solve for y^* and θ. To do this it is necessary to first solve for the steady state employment flows, which I do in the following section.

4.3 Equilibrium Employment Flows

In the steady state of any equilibrium of this model the unemployment rate cannot be changing, implying that the flows into and out of unemployment must be equal. These flows will differ by the workers' ability type, and by the sector of employment they choose in the hypothesised equilibrium.

Following Albrecht et al. (2009), I let $u(y)$ be the fraction of time an individual of productivity y spends unemployed, $n_e(y)$ be the fraction of time spent in wage employment and $n_s(y)$ be the fraction of time spent in self-employment, meaning that $u(y) + n_e(y) + n_s(y) = 1$[11].

[11] There is no non-participation in this model.

Individuals with productivity $y < y^*$ only move between wage employment and unemployment, implying that $n_s(y) = 0$. The steady state condition is thus $q_e(1-u(y)) = m(\theta)u(y)$, with $u(y) = \frac{q_e}{m(\theta)+q_e}$ and $n_e(y) = 1-u(y)$.

Individuals with $y^* < y < y^{**}$ take both wage and self-employment opportunities, meaning there are two steady state conditions for these individuals. Firstly, the flows into wage employment equal the flows into unemployment from wage employment, and secondly, the flows out of self-employment equal the flows into unemployment from self-employment. Hence $m(\theta)u(y) = q_e n_e(y)$ and $\alpha u(y) = q_s n_s(y)$. Combining these two implies

$$u(y) = \frac{q_e}{m(\theta) + q_e + \frac{\alpha q_e}{qs}}, \tag{20}$$

$$n_e(y) = \frac{m(\theta)}{m(\theta) + q_e + \frac{\alpha q_e}{qs}} \tag{21}$$

and

$$n_s(y) = \frac{q_e \alpha}{q_s(m(\theta) + q_e + \frac{\alpha q_e}{qs})}. \tag{22}$$

Individuals with $y > y^{**}$ take only move between self-employment and unemployment. This implies $\alpha u(y) = q_s n_s(y)$, so that $n_s(y) = 1 - u(y)$, $n_e(y) = 0$, $u(y) = \frac{q_s}{\alpha+q_s}$ and $n_s = \frac{\alpha}{\alpha+q_s}$.

Total unemployment can then be calculated as

$$u = \int_0^{y^*} u(y)f(y)dy + \int_{y^*}^{y^{**}} u(y)f(y)dy + \int_{y^{**}}^{1} u(y)f(y)dy, \tag{23}$$

where $u(y)$ varies for each of the three categories of worker, and $f(y)$ is the ability distribution in the population.

4.4 Using the Free Entry Condition

The model can be solved by finding a second equation containing the two unknowns y^* and θ, which can be obtained from the free entry condition $\prod_v(y) = 0$ that holds in the steady state equilibrium. This, together with equation (5), the vacant job asset equation, and the fact that in an equilibrium $\prod_e(y) \geqslant 0$, because firms will only match with workers that they do not make a loss from hiring, implies

$$c = \frac{m(\theta)}{\theta} E[\prod_e(y)]. \tag{24}$$

The value of a filled job to a firm is uncertain because the wage paid depends on the ability of the worker employed. In the proposed equilibrium only workers with productivity $y < y^{**}$ take wage employment offers. To calculate the expected value of a filled job, $E[\prod_e(y)]$, the distribution of workers with which firms can potentially match is required. The distribution of worker ability in the population is known but this is different to the distribution of unemployed workers who would take wage employment offers from firms. Above I assumed that the self-employed who do not search for wage employment do not congest the matching process between firms and workers who do take wage employment. Thus I denote the density of ability for those in unemployment who are searching for wage employment as $f_u(y)$ and then, following Albrecht et al. (2009), use Bayes' Law to write

$$f_u(y) = \frac{u(y)f(y)}{u_{02}}, \tag{25}$$

where u_{02} is the unemployment rate for all those who take wage employment opportunities, ie those with $y \leq y^{**}$.

Using the equation for a filled job, equation (4), with $\prod_v = 0$, and the value of $w(y)$ from equation (7), the outcome of the Nash Bargain with free entry it is possible to write the free entry condition as

$$c = \frac{m(\theta)(1-\gamma)}{\theta(r+q_e)} \int_0^{y^{**}} (y_w - V_u(y)) \frac{u(y)f(y)}{u_{02}} dy. \tag{26}$$

Equations (26) and (16) are two equations in the two unknowns y^* and θ, which can be solved to find the equilibrium values of these variables, after solving for the values for $V_u(y)$ for those workers with productivity below y^{**}.

