WPS4168

Mobility and Earnings in Ethiopia's Urban Labor Markets: 1994-2004

Arne Bigsten[♦]

Taye Mengistae[*]

Abebe Shimeles[♦]

Abstract

An analysis of panel data on individuals in a random selection of urban households in Ethiopia reveals large, sustained, and unexplained earnings gaps between public and private, and formal and informal sectors over the period 1994-2004. We have no formal evidence whether these gaps reflect segmentation of the labor market along either of these divides. In other words, we cannot show whether they are at least in part due to impediments to entry to the higher wage sector. However, we do have evidence that, if segmentation explains any part of the observed earnings gaps, then it could only have weakened over the survey decade. We find, first, that the rate of mobility increased between the two pairs of sectors. Sample transition rates grew across survey waves, while state dependence in sector choice decreased. Second, the sensitivity of sector choice to earnings gaps increased over the same period. In particular, the role of comparative earnings in selection into the informal sector was evident through out the survey decade and increased in magnitude over the second half of the period.

JEL Classification: J42, J62, J31, J45, O17

Key Words: Segmented labor markets, labor market mobility, wage differentials, informal sector, public sector labor markets

World Bank Policy Research Working Paper 4168, March 2007

[♦] University of Gothenburg, Gothenburg, Sweden.

[*] World Bank Group, Washington DC, USA.

1 Introduction[1]

In this paper we investigate the extent to which the structure of Ethiopia's urban labor market might have changed since the early 1990s by looking at changes in the rate of workers' mobility across sectors, in sectoral earnings gaps, and in the sensitivity of mobility rates to earnings differentials. The aim is to see if these have changed in ways consistent with the market getting more or less segmented in the sense of the existence of entry barriers to relatively high paying sectors of employment. Our analysis is based on data from the 1994 to 2004 waves of the Ethiopia Urban Household Socio Economic Survey of the Addis Ababa University and the University of Gothenburg, Sweden.[2] There are good enough theoretical reasons and anecdotes that could lead us to suspect that urban labor markets in low-income economies like Ethiopia's are segmented along the lines indicated in this paper. At the same time, there is consensus that the hypothesis of segmentation is difficult, if not impossible, to test formally in datasets like the one available to us.[3] We nonetheless think that the hypothesis provides a useful framework with the help of which one can analyze changes in the structure of labor markets. Although we cannot directly test its fundamental prediction that there are barriers to entry to certain sectors of the labor market, we can still interpret time series of mobility rates and wage gaps or the relationships thereof as indications of whether labor markets are getting more or less flexible. In other words, absent the possibility of segmentation, these same data would not be as interesting in terms of their implications for the structure or functioning of the market.

Segmentation is indicated whenever there are persistent sector wage premiums that are unexplained in the sense that we cannot attribute them to selectivity or to differences in comparative advantage, unmeasured ability, job attributes or taste.[4] One hypothesized

---

[1] This paper is based on an earlier version prepared as part of a World Bank study of Ethiopia's urban labor markets (World Bank, 2007). We thank Jeni Klugman and Caterina Laderchi , for very useful comments on that and a subsequent version as well as for the support and guidance they provided to our work on the paper.

[2] As discussed in greater detail in section 2 these data are not statistically representative of labor market indicators for urban areas in Ethiopia as a whole. Those can be found, together with a thorough discussion of different labor market data sources for Ethiopia in World Bank (2007).

[3] Although the balance of studies has so far found evidence against the segmentation hypothesis, almost everyone rejecting it are quick to point out serious limitations of data or methods used in the analysis.

[4] Dickens and Lang (1985) define segmentation in terms of the interrelated phenomena of sustained and unexplained wage premiums arising from endogenous wage rigidities and the non-price rationing of "primary-sector" or high-wage jobs arising from it. Magnac (1991) defines a segmented labor

instance of it is between formal sector wage employment and own account work or informal sector employment.[5] A second is between public sector employees and private sector wage earners.[6] If segmentation does indeed exist along either of these divides, then it is also possible that some proportion of the unemployed have been rationed out of paid jobs or out of self-employment by non-price factors.

Existing formal tests of segmentation rely on a host of restrictive assumptions including the log normality of potential earnings and the absence of state dependence in mobility costs, tastes and skills.[7] No attempt is therefore made in this paper to carry out any such tests, or formally establish if Ethiopia's urban labor market is segmented.[8] Instead, we look at changes in variables that define segmentation with a view to assessing whether these changes would have strengthened or reduced it if it in fact existed over the survey period. Unexplained and non-compensating wage differentials that cannot be attributed to differences in comparative advantage or to selectivity are sustainable only if the

market as one in which a) rewards vary across sectors for equally productive workers and b) wages are rigid at least in one (-presumably the high-paying) sector of the labor market as a state in which equally productive worker are rewarded.

[5] The classic story of segmentation along the formal vs. informal dichotomy is Fields' (1975) extension of the Harris -Todaro (1970) model of an underdeveloped economy's labor market quantity-adjusting to shocks through changes in unemployment and rural-to-urban migration in the face of wage-rigidity. Stiglitz (1974), Shapiro and Stigliz (1984) and Bulow and Summers (1985) provide efficiency wage explanations to the wage-rigidity assumed in the Harris-Todaro model. The extension by Fields introduces job search costs that rural-to-urban migrants financed by entering a low wage, low productivity no-barrier-to entry sector upon arrival to towns. Both entry barriers and wage premiums are essential features of Fields' model of segmentation. Rauch (1991) provides an alternative occupational-choice-cum-industrial-organization perspective to the modelling of the duality between the formal and the informal sectors of the economy. Although the dichotomy in Rauch's model is defined in terms of departures from an optimal size distribution of firms rather than from an optimal distribution of workers and earnings per se, it also arises from (government imposed) price rigidity in the labor market.

[6] The large body of empirical work comparing pay rates between the two sectors is overwhelmingly focused on advanced economies, and is concerned not so much with labor market segmentation as with the import of pay gaps to public finances or to the ability of government agencies to attract skilled workers. There has nonetheless been significant interest in public vs. private sector wage gaps in developing economies as well, partly out of similar policy concern as in advanced economies (e.g. Stevenson, 1992; Heller and Tait, 1994), and also in the context of testing the segmentation hypothesis (Lindauer, 1991). Examples include Lindaur and Sabot (1983) on Tanzania, Van der Gaag and Vijerberg (1988) on Cote d'Ivoire, Terrel on Haiti (1993), and Nielsen and Rosholm (2001) on Zambia. Regardless of their sign, unexplained and non-compensating public sector pay premia would be indicative of segmentation, since mobility must be impeded in one direction or the other for them to exist.

[7] Heckman and Hotz (1986) and Magnac (1991) discuss the assumptions in the context of the switching regression test proposed in Dickens and Lang (1985).

[8] On balance empirical studies of labor market studies on middle income developing economies reject the segmentation hypothesis. See, for example, Magnac (1991) on Colombia, Maloney (1999) on Mexico, and Pratap and Quintin (2005) on Argentina.

mobility of workers across sectors is somehow impeded. An increase in the rate of workers' mobility into high-wage sectors over the survey period would therefore indicate decrease in the degree of segmentation. In order to see if the rate of workers' mobility has increased over the survey period, we compare sample transition matrices across survey waves. We also estimate dynamic binary sector choice models for four states, namely, unemployment, informal sector employment, self-employment, and public sector employment.[9] A decrease or an increase in the mobility of workers across sectors over time should be reflected in a decrease or an increase the degree of state dependence in sector choice, the parameter of state dependence being the coefficient of the initial state in the sorting equation. It is important in this context that the parameter of state dependence we estimate measures true state persistence rather than merely reflects unobserved individual heterogeneity. Even then true state dependence could be a consequence as much of state dependence in mobility costs, taste or skills as it could be of entry barriers, of which only the latter defines segmentation.[10] Still, if we can assume that mobility costs and the skill and taste distribution of workers remain the same over time, we can tell whether or not the market is getting more segmented by looking at what happens to the persistence parameter over the same period.[11]

Given that the rate of workers' mobility is usually highly correlated with the size of sectoral earnings gaps, the change in the earnings gap itself is a second (alternative) indicator of change in the degree of segmentation. A third alternative indicator is change in the sensitivity of sector choice to earnings gaps. The less sensitive is state (or sector) choice to pay gaps, the more likely it is that mobility into the higher paying sector has been impeded. To look at these second set of indicators, we compute sector earnings differentials based on estimated sector specific earnings equations. The sensitivity of

[9] For completeness one could also estimate a separate model of binary choice into formal private sector employment. However, the estimation would not add information to that the models of choice into unemployment, informal sector employment and public sector employment provide between them.
[10] Heckman and Hotz (1985) highlight the distinction between mobility costs and entry barriers as an important issues in testing for segmentation.
[11] Maloney (1999) is probably the first to make use of mobility data in conjunction with post transition changes in earnings to test for segmentation. Gong and van Soest (2002) build on this idea to jointly model earnings and inter-sectoral mobility and estimate state dependence in sector choice as one of the test parameters of segmentation. Although we borrow from these two papers the idea of using mobility data, our goal here is less ambitious than theirs since we are not seeking to test for segmentation directly. We are in fact side stepping the task of testing for segmentation, while asking whether our data are consistent with becoming more or less pronounced over the period of observation.

sector choice to earnings gaps is estimated as a parameter of a dynamic structural sector choice model.[12]

Our estimates imply large earnings differentials between the public and the private sectors of wage employment. In the absence of direct evidence on impediments to mobility, these differentials would not necessarily imply segmentation. Still, they should make the hypothesis of segmentation more credible in the Ethiopian context than it would otherwise be. At the same time, the sensitivity of sector choice to earnings gaps seems to have gone up in more recent waves of the survey, not only between the private and the public sectors of wage employment, but also between formal sector wage employment and the informal sector. In particular, the role of relative earnings in selection into the informal sector seems to have gone up substantially.