The values of $V_u(y)$ differ for individuals with productivity above and below y^*. Workers with $y^* < y < y^{**}$ take both wage and self-employment, meaning that solving the self-employment, wage employment and unemployment asset equations (equations (1), (3)) and (2)) for rV_u implies

$$rV_u(y) = \frac{ky(r+q_s)(r+q_e) + \alpha y(r+q_e) + \gamma m(\theta) y_w (r+q_s)}{(r+q_s)(r+q_e) + \alpha(r+q_e) + \gamma m(\theta)(r+q_s)}. \tag{27}$$

Workers with $y < y^*$ do not take self-employment so solving the wage employment and unemployment asset equations ((3) and (2)) implies

$$rV_u(y) = \frac{ky(r+q_e) + \gamma m(\theta) y_w}{r + q_e + \gamma m(\theta)}. \tag{28}$$

$rV_u(y)$ is now expressed in terms of the exogenous parameters and y and θ and can be substituted into the free entry condition (26), along with the values of y^* from equation (16) and y^{**} from equation (19). This is a non-linear equation in theta, requiring a numerical solution, which I explain in more detail below, after proving that a unique two cutoff equilibrium to the model exists.

4.5 The Existence of a Unique Two Cutoff Equilibrium

In the sections above I proposed a two cut off equilibrium to the model laid out in section 3 and then verified that the proposed two cutoff equilibrium is indeed an equilibrium of the model. I now show that the proposed equilibrium is unique in the class of possible two cutoff equilibria. I have shown above that $u(y)$ and y^* are uniquely determined by θ, and that y^{**} is determined by the exogenous parameters of the model. A unique two cutoff equilibrium exists if there is a unique value of θ that solves the free entry condition (26). If I assume decreasing returns to scale in the matching function m, a standard assumption in the literature (Cahuc and Zylberberg 2004), then $\frac{m(\theta)}{\theta}$ is strictly decreasing in θ. $u(y)$ has different values for $y < y^*$ and $y^* < y < y^{**}$ but, as can be seen from section 4.3, both are strictly decreasing in θ. Since u is the sum of these two terms for $u(y)$ that are both strictly decreasing in θ, it is also strictly decreasing in θ[12].

To show that the free entry condition is decreasing in θ it is only required to show that $rV_u(y)$ is increasing in θ, where $rV_u(y)$ is defined by equation (28) for $y < y^*$ and by equation (27) for $y^* < y < y^{**}$. In both cases it can be shown that $\frac{\partial rV_u(y)}{\partial\theta}$ is positive and hence that the free entry condition is decreasing in θ.

To prove that the equilibrium is unique I must show that as $\theta \to \infty$ the right hand side of the free entry condition tends to zero and that it tends to ∞ as $\theta \to 0$. $\frac{m(\theta)(1-\gamma)}{\theta(r+q_e)}$ clearly tends to ∞ as $\theta \to 0$ and tends to zero as $\theta \to \infty$. Thus I only need to show that the definite integral is bounded to show that the equilibrium is unique. u and $u(y)$ are unemployment rates that are bounded between 0 and 1 for any value of θ. $f(y)$ is the distribution of abilities in the population and so does not depend on θ, and neither does y_w. $V_u(y)$ is the value of unemployment. As $\theta \to 0$ the labour market becomes very tight and matching with a firm increasingly unlikely. Thus, in the limit, individuals would not find wage employment. But as the model has been set up they can still find self-employment, where the arrival rate does not vary with θ. Thus the value of unemployment is still bounded in this case.

A similar argument holds when $\theta \to \infty$ and wage jobs are immediately available for those who become unemployed, except that some individuals still take self-employment, even though wage jobs are available instantly, because their employers would not pay a wage greater than the common level of productivity. Thus the definite integral is bounded and there is a unique value of θ that solves the free entry condition. In the following section I now explore the distributions of earnings implied by the equilibrium.

[12] u also includes a term for $u(y)$ above y^{**} but this does not depend on θ.

4.6 Equilibrium Earnings distributions in Wage and Self-Employment

In the model I have outlined wages depend on worker's outside options, which are themselves determined by individuals' self-employment productivities, whilst the self-employed earn their productivity. I now calculate the distributions of earnings in equilibrium.