These results suggest that, if Ethiopia's urban labor market is indeed segmented along the public vs. private or formal vs. informal sector dichotomies, it has got less and less so since 1994. The raw transition matrices we compute from the survey sample support this conclusion as do the parameters of the dynamic sector choice models that we have estimated on the same dataset. We see from the raw transition matrices that, although there was little mobility across sectors prior to the 1997 wave of the survey, mobility rates increased considerably between 1997 and 2000. Mobility rates of the 2004 wave were also at least as high as those of the 2000 wave. The increase in the rates of mobility across sectors of employment was accompanied by a small but persistent decline in the sample rate of open unemployment.

The increase in sample transition rates across survey waves is consistent with changes in state dependence parameters of the dynamic choice model that we estimate. The parameters are all positive and statistically significant in the equation for public sector employment, in the equation for informal sector employment and in the unemployment equation. At the same time, they have become smaller in more recent survey waves in each case.

[12] Our approach here is similar to that of Gong and van Soest (2002), who estimate a dynamic structural sector choice model by using predicted wages (though not premiums) as one of the potential determinants of mobility between the formal and informal sectors.

The organization of the rest of the paper is as follows. In the next section we describe our data. Section 3 discusses patterns and rates of mobility across states in terms of raw transition matrices. Section 4 describes our estimation and testing framework. We present results of estimation of (reduced form) dynamic sector choice models in Section 5, and those of their structural counterparts in Section 6. We conclude in Section 7.

2 Data

The first wave of the Ethiopia Urban Household Socio Economic Survey took place in 1994. There have been four waves since then, one in each of the years 1995, 1997, 2000 and 2004. Some 9,000 to 10,000 individuals in 1500 to 1600 households were covered in each of these waves. A major point of strength of the dataset generated by the survey is that it has a sizeable panel component. A significant proportion of the individuals covered in the first wave were also tracked by all four subsequent waves. An even higher proportion of them were covered by at least three waves. A weakness is that the survey samples were drawn exclusively from the country's seven largest urban centers. While these account for the bulk of the urban sector of the economy, we would hesitate to project details of our results to the broader (national) urban labor market. The smaller towns may differ in terms of their labor market structure from the major centers covered by the survey. That said, we would be surprised if the more qualitative aspects of our conclusions did not hold by and large for the urban sector countrywide.

Table 1 provides the breakdown of the survey sample by age groups and labor market states. In it we limit the labor force to those aged 15 to 64 years inclusive. This leads to a labor market participation rate of about 54%, which does not seem to have changed much over the decade spanned by the surveys. The table shows the division of the economically active between the unemployed, on the one hand, and the employed of various categories. The unemployed are defined as those household members of the 15-64 age group who were not self-employed, or working for anyone else and were actively looking for paid work. Among the employed, the table distinguishes between the self-employed and those working for someone else.

Self-employment is understood to include own account workers and owner manages of small businesses. We define formal sector employment as wage or salary employment in a private company, in a state-owned business or in a government agency. The term formal private sector worker refers to wage earners in private companies. We have put family workers, domestic workers and casual laborers under the category 'other private sector workers'. In the context of this paper the informal sector consists of the self-employed and these 'other private sector workers'. We include in the 'formal sector' employees of private companies and the public sector. [13]

We provide in table 2, descriptive statistics for monthly earnings, annual sales revenue of owned businesses and selected covariates of these and of labor market status. The covariates include age, gender, schooling, and occupation. Statistics are given by survey wave as well as for the observations pooled across waves. With the exception of earnings and business sales revenue all variables in the table are dichotomous. Schooling is presented using two classifications of school attainment. One of these is fairly standard, needs little explanation, and is used to report the sample proportions of those who attained a particular level for the 1994, 1997 and 2000 waves. Unfortunately the same classification was not used in the 2004 wave of the survey, which distinguished only between those who had no education, those who had attained grade 10 or lower, those who had completed preparatory school, and those who had some tertiary education. Other covariates highlighted in the analysis include location of employment, ethnicity and family background variables, including parental schooling and parental occupation.

3. Aggregate transition rates and patterns

As a prelude to our analysis of the determinants of labor mobility, we describe in this section the patterns of mobility observed in the data in terms of sample transition matrices between employment and unemployment, and across three sectors of employment, namely, the public sector, formal private sector employment, and informal sector employment.

---

[13] There is no employer size dimension to our definition of informality. People who have reported to be employees of a private sector organization or enterprise have all been classified here as formal sector workers. It is common to define informality in a way that takes into account the size of employers. There are a variety of operational definitions of informality with which our admittedly lose characterization is consistent. One definition identifies informality with lack of labor regulation (Pratap and Quintin, 2006). A second is in terms of employer size (Maloney 1999, Gong, van Soest and Villagomez, 2004). A third is in terms of the kind of work being done such as casual or piece rate work and also own account work, such as casual work or self employment (Magnac, 1991). A fourth alternative is to define firms as informal if they are not legally registered (Bigsten, Kimuyu, Lundvall, 2004)

There was not much transition across these sectors between the 1994 wave and the 1995 wave, or between the 1995 wave and the 1997 wave. On the other hand, transition rates were relatively high between the 1994 wave and the 1997 wave, the 1997 wave and the 2000 wave, and the 2000 wave and 2004 wave, and, as expected, even higher between any pair of waves that were further apart by more than four years. We report the sample transition matrices for these pairings in table 3. Two key patterns seem to emerge across the four waves. First, there was a small but statistically significant and persistent decline in the sample unemployment rate since the 1997 wave. Secondly, this was associated with an increase in the share of formal sector in employment.

Unemployment transitions

The probability of someone reporting unemployment in the 1994 sample landing a job or shifting to self-employment within a year was almost zero. However, the probability of someone in that state exiting unemployment increases as we expand the time horizon of observation. Panel a) of table 3 tracks the respondents of the 1994 wave through to the 1997 wave. Although the table registers some movement out of unemployment, this occurs at very low rates: less than 7 percent of those who were unemployed in the 1994 sample had some form of employment by 1997. A little over a third of these went into formal sector wage employment.

Things do change significantly over the next three years. By the 2000 wave, more than half of the unemployed of the 1994 sample had changed status (table 3, panel b). About one in eight had left the labor force altogether, presumably, having stopped looking for work. About one in five went into own account work or into informal wage employment. Another 11 percent got jobs in private firms. And approximately 9 percent landed public sector jobs, about one third of these being in state owned companies. By the 2004 wave, less than a third of the unemployed of the 1994 wave were still available for work (panel c). Of the rest, about 23 % left the labor force; some 19 % joined the informal sector (in self-employment or in informal wage contracts); 17 % landed jobs in private firms; and about 12 % got public sector jobs.

There are some interesting comparisons of rates of transition out of unemployment between the periods 1994-1997 and 1997-2000. The rate of absorption of the unemployed into the formal sector increased from 2 % to 18 %. The rate of absorption by the informal sector (both self-employment and wage contracts) increased from about 5% to 19%. The rate of absorption by the public sector increased from under 1% to almost 8%. The rate of absorption by the formal private sector increased from under 2% to about 10%.

The relatively high rates of transition from unemployment during the period 1997-2000 were also sustained in the next four years. The probability of someone reporting unemployment in the 2000 sample taking up some form of employment by the time of the 2004 survey was 41 percent. This is significantly higher than the corresponding transition rate for the period 1997-2000, but by an amount that could be explained entirely by the fact that the second interval is one year longer.[14]

Informal sector/self-employment transitions

Like rates of transitions out of unemployment, probabilities of transitions between the formal and informal sectors of employment would seem to be negligible over any given year. But, again, the probability of transition rises significantly when we expand the time horizon by three more years (table 3, panel b). This is largely on account of movement of workers to the private formal sector. The probability that someone reporting self-employment or informal employment in the 1994 survey would have joined the public sector by the 2000 survey was only 5 %. The probability that the same individual would have taken up a job in a private firm was about 15 %. These probabilities do not change much as we expand the horizon further to 10 years (table 3, panel c).

[14] Other related transition rates of interest for the years 2000-2004 are the following: About 24 % of the unemployed of 2000 sample landed formal sector jobs by 2004; another 17 % became self-employed, or entered into informal job contracts; some 13 % joined private firms; about 11 % joined the public sector. To illustrate the origins of the unemployed, some 17 % of those who were out of the labor force in the 2000 sample joined the unemployment pool by 2004 (table 3, panel e). Given that the out-of-the-labor force group is by far the largest category of survey respondents in any year, this suggests that fresh entry into the labor force must be the largest source of entry into unemployment in the economies from which the sample was drawn. The second largest source were private sector job losses were the second important source of gross addition to employment for the same period. The probability of a formal private sector job loss was 14 %. The probability of an informal sector job loss was even higher at 17%. This contrasts with 7-8 % rate of public sector job losses. Given that there were more private sector employees than there were public sector workers in the 2000 sample, a larger proportion of the unemployed of 2004 must have come from the private sector than did from the public sector.

To see if the rate of mobility from informal sector employment to formal sector jobs has increased over time, we again compare panel a, which captures transitions over 1994-1997 with panel d, which relates to 1997-2000. As is the case with transitions from unemployment to employment, there was a drastic change here also: the probability of transition from informal sector wage employment to formal sector jobs rose from less than 1 percent for 1994-1997 to 12 % for the period 1997-2000, and 15% for the period 2000-2004.[15]

Public sector transitions

There was a 48% chance that someone who worked for the government (as opposed to a state owned company) in the 1994 sample would have left that particular sector by 2004. We arrive at this rate by adding the following probabilities from panel c: that the person retired or otherwise left the labor force (20.5%); that they were laid off or pensioned off but did not leave the labor force (10%); and that they joined the private sector (17%). More than half of those who moved to the private sector would have worked for a formal sector firm. The corresponding probabilities of transition for a public enterprise worker were: 35% of leaving the labor force, 10% of being unemployed, and 16% of joining the private sector. Two-thirds of those joining the private sector would have joined a formal sector firm.

These rates reflect more the higher transition probabilities of the intervals between the fourth and fifth (2000-2004) and the third and fourth (1997-2000) waves of the survey (panel d and panel e) than the rates of the first three waves. As can be seen from panel a there was practically no movement from the public sector during the interval between the first and the third waves (1994-1997). On the other hand, there was a 12 % chance that someone who was a government employee in 2000 would have joined the private sector in self-employment (4 %) or as an employee (8%) by 2004. The probability of a

[15] To illustrate where the informally employed come from, more than one in ten of people who were out of the labor force in the 2000 sample would have become informal sector wage worker or self-employed by 2004. And just under 20% of those who were unemployed in the 2000 sample would be self employed or in informal wage work by 2004. These rates are much larger than the probability of transition from the formal sector. Because the combined share of the out-of-the labor force and the unemployed was 60% of the full sample of individuals in the 2000 sample, this means that by far the largest fraction of those who joined the informal sector by 2004 could only have come from these sources. Indeed, the probability that anyone who had a formal sector job in the 2000 sample would be found in self employment or in informal paid jobs in 2004 was quite low, standing at 6% for a government worker, at 8% for a public enterprise employee, and at 12 % for an employee of a private firm.

public enterprise employee in the 2000 sample moving to the private sector by 2004 was even higher, at 19%.[16]

4 Estimation and testing

Suppose we would like to know whether the labor market is getting more or less segmented between any two sectors over time: say, formal sector wage employment vs. informal sector employment. One way of investigating this is to see if the probability that anyone who starts out in the informal sector in year $t-1$ will have remained in the same sector in year t has increased, decreased, or remained the same over time. The more mobile are workers across any divide, the poorer is status at time $t-1$ as a predictor of status at time t. A way of measuring changes in the degree of segmentation is therefore to use changes over time in the extent of state dependence in sector choice. In doing so we leave open the possibility that mobility costs are state dependent in the sense, for example, that prospective entrants to formal sector or public sector employment might have to migrate from elsewhere at significant transport costs, while similar costs could be sunk or entirely absent for those already in those sectors. We also allow for the possibility of sector "scarring effects" as defined below as well as for skill formation and skill depreciation rates to vary by labor market state. As already noted, any of these factors would affect the rate of inter-sectoral mobility and, consequently, result in true state dependence in sector choice, but is not part of what defines market segmentation.

In estimating the degree of state dependence based on our sample, we assume that the current distribution of individuals in terms of belonging to a sector or not is generated by the process

$$(1) \qquad \begin{cases} y_{it} = 1(y^*_{it} > 0) \\ where \\ y^*_{it} = X_{it}\beta + \delta y_{it-1} + \varepsilon_{it}, \end{cases}$$

[16] Looking at transitions in the reverse direction, some 12 per cent of those who were working for private firms in the 2000 sample had joined the public sector by 2004. Some 9.5% of informal sector employees and 3 % of self-employees in the same sample moved to the public sector over the same period. Again these rates are much higher than the negligible transition probabilities between 1994 and 1997, but are only a little higher than the rates for the period 1997-2000.

X_{it} is a vector of exogenous observable determinants of such status, ε_{it} is a random variable summing up all unobservable influences on the same; β and δ are constants; and y_{it}^{*} is a latent variable registering a threshold value of comparative earnings or of utility governing individual's sector choice, or a critical value of a measure of entry barriers to which the choice is subject. Following most of the relevant empirical literature on labor market transitions, we assume that ε_{it} is distributed normal with zero mean and constant variance σ_{ε}^{2} . There is state dependence in the process when $\delta \neq 0$, so that the probability $\Pr(y_{it} = 1 | X_{it}) \equiv E(y_{it} | X_{it})$ is a function of y_{it-1}. In the present context we should expect δ to be positive in the event of its being different from zero. In that case a larger value of δ would imply greater state dependence in the sense that y_{it-1} would be a stronger predictor of y_{it}.

We use equation (1) to describe any one of four distinct processes in Ethiopia's labor market: unemployment vs. employment in general, informal vs. formal sector employment, self-employment vs. wage employment, and public vs. private sector employment. Suppose (1) relates to the unemployment process, where $y_{it} = 1$ if i is unemployed at time t, but $y_{it} = 0$ otherwise. In the literature on unemployment, a larger value of δ could be a result of longer unemployment leading to loss of human capital through disuse, which would make the unemployed less attractive than the employed as job candidates (Pissarides, 1992). It could also reflect the "scarring effect" of unemployment if prospective employers use unemployment history as a rule for screening out inferior workers (Arulampalam et al. 2002, Phelps, 1972). A third possibility is that getting a new job involves a fixed relocation or transport cost that becomes sunk once incurred. While each of these three scenarios is consistent with competitive labor markets, a higher δ could also be the result of segmentation, that is, a result of barriers of entry into employment on the part of the unemployed, due to employment protection laws, the exercise of insider power as in Lindbeck and Snower (1986), or jobs being rationed via efficiency wage payment (as in Shapiro and Stiglitz, 1984; and Bulow and Summers, 1986).

Likewise, state dependence in the sorting of workers between the informal and formal sectors could also arise from segmentation (by entry barriers to the formal sector) or from any one of the three "competitive" sources of state dependence listed above for the case of unemployment. If we let $y_{it} = 1$ when i is an informal sector worker, and $y_{it} = 0$ otherwise, a larger value of δ could reflect the fact higher barriers of entry to the formal sector arising from the protection of formal sector jobs by law or via insider power, or some combination of formal sector efficiency wage payment (Stiglitz, 1974) and search costs of formal sector jobs (Fields, 1975) in a Harris-Todaro setting (Harris and Todaro, 1970). This would be the segmentation scenario. Alternatively, a large δ might simply be caused by the fact that staying in the informal sector involves loss of skills relevant to formal sector employment or has some kind of scarring effect similar to that associated with unemployment, or because the informal sector workers would have to incur reallocation costs to formal sector jobs.

Similarly, a higher value of δ could reflect higher entry barriers to the public sector if equation (1) describes the allocation of workers between that sector and private firms. Public sector jobs are probably more often life time jobs than private sector jobs in many low-income economies. At the same time their availability for new entrants to the labor force is limited by government finances.