$f(y)$ is the density of abilities in the population. If $h(y)$ is the density of ability in wage employment and $P(E)$ is the probability of being in wage employment, Bayes Law then implies that

$$h(y) = \frac{P(E|y)f(y)}{P(E)} = \frac{n_e(y)f(y)}{\int_0^1 n_e(y)f(y)dy}. \tag{29}$$

For simplicity I assume a uniform distribution of ability between 0 and 1. It is clear from section 4.3 above that the probability of wage employment, n_e, differs for different values of y. This means

$$\begin{aligned} h(y) &= \frac{\frac{m(\theta)}{m(\theta)+q_e}}{\frac{y^* m(\theta)}{m(\theta)+q_e} + (y^{**} - y^*)\frac{m(\theta)}{m(\theta)+q_e+\frac{\alpha q_e}{q_s}}} && 0 \leq y < y^* \\ &= \frac{\frac{m(\theta)}{m(\theta)+q_e+\frac{\alpha q_e}{q_s}}}{\frac{y^* m(\theta)}{m(\theta)+q_e} + (y^{**} - y^*)\frac{m(\theta)}{m(\theta)+q_e+\frac{\alpha q_e}{q_s}}} && y^* < y < y^{**} \\ &= 0 && y^{**} < y \leq 1 \end{aligned}$$

Since $h(y)$ is a different constant for each of the two groups of lower productivity individuals calculating the distribution of wages $m(w)$ is relatively simple. The distribution will also have two constant parts on the domain of w, but with a much narrower base than the ability distribution. If I let the height of the first block of the $h(y)$ distribution be a then the probability mass within this block is ay^*. The area of the first block of the distribution of w will be the same and thus the height of this block will be $\frac{ay^*}{w(y^*)-w(0)}$. Using similar reasoning, the height of the second block will be $\frac{b(y^{**}-y^*)}{w(y^{**})-w(y^*)}$. Solving the model numerically I will be able to calculate the equilibrium wages at $y = 0$, $y = y^*$ and $y = y^{**}$, obtain the density and then be able to solve for the average wage paid in the equilibrium. The average wage is calculated as $\int_0^{y^{**}} wm(w)dw$.

Calculating the density of earnings in self-employment is also simple. Again, the density will differ for those with ability y above and below y^{**}

since the probability of self-employment varies with ability. Denoting the probability of self-employment as $P(S)$ I can write the distribution of ability in self-employment, again using Bayes' rule, as

$$g(y) = \frac{P(S|y)f(y)}{P(S)} = \frac{n_s(y)f(y)}{\int_0^1 n_s(y)f(y)dy}. \tag{30}$$

From section 4.3 I have the probability of being in self-employment for individuals with $y > y^*$ and I know that $P(S) = 0$ if $y < y^*$. Thus it is possible to write down the density of ability in self-employment as

$$\begin{aligned} g(y) &= 0 && 0 < y < y^* \\ &= \frac{\frac{q_e\alpha}{q_s(m(\theta)+q_e+\frac{\alpha q_e}{q_s})}}{((y^{**}-y^*)\frac{q_e\alpha}{q_s(m(\theta)+q_e+\frac{\alpha q_e}{q_s})})+(1-y^{**})\frac{\alpha}{\alpha+q_s}} && y^* < y < y^{**} \\ &= \frac{\frac{\alpha}{\alpha+q_s}}{((y^{**}-y^*)\frac{q_e\alpha}{q_s(m(\theta)+q_e+\frac{\alpha q_e}{q_s})})+(1-y^{**})\frac{\alpha}{\alpha+q_s}} && y^{**} < y \leq 1 \end{aligned}$$

Since earnings in self-employment are simply the value of individual productivity in self-employment, the ability distribution is also the earnings distribution in self-employment.

5 Numerical Solutions to the Model

Solving for θ in a non-linear equation requires numerical methods and so I use Broyden's (1965) quasi-Newton method, implemented in Matlab. A solution requires assuming values for the exogenous parameters and functional forms of the model, or estimating the model and parameters using data. As far as I am aware, there have been no matching models estimated on African data, and as a result there are no empirical estimates of most of the parameters for which I require values. Thus I simply follow other papers in using reasonable parameter values and then explore what the model predicts about the earnings distributions across wage and self-employment, transitions between the two sectors, and average wages in each sector.