Given any of the three binary sector choice situations discussed above, we would have no way of telling what component of the δ we might estimate for any of them reflects segmentation and what part does not. We can nonetheless tell if segmentation has increased, decreased or remained the same along any of the dichotomies by estimating δ over the various segments of the survey decade under an identifying assumption. Specifically, we assume that the distribution, across sectors, of mobility costs, skill formation/depreciation rates, scarring effects, and other possible 'competitive' sources of state dependence across sectors does not change over the period of interest, so that the share of δ that we can attribute to these sources is constant over time. A fall in the degree of state dependence in sector choice would then indicate decline in segmentation while an increase in the same would mean greater segmentation.

However, this is only one aspect of the problem of identification that we face in trying to estimate equation (1). Segmentation or not, our estimate of δ as a measure of true state dependence in general would be consistent only if the initial values, y_{i0}, of labor market states as observed at the start of the survey are exogenous (Heckman, 1981).[17] If y_{i0} is endogenous, a positive estimated value of δ would signify wholly or partially what Heckman (1981) characterised as spurious state dependence, that is, as a case of upward bias that the endogeneity gives rise to in the estimation. The initial condition would be exogenous only if (a) ε_{it} are serially independent, and (b) the start of the survey coincides with the onset of the process itself. The latter being unlikely the case, it is conventional to assume that y_{i0} is endogenous, determined as it is by the selection rules of the survey sample. The assumption that ε_{it} is serially uncorrelated would also be unrealistic since it would rule out ε_{it} including persistent individual effects. Yet such effects could well be present since we are unlikely to observe all potential determinants of sector choice or the all the constraints under which it is made. Unobserved heterogeneity of this kind generate serial correlation in the error term, which in turn produces correlation between the error term and y_{it-1} and y_{i0}. We will therefore think of ε_{it} as consisting of two components: a term summing up persistent elements of unobserved heterogeneity, c_i, that is potentially correlated with y_{i0}, and an i.i.d component, u_{it} uncorrelated with y_{i0}, y_{it-1} and X_{it} such that

(2) $$\varepsilon_{it} = c_i + u_{it}$$

This in turn makes ε_{it} serially correlated with correlation coefficient $\rho = \dfrac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2}$, The correlation between c_i and y_{i0} in turn biases δ upwards (Stewart, 2005). Heckman (1981), Arulampalam et al. (2000), Wooldridge (2005), and Stewart (2005), discuss alternative ways of avoiding this bias. Heckman's estimator involves estimating (1) jointly with a probit model for the initial condition

(3) $$y_{i0} = 1(W_i\gamma + u_i > 0)$$

where W_i is a vector of exogenous variables including pre-survey characteristics of the sample not included in X_{it} possibly along with some of the exogenous elements of X_{it}, γ is a vector of constants, and u_i is a normal error term uncorrelated with u_{it}, but

[17] Consistent estimation would also be possible with serially uncorrelated errors if it can be assumed that the process is in equilibrium at the start of the survey. However, the later assumption would be unrealistic as it excludes the time varying exogenous variables from the process (Heckman, 1981).

possibly correlated with c_i. Heckman's estimator is obtained by maximizing the likelihood of $(y_{i0}, y_{i1,} ..., y_{it}, ..., y_{iT})$ conditional on (X_{it}, Z_i) as given by (1) to (3), while treating c_i as random effects. The reason for the random effects formulation is that MLEs of parameters of (1) would be biased for any finite time series length, T, of the panel under the assumption of fixed effects. On the other hand, MLEs of parameters of (1) are consistent as the cross-section size, N, tends to infinity for given series length if individual effects are random (Heckman, 1981).[18]

Treating c_i as random effects in turn amounts to assuming that they are uncorrelated with observable influences on y_{it}^*. This, however, is not as restrictive as it might seem. It can be relaxed by assuming that c_i is a linear combination of time means, $\overline{X}_i$,of the time varying elements of X_{it}, and a random effects component α_i distributed normal with mean zero and variance σ_a^2, but assumed to be uncorrelated with X_{it} such that $c_i = \overline{X}_i \lambda + \alpha_i$ (Wooldridge 2000). In that case the primary equation of Heckman's estimator should be written as the following random effects probit:[19]

(4) $$y_{it} = 1(X_{it}\beta + \delta y_{it-1} + \overline{X}_i \lambda + \alpha_i + u_{it} > 0)$$

Heckman's estimator deals with the initial condition problem of identification of true state dependence by instrumenting the initial conditions, that is, by maximizing the joint likelihood of $(y_{i0}, y_{i1,} ..., y_{it}, ..., y_{iT}) \mid X_{it}, W_i$. Wooldridge (2000, 2005), proposes an alternative estimator that maximizes the joint likelihood of $(y_{i1,} ..., y_{it}, ..., y_{iT}) \mid X_{it}, y_{i0}$, that is, maximizes the post-initial sub sample conditional on the initial condition and exogenous covariates of y_{it}. The underlying idea is that, if y_{i0} is endogenous while u_{it} is serially uncorrelated, then c_i must be correlated with y_{i0}. Since there can also be exogenous covariates of y_{i0} that may produce individual effects independently of their influence on initial conditions, Wooldridge assumes that $c_i = \pi_0 y_{i0} + Z_i \pi + a_i$, where Z_i are elements of W_i correlated with c_i independently of y_{i0}, and a_i is distributed

[18] See Wooldridge (2000), Hsiao (2003) and Stewart (2005) for reviews of the literature on consistent estimation of state dependence in the context of a probit model with unobserved heterogeneity.
[19] This is as opposed to the traditional random effects probit defined by equations (1) and (2).

normal with zero mean and variance σ_a^2 so that the likelihood function is written in terms of

(5) $$y_{it} = 1(X_{it}\beta + \delta y_{it-1} + \pi_0 y_{i0} + Z_i\pi + a_i + u_{it} > 0)$$

Again the likelihood is maximized by treating a_i as random effects.

Because it is easier to implement than Heckman's estimator, we use Wooldridge's estimator. It turns out that the results that we obtain are quite similar to those we get by using a two stage procedure proposed by Arulampalam, Booth and Taylor (2000).[20] The procedure estimates in two steps Heckman's specification (Heckman, 1981) as given by equations (1) to (4) above. The first stage involves the estimation of (3) by maximum likelihood as a simple probit. In the second stage one estimates the following random effects probit by maximum likelihood:

(6) $$y_{it} = 1(X_{it}\beta + \delta y_{it-1} + \overline{X}_i\lambda + b_1\hat{u}_i + \tilde{v}_i + u_{it} > 0)$$

where $\hat{u}_i$ is the generalized residual from the first stage probit, replacing the true error component u_i in the equation

(7) $$y_{it} = 1(X_{it}\beta + \delta y_{it-1} + \overline{X}_i\lambda + b_1 u_i + v_i + u_{it} > 0)$$ and

$\tilde{v}$ is such that $b_1\hat{u}_i + \tilde{v}_i = b_1 u_i + v_i = a_i$.

We can test for the endogeneity of initial conditions based on Wooldridge's estimator by testing for the statistical significance of π_0 in equation (5). The corresponds to the test for the statistical significance of b_1 in equation (6) in the context of the two-step estimator. Since endogeneity is controlled for in each case, the estimate of δ we obtain by maximizing the log likelihood of the ample according to either equation should be consistent for true state dependence in choice into the labor market state of interest. In order to assess the merit of the two estimators relative to estimates from maximum likelihood estimates of a simple dynamic probit, or a traditional random effects probit that would assume the exogeneity of initial conditions, we have estimated equations (5) and (6) on the full five-wave panel over the period 1994-2004.

The objective of the analysis being to gauge possible changes in δ over the same period, we also estimate each specification over two consecutive segments of the decade, namely, the period 1994-2000 and the period 1997-2004. The panel consisting only of five waves, we had to choose overlapping periods in order to make sure that the estimation of δ

[20] Henley (2004) is an example of more recent work using the two-step estimator.

controls for unobserved effects. As already noted the five survey years are 1994, 1995, 1997, 2000 and 2004. This means the two periods overlap over the years 1997-2000. The value of δ therefore can change between the two periods only in as far as the value of δ for 1994-1997 is different from the value of the same for 2000-2004.

We would interpret a higher value of δ for the 1994-2000 observation than that for the 1997-2004 observations as indication increase in the degree of segmentation of the labor market over the survey decade. An alternative measure of change in the degree of segmentation is the corresponding change in the sensitivity of mobility to wage gaps that one observes over the period of interest. Considering equation (3), let w_{it} be the earnings of person i in the sector of current employment, and w_{0it} the person's potential earnings in the alternative sector. If workers respond to sector wage gaps in deciding where to work, then $w_{it} - w_{0it}$ should be one of the determinants of employment status, y_{it}. We assume that the relationship between pay gaps and sector choice is such that

(8) $$y_{it} = 1(a_1(w_{it} - w_{0it}) + Z_{1it}\theta + \xi_{it} > 0)$$

where a_1 is a constant; Z_{1it} is a vector of exogenous individual characteristics, θ is a vector of parameters (including a constant term), and ξ_{it} is a zero mean iid error term orthogonal to pay gaps and to Z_{1it}. The proposition that workers' sector choice and mobility decisions depend on pay gaps can be tested by estimating equation (8) and testing for the statistical significance of a_1. More importantly in the present context, an increase in a_1 over time would suggest that segmentation has increased in the sense that mobility has become more and more sensitive to sectoral earnings gaps. Conversely, if segmentation grows over time in the sense that workers have become less responsive to pay gaps this should be reflected in a_1 diminishing over time.