5.1 Exogenous Parameter Values

The exogenous parameters required in order to solve the model are those of the interest rate, the job destruction rate in self-employment and wage employment, the arrival rate of self-employment opportunities, firm productivity, the cost to a firm of holding a job open and the bargaining power parameter. A functional form for the matching function $m(\theta)$ is also required. The time unit is assumed to be one year.

I follow some of the assumptions of Albrecht et al. (2009), the only other two sector developing country microeconomic matching model I am aware of, in assuming values for the parameters and functional forms, but modify these for the Tanzanian context in some cases. I assume the matching function to be of the form $m = 2.5\theta^{.5}$, the cost of holding a job open $c = .3$, the arrival rate of self-employment offers $\alpha = 2.5$, the job destruction rates in self-employment and wage employment $q_s = q_e = .5$, the real interest rate $r = 0.05$ and a uniform distribution of self-employment productivity between 0 and 1. I assume that $y_w = .5$ which, given $y \sim$ uniform$(0, 1)$, is the mean and median of this productivity[13]. I finally assume that $k = .3$, so that individuals have unemployment income that is 30 percent of their self-employment productivity.

5.2 Equilibrium Solution

The equilibrium outcome with the chosen parameter values and functional forms is shown in the third column of Table 4. In this baseline equilibrium $y^* = .35$, $y^{**} = .57$ and $\theta = 0.52$. This means that 35 percent of the population choose wage employment only, around 22 percent choose either wage or self-employment and 43 percent choose only self-employment. Total unemployment is 17 percent, with unemployment being 22 percent for individuals with $y < y^*$, 10 percent for individuals with $y^* < y < y^{**}$ and 17 percent for high productivity individuals with $y > y^{**}$. Individuals with productivity $y^* < y < y^{**}$ take both wage and self-employment opportunities so their unemployment rate is much lower than for those with $y < y^*$ or $y > y^{**}$ who take only wage and self-employment offers respectively. The vacancy rate is defined by $v = u_{02}\theta$, where u_{02} here is the unemployment rate for all individuals who actually take wage employment opportunities, meaning that $y > y^{**}$ individuals do not congest the labour market. Thus $v = 9$ percent. The average wage paid is .43 and the average earnings in self-employment is .7. I discuss this further in section 5.3 below.

5.2.1 Earnings variation in wage and self-employment

Those who take only wage employment in the equilibrium ($0 < y < y^*$) have an extremely low variance of earnings since $w(0) = .405$ whilst $w(y^*) =$

[13] I discuss the implications of other potential underlying distributions below.

$w(.35) = .425$. This narrow distribution is a result of identical firm productivity and the small variation in the outside options of those who choose only wage employment in the equilibrium. For these individuals, their self-employment productivity does not enter directly into their outside option in the wage bargain with firms, because they never actually exercise their self-employment option in the equilibrium of the model. Thus only the difference in workers' unemployment income, which is assumed to be a function of self-employment productivity, generates wage differentials amongst these workers.

There is more wage variation for those who also take self-employment opportunities ($y^* < y < y^{**}$) since $w(y^*) = w(.39) = .425$ and $w(y^{**}) = w(.57) = .5$. This is because these individuals do exercise their self-employment options, which enter the wage bargain and generate greater variation in individuals' outside options.

Having shown how it is possible to generate the distribution of earnings conditional on wage employment in Section 4.6 above, the corresponding wage cdf is shown in Figure 2 (labeled baseline), which emphasises the low variation in wages.

Those taking only self-employment simply earn their self-employment productivity, which means that there is a much larger amount of variation in self-employment earnings compared to wages earned working for firms with identical productivity. The minimum amount earned in self-employment is y^*, whilst the maximum is the highest productivity in the distribution of y, which are, respectively, .35 and 1 in the baseline model. Thus the maximum and minimum earnings for the population are respectively higher and lower than the maximum and minimum earnings in wage employment. The cdf for self-employment earnings is shown in 2 and is labeled "Self-employment earnings cdf."

5.3 Analysis of the Model Solution

The modeling of comparative advantage of individuals in one or other sector means that most individuals prefer to take opportunities in only one of the sectors, with only around 10% of individuals moving between the two sectors, as in the TUPS panel data. The model also generates higher variance in self-employment earnings than in wages, which is consistent with both the TUPS and ILFS data. The solution to the model also agrees with the earnings data from the 2006 ILFS, in which mean earnings for employees in small firms are around 40% lower than mean earnings in self-employment without employees, and the TUPS, where even with controls for observable human capital employees in small firms earn less than the self-employed without employees.