A problem that has to be tackled in trying to estimate a_1 on the basis of (8) is that, although each worker might have a reasonable idea of what they would earn in alternative states of employment, we can observe only one of these, which is the wages of current employment. A commonly used solution to this problem is to replace the unobserved wage by a counterfactual pay rate implied by an estimated earnings equation.

To obtain the estimates we assume that w_{it} and w_{0it} are linear in a set of parameters such that

(9) $$w_{it} = X_{1it}\Pi_1 + \zeta_{1it}$$

and

(10) $$w_{0it} = X_{1it}\Pi_2 + \zeta_{2it}$$

where X_{1it} is a vector of exogenous covariates of earnings, Π_1 and Π_2 are vectors of parameters (including constant terms), and ζ_{1it} and ζ_{2it} are iid error terms orthogonal to X_{1it}.

The counterfactuals that one would compute based on the OLS estimation of equations (9) and (10) will be biased if there are unobservable worker attributes that induce workers to self select into either sector and at the same time enhance or reduce their earnings in the same sector. In order to avoid this bias we estimate the two equations having added a selectivity term based on the estimation of equation (6). Let $\Lambda_1(\hat{y}_{it}) = \frac{\phi(\hat{y}_{it})}{\Phi(\hat{y}_{it})}$ where ϕ is the density function of ξ_{it}, Φ the corresponding cumulative distribution function, and $\hat{y}_{it}$ is predicted value of y_{it} that one would obtain from parameter estimates of equation (2) and $\Lambda_2(\hat{y}_{it}) = \frac{\phi(\hat{y}_{it})}{1-\Phi(\hat{y}_{it})}$. We estimate the earnings equations by applying OLS to

(9b) $$w_{it} = X_{1it}\Pi_1 + \sigma_1\Lambda(\hat{y}_{it}) + \varsigma_{1it}$$

and

(10b) $$w_{0it} = X_{1it}\Pi_2 + \sigma_2\Lambda(\hat{y}_{it}) + \varsigma_{2it}$$

where σ_1 and σ_2 are constants and ς_{1it} and ς_{2it} are iid error terms orthogonal to all other right hand side variables in their respective equations. The structural sector choice equation we actually estimate is

(8b) $$y_{it} = 1(a_1(\hat{w}_{it} - \hat{w}_{0it}) + Z_{1it}\theta + v_{it} > 0)$$

where the hat symbol indicates estimates based on the estimation of (9b) and (10b).

5 State dependence in sector choice

The most important finding emerging from our estimates of dynamic binary sector choice models is that, while there has always been state dependence along the three divides of Ethiopia's urban labor market, it has grown significantly smaller over the survey decade. The estimates show that the probability of being in any one of the four sectors during a particular survey wave is higher for those who were in that state in the preceding wave, but this effect of last-period state on current choice has grown weaker in more recent waves. We interpret this as evidence that the rate of inter-sectoral mobility has increased over the years.

Details of the finding are reported in tables 4, 5 and 6, where we estimate a dynamic probit for each of the four labor market states. The probit for informal employment is estimated under alternative definitions, namely, one in which we equate informal sector employment with self-employment, and a second in which the category also includes those who work for others under informal wage contracts. In the first panel of table 4, we estimate a dynamic probit for each of the four labor market states on observations pulled across all five waves on the assumption that there are no unobserved individual effects and that the initial labor market state observed in the first wave of the survey is exogenous. Since neither is necessarily true in the light of the discussion in section 4, what we are reporting in the panel are estimates of a baseline model rather than estimates of the true data generating process.

Indeed we do show that neither assumption is correct in table 5 with respect to our data. Table 5 is also where we compare the results of addressing the initial condition problem using the Wooldridge's estimator and the two step estimator. As part of implementation of the latter we estimate a simple probit model of employment status at the time of the 1994 wave of the survey and report it in the second panel of table 4. As a selectivity equation, this is identified by including family background variables in it having excluded them from the first panel and from the specifications in tables 5 and 6. The underlying assumption is that these variables affect current status only in as far as they influence initial conditions. It turns out that both paternal occupation and maternal occupation are significant influences on initial conditions. Interestingly, individuals whose fathers owned non-farm businesses were less likely to be to be found in the public sector in 1994, but they were also less likely to be unemployed, as they were more likely to be self-employed or to work for someone else in informal contracts. Also, maternal occupation did not seem to influence the chances of initial public sector employment. Those who had working mothers were more likely to be unemployed as they were less likely to work in the informal sector as employees or as own account workers. Surprisingly, parental education seems to be only weakly correlated with initial labor market status.

The level of own education does influences the same status. Notably, those who had secondary education as the highest level of schooling were more likely to be unemployed in 1994 than those who had less schooling than that and those who had some tertiary education. This is in spite of the fact that there was no significant correlation between education and public sector employment, and reflects the fact that the more educated were less likely to be found in the informal sector or in self-employment at the time. Nevertheless, the public sector attracted the more skilled than the private sector in 1994 while informal sector employment was associated with lower skills, when skills are measured in terms of broad occupation groups. Gender and age were also significant factors in sector choice in 1994. In particular, women were more likely to be self-employed. Public sector employment and self-employment were also positively correlated with age, while the unemployed were predominantly in the 15-29 age group.

Turning to table 5, we report Wooldridge's estimator of each sector's probit in the first panel, and the corresponding two-step estimates in the second panel. Results of the first panel are obtained by estimating equation (5) by maximum likelihood. Those in the second are also maximum likelihood estimates of the specification given by (6). There is clear evidence of endogeneity of initial conditions in both panels. This can be seen in the fact that all coefficients of y_{1994} are statistically significant in the first panel. Likewise, the coefficients of the generalized residual of the initial status probit of the first stage are statistically significant. The fact that both sets of coefficients are positive means that the coefficients of y_{t-1} in the first panel of table 4 would be biased upwards for true state dependence due to the endogeneity of initial conditions. However, what is left when we remove the bias is still statistically significant: the coefficients of y_{t-1} are positive and significantly different from zero in both panels of table 5. It is also clear that there are unobservable influences on sector choice beyond the observable factors controlled for in the table. However, the same unobservable variables happen to be highly correlated with time variant observables, namely, education, age, and occupation. Indeed, there is no evidence of unobserved effects that are uncorrelated with those observables. This can be seen from our estimates of σ_a and σ_v, which are reported in the first and second panels of the table, respectively, and are not statistically significant.

Since estimation results are quite similar across the two panels of table 5, we confine ourselves to the use of Wooldridge's estimator in table 6, where we compare state dependence in sector choice between the periods 1994-2000 and 1997-2004. As already noted, this amounts to comparing estimates for 1994-1997 with those for 2000-2004. The conclusion that state dependence in sector choice has diminished over time can be read from the first row of the table, where each coefficient in the first panel is significantly lower than the corresponding entry in the second panel. For example, the coefficient of y_{t-1} in column 1 is significantly smaller than the coefficient of the same in column 5, suggesting that the average unemployment duration for 2000-2004 was smaller than that for 1994-1997. Likewise, the coefficient of y_{t-1} in column 4 is significantly smaller than the corresponding coefficient in column 8, implying that past public sector employment was a weaker predictor of current public sector employment during the period 2000-2004 than it was during 1994-1997. The same can be said of informal employment in general and self-employment in particular: a history of employment in either form was a weaker predictor of being in the same type of employment in 2000-2004 than it was in 1994-1997. [21]On the assumption that the combined effect of competitive sources of true state dependence of sector choice has not changed significantly over the survey decade in all three cases, the observed fall in the persistence parameters over the same period can only indicate weakening segmentation if there are indeed entry barriers to formal sector employment in Ethiopia.