The low variance of wage earnings in the model solution is less realistic. Allowing unemployment income to vary by self-employment ability buys some degree of realism at the lower end of the earnings distribution, in that there is some variation in the bargained wage for individuals taking only wage employment, which would not be the case if unemployment income was fixed (there would be a mass point). Figure 2 suggests, however, that the amount of variation is limited. It is possible to raise the variance by increasing the level of frictions in both wage and self-employment. This is shown by column 1 of Table 5 and by the "high frictions" cdf in Figure 2. The variance is raised slightly, but the cost is that the model unemployment rate is much higher than the measured Tanzanian urban unemployment rate of 17%. Table 5 also shows that lower levels of frictions lower the unemployment rate but only slightly decrease the variance of wage earnings.

I have shown that earnings in small firms are much lower than in self-employment in the data, and the model generates a similar outcome. Raising firm level productivity, which could be thought of as exogenous development progress, perhaps coming from increased product market competition due to declining trade barriers, has the effect of shifting most of total employment towards wage-employment. It also raises wage earnings towards self-employment earnings, as shown in Table 4, both of which have been commonly experienced in the longer term development of economies that have made the jump from low to higher levels of development.

6 Conclusion

In this paper I have presented some stylised facts from the urban Tanzanian labour market and built a model that helps to explain these, based on the work of Albrecht et al. (2009), a model itself inspired by the Roy (1951) model of earnings and sector choice. I improved this model by completing the equilibrium analysis and by generating outcomes that more closely match those observed in the Tanzanian data.

The model I have built moved beyond a focus on average earnings differentials between formal and informal sector employment, instead positing a central role for individual heterogeneity in both the choice of whether to work in wage or self-employment and in earnings determination in urban Tanzania. The model makes sense of the differences between earnings determination mechanisms in wage and self-employment- wage bargaining only occurs if one has an employer, but many workers do not have an employer in the urban Tanzanian labour market and these workers earn their self-employment productivity in the model. By using these categories in a model embedded in the equilibrium search framework, it was possible to generate distributions of earnings in wage and self-employment that exhibit characteristics that are broadly similar to the empirical distributions, including higher variance in self-employment earnings than wage employment earnings, and lower average earnings in wage employment than self-employment. The transitions predicted by the model are also similar to those observed empirically, in that there is little movement between wage and self-employment.

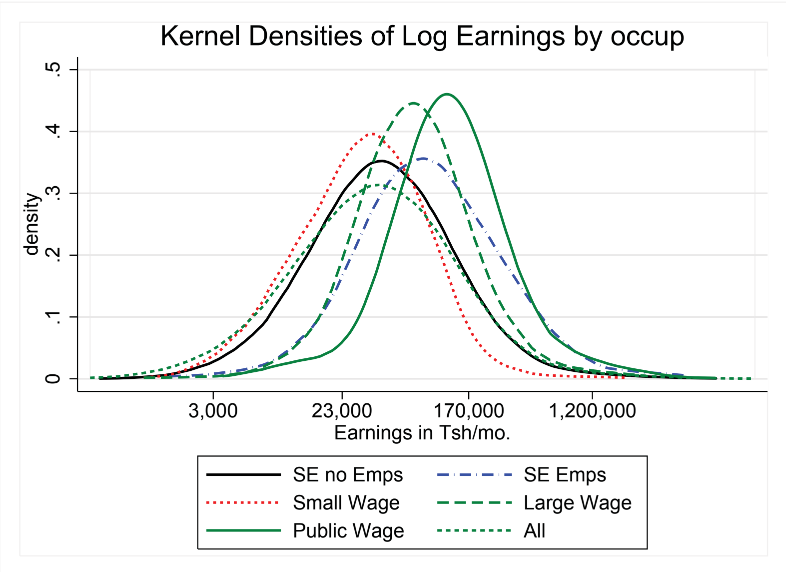
Labour market analysis in Africa has traditionally focused on earnings differentials between the formal and informal sectors, following Harris and Todaro (1970) and Fields (1975). I have shown that this focus does not accurately capture important empirical features of the urban Tanzania labour market, nor does it make sense of the obvious differences in earnings determination in wage and self-employment. The strength of the model presented is its relative simplicity in capturing this distinction between wage and self-employment within the matching framework and its prediction of some of the stylised facts from the urban Tanzanian labour market.