6 Earnings gaps and sector choice

We now turn to our second approach to gauging changes in the degree of segmentation, which is estimating the sensitivity of sector choice to sectoral earnings gaps. In tables 7, 8, and 9 we report results of estimation of a structural mode (equation 8b) of the sorting of workers between the public and private sectors of wage employment. In column 4 of

[21] At the same time, there is no evidence that the influence of any of the key observable determinants of labor market state has changed between the two periods. This includes gender, age and schooling. The only exception to this is that age effects in public sector employment seem to have come down significantly.

table 7, the model is estimated by maximum likelihood, on the formal sector employees sub-sample of the 2004 wave that were also covered in the 2000 wave. The earnings gap that enters the equation is obtained from sector specific earnings equations that we report in the first (public sector) and the second (private sector) columns of the same table. These are obtained by applying OLS to the specifications given by equations 9b and 10b, respectively. The selectivity terms of these columns are obtained from a binary dynamic public sector choice model of the same specification as column 6 of table 4-that is, as given by equation 5- but this time on data points of wage workers only. In the third column of the table we estimate a single earnings equation for both private and public sector workers having included a public sector dummy. Since the dummy is assumed to be endogenous in this context, we include the generalized residual of equation 5 as a selectivity correction term in this column.

Tables 8 and 9 are in similar to table 7 in content, in terms of underlying specification and identification and estimation methods. Table 8 reports results of estimation on data on wage earners in the 2000 survey sample who were also covered in the 1994 survey. Table 9 relates to wage earners of the 1997 wave who were also covered in the 1994 wave.

In table 10 the focus is on the estimation of a structural model of the sorting of workers between formal sector wage employment and informal employment. Here too the assumed structure is given by equation 8b. Since there are practically no earnings data for informal sector wage workers, we have been forced to confine the analysis to data on formal sector wage workers and own account workers reporting business sales revenue. We use the latter as a proxy for earnings from self employment. What we refer to in the table as self-employment earnings premium is in fact the log difference between annual business revenue and the counterfactual annual wage, which we assume to be monotonically increasing in the true gap between earnings from self-employment and wages. The equations we estimate in table 10 are very similar to those of tables 7 to 9 except that we do not include a common earnings equation across the two divides in the case of table 10. The sector-specific earnings equations are identified and estimated in the same way in both sets of tables, as are the underlying reduced form selection equations.

Looking at the coefficients of the public sector dummy of the third column of tables 7 through to 9, we see that there has been a sizeable public sector wage premium throughout the survey decade, which, if anything, has been growing over time. The implied public sector premium (at the mean of the wage distributions) was 49 % for 2004 as compared to 40% for the year 2000 wave, and 32% for 1997. However, the coefficients of the public sector wage premium in the last columns of the three tables also suggest that the premium has become an increasingly more powerful driver of selection into public sector employment. Other things being the same, this would indicate diminishing segmentation over time, which is consistent with what we see both in terms of increase in raw transition rates and in the form of diminishing state dependence in selection into the public sector.

It should perhaps be stressed that the public sector premium is computed over and above possible sector differences in rates of return to schooling and to market experience. It turns out that in spite of the positive public sector premium, the rate of return to observed human capital-i.e., to schooling and experience-was higher in the private sector in the 2000 wave. This was reversed quite drastically in the 2004 wave of the survey, for which public sector rates of return were significantly higher. This was in addition to the 49 percent residual public sector pay premium.[22]

Turning to table 10, the key finding that emerges from it is that engagement in the informal sector is the outcome of active choice based on comparative advantage for at least for some proportion of those found in the sector. The evidence for this is that the coefficient of relative earning in the last column of each of the three tables is always positive and statistically significant, suggesting that anticipated earnings gains are a driver of selection into informal sector employment. The size of the coefficient dropped precipitously between the 1997 and the 2000 waves, but then picked up in 2004 to more than double of the estimate for the 2000 wave.

This role of comparative advantage in informal sector employment could be concealed by the fact that schooling is negatively correlated with selection into self-employment.

[22] We should hasten to add that this result applies to comparison at the mean of the distributions. It probably is the case that the public sector premium gets more pronounced towards the lower end of the distribution, but turns increasingly negative as we moved to upper quintiles.

However, it also turns out that the rate of return to education in informal sector employment that we read from the three tables is comparable to that of the formal sector. In other words, while the less educated have greater propensity for self-employment, the more educated are more successful among the self-employed.

7 Conclusion

Based on data from the Ethiopia Urban Household Socio Economic Survey over the period 1994 to 2004, this paper has assessed the extent to which the structure of the urban labor market has changed over the same period. Specifically, we have looked at what has happened to entry rates to the public and the formal sectors of the labor market and to sectoral earnings premiums over the survey period, as well as to the sensitivity of inter-sectoral mobility to earnings gaps. Our data show a large, persistent and unexplained public sector wage premium. However, the sensitivity of sector choice to earnings gaps has become more pronounced, not only in relation to the public vs. private sector divide, but also vis-à-vis the formal and informal sector dichotomy. In particular, the role of comparative earnings in selection into the informal sector has increased in recent years. The rate of workers' mobility has also increased between the two pairs of sectors since the late 1990s as indicated by sample transitions rates. More importantly, state dependence in sector choice has decreased. In other words, the fact that we observe a randomly chosen individual in a given labor market state today is becoming a less and less powerful predictor of the probability that the individual will be found in the same state at a future date.

References

Arulampalam,W., A. Booth, and M. Taylor. 2000. 'Unemployment Persistence,' Oxford Economic Papers, 52: 24-50.

Bigsten, A., P. Kimuyu, and K. Lundvall. 2004. 'What to Do with the Informal Sector?,' Development Policy Review, 22: 701-715.

Bulow, J. and L. Summers. 1986. A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment,' Journal of Labor Economics, 4:376-414.

Chamberlain, G. 'Panel data,' in Z. Griliches and M. Intrilligator (eds.) *Handbook of Econometrics*. Amsterdam: North-Holland.

Dickens, W. and K. Lang. 1985. 'A Test of Dual Labor Market Theory,' American Economic Review, 75: 792-805.

Fields, G. 1975. 'Rural-urban migration, urban unemployment and underemployment, and job search activity in LDCs,' Journal of Development Economics, 2: 165-187.

Gong, X. and A.van Soest. 2002. 'Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico,' Labour Economics 9, 513-529.

Gong, X. A. van Soest, and E. Villagmez. 2004. 'Mobility in the Urban Labor Market: A Panel Analysis for Mexico,' Economic Development and Cultural Change 52, pp.1-36.

Harris J. and M. Todaro. 1970. 'Migration, unemployment and development,' A two sector analysis,' American Economic Review, 60: 126-142.

Heckman, J. 'The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process,' in C. Manski and D. McFadden (eds.) Structural Analysis of Discrete Data in with Econometric Applications, Cambridge MA: MIT Press.

Heckman, J. and V. Hotz. 1986. 'An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality,' Journal of Human Resources, 21: 507-542.

Heller, P. and Tait, A. 1994. `Government Employment and Pay: Some International Comparisons.' IMF Occasional Papers No. 24.

Henley, A. 2004. 'Self-Employment Status: The Role of State Dependence and Initial Circumstances,' Small Business Economics, 22: 67-82.

Hsiao, C. 2003. *The Analysis of Panel Data*. Second edition. Cambridge: Cambridge University Press.

Lindbeck, A. and D. Snower. 1986. 'Wage Setting, Unemployment and Insider-Outsider Relations,' American Economic Review, 76:235-239.

Lindauer, D.L. 1991. `Government Pay and Employment Policy: A Parallel Market in Labor.' in M. Roemer and C. Jones (eds.), Markets in Developing Countries: Parallel, Fragmented and Black. ICS Press: San Francisco.

------ and R. Sabot. 1983. `The Public-Private Sector Wage Differential in a Poor Urban Economy.' Journal of Development Economics 12, 137-153.

Magnac, T. 1991. 'Segmented or competitive labor markets,' Econometrica 59, pp. 165-187.

Maloney, W.F. 1999. 'Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico,' The World Bank Economic Review 13, 275-302.

Nielsen, H. and M. Rosholm. 2001. 'The public-private wage gap in Zambia in the 1990s: A quantile regression approach,' Empirical Economics 26: 169-182.

Pratap, S. and E. Quintin. 2006. 'Are labor markets segmented in developing countries? A semi-parametric approach,' European Economic Review 50, pp. 1817-1841.

Phelps, E. 1972. Inflation Policy and Unemployment Theory: the Cost Benefit Approach to Monetary Planning, Macmillan: London.

Pissarides, C. 1992. 'Loss of skills during unemployment and the persistence of employment shocks,' Quarterly Journal of Economics, 107: 1371-1391.

Rauch, J. 1991. 'Modelling the informal sector formally,' Journal of Development Economics 31: 33-47

Shapiro, C. and J. Stiglitz. 1984. 'Involuntary Unemployment as a Worker Disciplining Device,' American Economic Review, 74:433-444.

Stevenson, G. 1992. `How Public Sector Pay and Employment Affect Labor Markets.' World Bank Policy Research Working Paper No. 944.

Stewart, M. 2005. 'The Inter-related Dynamics of Unemployment and Low Wage Employment,' Warwick Economic Research Papers, No 741, December.

Stiglitz, J. 1974. 'Alternative Theories of Wage Determination and Unemployment in LDCs: the Labor Turnover Model,' Quarterly Journal of Economics, 88: 194-227.

Terrel, K. 1993. `Public-Private Wage Differentials in Haiti: Do Public Servants Earn Rent?' Journal of Development Economics 42: 293-314.

Van der Gaag, J. and W. Vijerberg. 1988. ' A switching regression model for wage determination in the public and private sectors of a developing country,' Review of Economics and Statistics 70: 244-252.

Wooldridge, J. 2000. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge MA: MIT Press.

__________. 2005. 'Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity,' *Journal of Applied Econometrics* 20: 39-54.

World Bank Group. 2007. 'Urban labour markets in Ethiopia: Challenges and Prospects,' Washington DC, February.

Table 1: Ethiopia Urban Socio-economic Survey, distribution of sample labor market state (age group 15-64 only)

	1994		1995		1997		2000		2004	
Labor market state	Number	%	Number	%	Number	%	Number	%	Number	%
Self-employed	609	10.47	636	9.73	530	9.8	556	10.05	574	9.32
Government worker	523	8.99	561	8.58	469	8.67	453	8.19	548	8.9
Public enterprise worker	227	3.9	244	3.73	182	3.37	190	3.43	143	2.32
Formal private sector worker	239	4.11	251	3.84	233	4.31	372	6.72	561	9.11
Other private sector worker	516	8.87	586	8.97	482	8.91	536	9.69	615	9.99
Unemployed	1,075	18.48	1,140	17.44	938	17.35	926	16.73	1,021	16.58
Out of the labor force	2,627	45.17	3,118	47.71	2,573	47.59	2,501	45.19	2,695	43.77
Total	5,816	100	6,536	100	5,407	100	5,534	100	6,157	100

Table 2: Descriptive statistics for selected variables (age group 15-64 only)

	1994	1997	2000	2004	Pooled (all waves)
Earnings and business sales revenue:					
Monthly wages (in log current Birr)					
Mean	7.750	8.076	8.061	8.185	8.007
Standard deviation	1.132	1.043	1.092	0.976	1.074
Annual business revenue (in log current Birr)					
Mean	9.758	9.440	9.562	9.376	9.395
Standard deviation	1.920	1.801	1.740	2.044	1.924
Occupation:					
Professional/technical	0.254	0.171	0.183	0.193	0.193
Other admin/clerical/sales	0.234	0.139	0.218	0.182	0.182
Skilled production	0.148	0.107	0.214	0.155	0.152
Unskilled production	0.333	0.223	0.294	0.292	0.272
Education (classification 1):					
Primary incomplete	0.161	0.161	0.124		0.121
Primary complete	0.390	0.393	0.389		0.304
Secondary complete	0.215	0.186	0.209		0.158
Some tertiary education	0.059	0.051	0.061	0.087	0.062
Education (classification 2):					
Grade 1-10	0.509	0.512	0.513	0.445	0.494
Preparatory	0.257	0.228	0.209	0.250	0.235
Demographics:					
Female	0.554	0.552	0.555	0.551	0.553
Married	0.287	0.252	0.273	0.258	0.265
Age 30-44 yrs.	0.239	0.241	0.225	0.224	0.234
Age 45-54 yrs.	0.100	0.095	0.103	0.098	0.098
Age 55-64 yrs.	0.062	0.068	0.066	0.066	0.066

Table 3. Transtions probabilities across sectors between survey waves (age group 15-64 only)

	New status							
Initial status	1	2	3	4	5	6	7	Total
a. Transition between the 1994 wave and the 1997 wave								
1. self-employed	94.7	0.2	0.4	0.2	0.6	1.3	2.5	100
2. government worker	0.7	97.3	0.0	0.7	0.0	0.5	0.9	100
3. public enterprise worker	1.2	0.0	92.9	1.8	0.0	0.0	4.1	100
4. formal private sector worker	0.0	0.5	1.0	94.4	0.0	1.5	2.5	100
5. other private sector worker	0.8	0.3	0.5	0.0	95.7	2.8	0.0	100
6. unemployed	1.5	0.6	0.3	1.7	2.4	93.2	0.3	100
7. out of the labor force	1.4	0.1	0.1	0.4	0.8	1.8	95.5	100
Total	10.6	9.4	3.6	4.6	8.9	18.8	44.1	100
b. Transition between the 1994 wave and the 2000 wave								
1. self-employed	59.4	2.4	0.3	4.4	9.6	4.4	19.5	100
2. government worker	2.0	58.6	14.9	6.0	0.8	4.8	12.9	100
3. public enterprise worker	8.1	23.4	29.8	7.3	2.4	4.8	24.2	100
4. formal private sector worker	11.1	0.9	7.7	46.2	12.0	8.6	13.7	100
5. other private sector worker	11.4	2.1	2.1	14.9	40.4	14.2	14.9	100
6. unemployed	8.7	5.9	2.5	11.0	11.6	48.7	11.6	100
7. out of the labor force	7.9	3.4	1.9	4.5	5.1	19.0	58.1	100
Total	13.6	9.9	4.7	8.4	8.5	20.0	34.9	100
c. Transition between the 1994 wave and the 2004 wave								
1. self-employed	49.3	3.0	1.5	6.3	8.2	7.8	24.1	100
2. government worker	3.5	45.4	7.0	9.7	3.5	10.5	20.5	100
3. public enterprise worker	3.2	21.3	17.3	10.2	3.2	9.5	35.4	100
4. formal private sector worker	8.9	8.9	4.0	37.6	7.9	14.9	17.8	100
5. other private sector worker	15.7	4.7	1.2	12.2	32.0	9.3	25.0	100
6. unemployed	8.1	9.4	2.4	16.6	10.5	30.4	22.7	100
7. out of the labor force	8.5	6.1	2.0	9.0	6.7	20.4	47.3	100
Total	12.5	11.2	3.3	11.6	8.8	18.3	34.3	100
d. Transition between the 1997 wave and the 2000 wave								
1. self-employed	58.3	1.0	0.7	6.2	9.3	5.5	19.0	100
2. government worker	2.4	59.4	14.9	6.8	0.8	4.4	11.2	100
3. public enterprise worker	3.8	27.4	33.0	9.4	4.7	4.7	17.0	100
4. formal private sector worker	9.8	0.0	8.9	44.7	10.6	13.0	13.0	100
5. other private sector worker	9.9	3.1	1.9	12.4	40.4	15.5	16.8	100
6. unemployed	7.8	5.6	2.0	10.2	11.6	51.3	11.6	100
7. out of the labor force	7.2	3.2	1.2	4.0	4.7	16.9	62.8	100
Total	12.4	9.3	4.2	8.1	8.3	19.5	38.2	100
e. Transition between the 2000 wave and the 2004 wave								
1. self-employed	50.0	2.0	1.3	3.7	6.7	9.3	27.0	100
2. government worker	3.6	55.7	11.3	6.1	2.4	7.3	13.7	100
3. public enterprise worker	2.7	37.3	21.8	10.0	5.5	8.2	14.6	100
4. formal private sector worker	5.0	7.7	3.9	45.9	7.2	13.8	16.6	100
5. other private sector worker	8.7	8.2	1.4	14.9	26.9	16.4	23.6	100
6. unemployed	6.8	8.5	2.9	12.6	10.5	35.7	22.9	100
7. out of the labor force	5.9	4.0	1.1	6.3	4.5	16.9	61.4	100
Total	10.6	11.0	3.4	10.4	7.5	17.9	39.3	100

Table 4: ML estimates of simple probit models of current and initial employment status, 1994-2004