Table 1: Urban Occupational Statistics

Employment Categories	Number	O prop	Mean Earn
Self Emp no Employees	770059	0.105	85482
	(63502)	(0.010)	(8136)
Self Emp with Employees	129370	0.262	201781
	(14031)	(0.030)	(26855)
Small Firm Wage Employment	206788	0.150	54807
	(19704)	(0.019)	(4039)
Large Firm Wage Employment	214959	0.342	127834
	(22533)	(0.033)	(16442)
Public Employment	194215	0.633	265889
	(24195)	(0.035)	(49979)
Agric	217093	0.068	–
	(37965)	(0.013)	–

The data are weighted. Standard Errors in Parenthesis. O Prop is the proportion of individuals in each category with O levels (11 years of schooling under the new system). The average exchange rate was Tsh 1261 to the US dollar in 2006 (International Monetary Fund 2010)

Figure 1: earnings distributions, ILFS



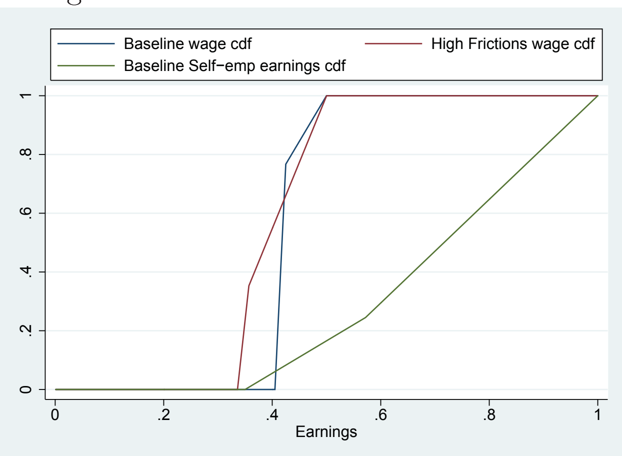
Source: pooled earnings data from Tanzanian ILFS 2006. The kernel density was estimated using Stata.

Table 2: OLS earnings regressions from the TUPS

	2004	2005	Pooled
	(1)	(2)	(3)
Total Exp	0.024**	0.014	0.019**
	(0.012)	(0.01)	(0.009)
$exp^2/100$	-.030	-.015	-.023
	(0.021)	(0.019)	(0.015)
Ln(Hours)	0.025	0.298**	0.132
	(0.117)	(0.136)	(0.093)
Ln(firsiz)	0.18***	0.168***	0.18***
	(0.033)	(0.048)	(0.028)
Ln(emps)	0.615***	0.492***	0.556***
	(0.123)	(0.135)	(0.097)
Tenure	-.006	0.017	0.009
	(0.013)	(0.012)	(0.009)
$Tenu^2/100$	0.011	-.040	-.022
	(0.033)	(0.037)	(0.028)
Male	0.217***	0.299***	0.256***
	(0.065)	(0.065)	(0.051)
Educ	0.0004	-.008	-.002
	(0.024)	(0.026)	(0.02)
$Educ^2/100$	0.634***	0.496***	0.557***
	(0.183)	(0.172)	(0.144)
Pub Ent	-.085	0.193	0.008
	(0.215)	(0.236)	(0.16)
Civ Ser	0.985***	0.908***	0.886***
	(0.17)	(0.125)	(0.111)
Priv Wag	-.241*	-.177	-.216**
	(0.124)	(0.116)	(0.09)
Const.	9.626***	8.735***	9.311***
	(0.483)	(0.569)	(0.394)
Obs.	568	579	1147
R^2	0.354	0.35	0.345

Source: Tanzanian UPS 2004-2005. A year dummy for 2004 was included but not reported.

Figure 2: Wage CDFs for baseline and high frictions model



Source: Baseline: Column 3, Table 4. High Frictions: Column 1, Table 5.

Table 3: Occupation Transitions

	Occupation						
Occupation	S/E no emps	S/E with emps	Small firm	Large firm	Public	none	Total
	%	%	%	%	%	%	%
S/E no emps	87	4	1	0	1	7	100
S/E with emps	43	48	0	2	0	7	100
Small firm	2	2	49	22	10	15	100
Large firm	2	0	6	34	52	6	100
Public	4	0	4	11	70	11	100
none	55	16	2	2	4	20	100
Total	55	9	6	8	13	9	100

Source: 2004 and 2005 rounds of the Tanzanian Urban Panel Survey. Total number of observations observed in both 2004 and 2005 in this table is 457

Table 4: VARYING FIRM PRODUCTIVITY

θ	0.15806	0.31099	0.52025	0.78703	1.112	1.4957	1.9383	2.4402
y^*	0.16904	0.25768	0.35038	0.44537	0.54176	0.63906	0.73696	0.8353
y^{**}	0.34334	0.45779	0.57223	0.68668	0.80113	0.91557	1.03	1.1445
m	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
α	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
y_w	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$w(0)$	0.2212	0.31179	0.40528	0.50054	0.59696	0.69417	0.79194	0.89012
$w(y^*)$	0.23452	0.32884	0.42519	0.52269	0.62088	0.71953	0.81848	0.91765
$w(y^{**})$	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
avg wage	0.23883	0.33138	0.42626	0.52256	0.61976	0.71757	0.81582	0.91438
wmass 1	0.72167	0.74921	0.7671	0.7799	0.78965	0.79738	0.8037	0.80898
wmass 2	0.27833	0.25079	0.2329	0.2201	0.21035	0.20262	0.1963	0.19102
sbar	0.6026	0.6542	0.70665	0.75823	0.80669	0.84614	0.84612	2.162
smass1	0.16623	0.20125	0.24467	0.3069	0.40973	0.61862	1.2842	-13.2337
smass2	0.83377	0.79875	0.75533	0.6931	0.59027	0.38138	-0.28415	14.2337
u	0.18784	0.18116	0.17045	0.15728	0.14252	0.12672	0.11018	0.093116
$u1$	0.33469	0.26397	0.21709	0.18397	0.15942	0.14055	0.12561	0.1135
$u2$	0.12519	0.11379	0.1041	0.095825	0.088711	0.082543	0.077153	0.072409
$u3$	0.16667	0.16667	0.16667	0.16667	0.16667	0.16667	0.16667	0.16667
u_{02}	0.22834	0.19832	0.17328	0.15299	0.13653	0.12303	0.11182	0.1024
v	0.03609	0.061676	0.09015	0.12041	0.15182	0.18402	0.21675	0.24987

Note: each column represents a model solution, which changes as firm productivity varies, shown in the forth row. wmass1 is the percentage of wage employees who only take wage employment in the equilibrium and wmass2 is the percentage of wage employees who also take self-employment in the equilibrium. smass1 is the percentage of the self-employed who also take wage employment in the equilibrium and smass2 is the percentage of the self-employed who only take self-employment in the equilibrium. u is total unemployment, $u1$ is the unemployment rate for $y < y^*$ types, $u2$ is the unemployment rate for $y^* < y < y^{**}$ types and $u3$ is the unemployment rate for $y > y^{**}$ types.

Table 5: VARYING THE LEVEL OF FRICTIONS IN WAGE AND SELF EMPLOYMENT

θ	0.33774	0.51557	0.60718	0.37063	0.48592	0.54655	0.39429	0.49548	0.54885
y^*	0.21506	0.32548	0.37611	0.22077	0.3221	0.37115	0.22459	0.32322	0.37135
y^{**}	0.78947	0.78947	0.78947	0.57223	0.57223	0.57223	0.54126	0.54126	0.54126
m	1	2	3	1	2	3	1	2	3
α	0.5	0.5	0.5	2.5	2.5	2.5	4.5	4.5	4.5
y_w	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$w(0)$	0.33642	0.39156	0.42	0.33907	0.38974	0.41712	0.34085	0.39034	0.41723
$w(y^*)$	0.35753	0.41274	0.43806	0.36039	0.41105	0.43557	0.3623	0.41161	0.43567
$w(y^{**})$	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
avg wage	0.39983	0.43095	0.44835	0.37616	0.41422	0.43547	0.3691	0.41015	0.43254
wmass 1	0.35381	0.46883	0.51696	0.67153	0.74921	0.7799	0.77968	0.83274	0.85289
wmass2	0.64619	0.53117	0.48304	0.32847	0.25079	0.2201	0.22032	0.16726	0.14711
sbar	0.64626	0.73454	0.7792	0.628	0.68941	0.71922	0.62325	0.67908	0.70648
smass1	0.63311	0.47499	0.37039	0.40583	0.28531	0.21277	0.38015	0.27055	0.20409
smass2	0.36689	0.52501	0.62961	0.59417	0.71469	0.78723	0.61985	0.72945	0.79591
u	0.38637	0.28456	0.23346	0.21954	0.18478	0.15884	0.17357	0.1476	0.12584
u1	0.46247	0.25826	0.1762	0.45094	0.26397	0.18397	0.44329	0.26208	0.18365
u2	0.31622	0.20525	0.14981	0.13855	0.11379	0.095825	0.088843	0.07803	0.069228
u3	0.5	0.5	0.5	0.16667	0.16667	0.16667	0.1	0.1	0.1
u02	0.35606	0.2271	0.16238	0.25907	0.19832	0.15299	0.23592	0.18794	0.14773
v	0.12026	0.11709	0.098595	0.09602	0.096368	0.083618	0.093019	0.093119	0.081082

Note: each column represents a model solution, which changes as frictions vary (m and α). $y_w = .5$ for all solutions in this table. wmass1 is the percentage of wage employees who only take wage employment in the equilibrium and wmass2 is the percentage of wage employees who also take self-employment in the equilibrium. smass1 is the percentage of the self-employed who also take wage employment in the equilibrium and smass2 is the percentage of the self-employed who only take self-employment in the equilibrium. u is total unemployment, $u1$ is the unemployment rate for $y < y^*$ types, $u2$ is the unemployment rate for $y^* < y < y^{**}$ types and $u3$ is the unemployment rate for $y > y^{**}$ types.

References

Albrecht, James, Lucas Navarro, and Susan Vroman, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal*, July 2009, *119* (539), 1105–1129.

Badaoui, Eliane, Eric Strobl, and Frank Walsh, "Is There An Informal Sector Wage Penalty? Evidence from South Africa," *Economic Development and Cultural Change*, 2008, *56*, 683–710.

Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky, "The Nash Bargaining Solution in Economic Modelling," *Rand Journal of Economics*, 1986, *17* (1), 176–188.

Broyden, C G, "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Mathematics of Computation*, 1965, *19* (92), 577–593.

Cahuc, Pierre and Andre Zylberberg, *Labor Economics*, The MIT Press, 2004.

Fafchamps, Marcel and Agnes R. Quisumbing, "Marriage, Bequest, and Assortative Matching in Rural Ethiopia," *Economic Development and Cultural Change*, January 2005, *53* (2), 347–80.

Falco, Paolo, Andrew Kerr, Neil Rankin, Justin Sandefur, and Francis Teal, "The returns to formality and informality in urban Africa.," *Journal of Labour Economics*, 2011, *18, Supplement 1*, S23 S31.

Fields, Gary S., "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs," *Journal of Development Economics*, 1975, *2* (2), 165–187.

_ , "Dualism In The Labor Market: A Perspective On The Lewis Model After Half A Century," *Manchester School*, December 2004, *72* (6), 724–735.

Harris, John and Michael Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 1970, *60* (1), 126–142.

Heckman, James J., "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, August 2001, *109* (4), 673–748.

International Monetary Fund, "Internation Financial Statistics April 2010, Series ..DE.ZF," 2010.

Maloney, William, "Informality Revisited," *World Development*, July 2004, *32* (7), 1159–1178.

Mortensen, Dale T., *Wage dispersion : why are similar workers paid differently*, Cambridge, MA. London: MIT Press, 2003.

_ **and Christopher A. Pissarides**, "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, July 1994, *61* (3), 397–415.

_ **and** _ , "New developments in models of search in the labor market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, April 1999, chapter 39, pp. 2567–2627.

Nash, John, "Two-person cooperative games," *Econometrica*, 1953, *21* (1), 128–140.

Rankin, Neil, Justin Sandefur, and Francis Teal, "Learning and earning in Africa: Where are the returns to education high?," 2010. CSAE Working Paper WPS/2010-02.

Roy, A, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Satchi, Mathan and Jonathan Temple, "Labor Markets and Productivity in Developing Countries," *Review of Economic Dynamics*, January 2009, *12* (1), 183–204.

Ulyssea, Gabriel, "Regulation of entry, labor market institutions and the informal sector," *Journal of Development Economics*, January 2010, *91* (1), 87–99.

Wong, Linda Y., "Can the Mortensen-Pissarides Model with Productivity Changes Explain U.S. Wage Inequality?," *Journal of Labor Economics*, January 2003, *21* (1), 70–105.

Zenou, Yves, "Job search and mobility in developing countries. Theory and policy implications," *Journal of Development Economics*, June 2008, *86* (2), 336–355.