	Dynamic pooled probit of current employment Status, Y_T				**Simple probit of employment status in 1994**			
	Unemployed	Informal Sector worker	Self-Employed	Public Sector worker	Unemployed	Informal Sector worker	Self-employed	Public sector worker
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Y_{t-1}	2.393	2.679	3.026	3.015				
	(48.89)**	(56.48)**	(51.42)**	(49.87)**				
Female	-0.025	0.088	0.121	-0.022	-0.092	0.086	0.208	0.109
	(0.52)	(1.81)	(2.00)*	(0.38)	(1.50)	(1.49)	(3.13)**	(1.62)
Married	-0.459	0.266	0.175	0.266	-0.325	-0.029	-0.456	-0.013
	(3.47)**	(2.20)*	(1.26)	(1.97)*	(1.06)	(0.10)	(1.51)	(0.04)
Age groups (reference group=Age 15-29):								
Age 30-44	-0.199	0.121	-0.073	0.393	-0.250	-0.023	0.207	0.023
	(1.84)	(1.12)	(0.54)	(3.22)**	(1.51)	(0.14)	(1.12)	(0.13)
Age 45-54	-0.184	-0.129	-0.191	0.282	-0.412	-0.145	0.092	0.045
	(0.93)	(0.76)	(0.98)	(1.52)	(1.25)	(0.56)	(0.34)	(0.16)
Age 55-64	0.277	0.015	0.114	-0.088	-0.564	0.275	0.293	-0.350
	(0.88)	(0.06)	(0.45)	(0.31)	(1.25)	(0.88)	(0.95)	(0.98)
Completed schooling (reference group=no formal schooling or primary incomplete):								
Primary	0.386	-0.152	-0.355	-0.307	0.402	-0.302	-0.224	-0.046
	(3.51)**	(1.54)	(2.89)**	(2.32)*	(1.49)	(1.28)	(0.85)	(0.16)
Secondary	0.656	-0.598	-0.204	-0.039	0.692	-0.670	-0.545	0.042
	(5.30)**	(5.02)**	(1.32)	(0.28)	(2.12)*	(2.13)*	(1.38)	(0.12)
Some tertiary	-0.031	-1.138	-1.053	0.535	0.311	-0.699	-0.929	0.351
	(0.15)	(5.47)**	(3.90)**	(2.84)**	(0.60)	(1.29)	(1.37)	(0.72)
Time (or group) means of time varying characteristics								
Married	0.187	-0.141	0.072	-0.266	-0.293	0.107	0.979	0.441
	(1.21)	(1.02)	(0.45)	(1.72)	(0.90)	(0.36)	(3.03)**	(1.40)
Age 30-44	0.023	-0.013	0.363	-0.244	-0.158	0.053	0.328	0.736
	(0.18)	(0.10)	(2.31)*	(1.66)	(0.88)	(0.31)	(1.62)	(3.77)**
Age 45-54	0.101	0.077	0.508	-0.002	-0.088	0.136	0.582	0.786
	(0.44)	(0.39)	(2.25)*	(0.01)	(0.25)	(0.49)	(1.97)*	(2.57)*
Age 55-64	-0.942	0.245	0.365	0.124	-0.302	0.175	0.894	0.859
	(2.34)*	(0.85)	(1.22)	(0.37)	(0.63)	(0.52)	(2.61)**	(2.26)*
Primary school	0.021	-0.239	0.187	0.562	0.466	-0.530	-0.106	0.466
	(0.17)	(2.06)*	(1.31)	(3.65)**	(1.55)	(2.03)*	(0.37)	(1.42)
Secondary	0.064	-0.047	0.031	0.330	1.115	-1.087	-0.189	0.361
	(0.46)	(0.34)	(0.17)	(2.03)*	(3.13)**	(3.15)**	(0.44)	(0.92)
Some tertiary	0.438	0.653	1.082	-0.134	0.808	-0.719	0.800	0.350
	(1.81)	(2.69)**	(3.59)**	(0.58)	(1.44)	(1.23)	(1.11)	(0.66)
Occupation groups								
Professional/technical	-1.602	-1.378	-1.468	0.998	-4.138	-2.185	-4.009	2.923
	(5.86)**	(5.78)**	(4.57)**	(5.83)**	(10.16)**	(6.84)**	(6.85)**	(13.95)**
Admin/clerical	-1.348	-1.072	-1.323	0.593	-3.665	-1.590	-4.226	2.416
	(6.43)**	(5.53)**	(4.52)**	(3.80)**	(12.11)**	(6.06)**	(7.42)**	(12.91)**
Skilled production	-1.124	-0.871	-1.386	0.592	-2.761	-1.073	-2.962	1.717
	(5.31)**	(5.25)**	(4.84)**	(3.76)**	(9.35)**	(5.15)**	(7.74)**	(9.30)**

Table 4 continued:

	Pooled dynamic probit of current employment status, Y_t				Simple probit of employment status in 1994			
	Unemployed	Informal sector employment	Self employment	Public sector employment	Unemployment	Informal sector employment	Self employment	Public sector emploment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Family background variables:								
Fathers' schooling								
Primary					-0.108	0.062	-0.030	-0.036
					(0.83)	(0.48)	(0.20)	(0.25)
Secondary					-0.307	0.200	-0.273	-0.316
					(2.01)*	(1.27)	(1.27)	(1.72)
Some tertiary					-0.081	-0.214	-0.500	0.137
					(0.30)	(0.65)	(1.00)	(0.47)
Mother's schooling (completed:								
Primary					0.017	0.123	-0.065	-0.170
					(0.08)	(0.63)	(0.28)	(0.74)
Secondary completed					-0.478	0.358	0.597	0.367
					(1.16)	(0.89)	(1.31)	(0.83)
Father's occupation (reference group small farmer):								
Non farm business owner/own account worker					-0.203	0.314	0.592	-0.168
					(2.05)*	(3.50)**	(6.16)**	(1.58)
paternaloccupation==public sector worker					-0.012	-0.005	0.087	0.078
					(0.13)	(0.05)	(0.74)	(0.78)
Mother is a homemaker					-0.413	0.331	0.202	0.022
					(6.58)**	(5.81)**	(3.21)**	(0.33)
Ethnicity dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Religion dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-2.016	-0.722	-1.576	-2.374	-0.574	0.730	-0.824	-2.625
	(15.27)**	(5.71)**	(10.91)**	(15.76)**	(3.87)**	(5.32)**	(5.39)**	(15.44)**
Observations	8513	8513	8513	8513	3312	3312	3312	3312
Log likelihood	-1910	-1881	-1239	-1289	-1251	-1479	-1176	-1081
Pseudo R-sq.	0.63	0.66	0.72	0.72	0.40	0.32	0.30	0.39

Absolute value of z statistics in parentheses

* significant at 5%; ** significant at 1%

Table 5: Estimates of dynamic random effects probits of current employment status, 1994-2004

	Wooldridge's estimator				***Two-stage estimator***			
	Current employment status=Y_t				Current employment status=Y_t			
	Unemployed	Informal sector worker	Self-employed	Public sector worker	Unemployed	Informal sector worker	Self-employed	Public sector worker
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Y_{t-1}	1.413	1.576	1.917	1.904	1.635	1.783	2.356	1.906
	(12.88)**	(15.05)**	(14.20)**	(13.44)**	(15.80)**	(18.00)**	(19.36)**	(13.94)**
Y_{1994}	1.621	1.677	1.582	1.642				
	(14.13)**	(15.63)**	(11.41)**	(11.26)**				
Generalized residual of initial status probit					0.783	0.834	0.628	0.906
					(12.70)**	(14.24)**	(8.75)**	(11.64)**
Female	-0.048	0.094	0.132	0.002	-0.085	0.122	0.175	0.062
	(0.87)	(1.63)	(1.88)	(0.03)	(1.53)	(2.13)*	(2.50)*	(0.90)
Married	-0.596	0.276	0.147	0.111	-0.597	0.280	0.111	0.109
	(3.62)**	(1.82)	(0.87)	(0.67)	(3.64)**	(1.85)	(0.66)	(0.65)
Age groups (reference group=Age 15-29):								
Age 30-44	-0.313	0.115	-0.078	0.636	-0.302	0.110	-0.107	0.617
	(2.45)*	(0.89)	(0.49)	(4.33)**	(2.38)*	(0.86)	(0.67)	(4.21)**
Age 45-54	-0.416	-0.107	-0.169	0.625	-0.444	-0.109	-0.203	0.607
	(1.79)	(0.52)	(0.75)	(2.79)**	(1.89)	(0.53)	(0.91)	(2.71)**
Age 55-64	-0.005	-0.020	0.074	0.325	-0.022	0.026	0.074	0.274
	(0.01)	(0.07)	(0.25)	(0.94)	(0.06)	(0.09)	(0.26)	(0.80)
Completed schooling (reference group=no formal education)								
Primary	0.075	-0.034	-0.410	-0.166	0.062	-0.046	-0.438	-0.173
	(0.52)	(0.27)	(2.71)**	(0.99)	(0.44)	(0.37)	(2.94)**	(1.03)
Secondary	0.603	-0.568	-0.251	-0.190	0.578	-0.561	-0.246	-0.213
	(3.86)**	(3.86)**	(1.33)	(1.08)	(3.71)**	(3.83)**	(1.30)	(1.20)
Some tertiary	0.018	-0.940	-0.849	0.489	-0.027	-0.977	-0.918	0.478
	(0.07)	(3.88)**	(2.85)**	(2.12)*	(0.11)	(3.95)**	(3.02)**	(2.07)*
Time (or group) means of time varying characteristics								
Married	0.472	-0.227	0.022	-0.149	0.196	-0.153	0.241	0.037
	(2.50)*	(1.33)	(0.12)	(0.79)	(1.04)	(0.90)	(1.28)	(0.19)
Age 30-44	0.185	-0.102	0.209	-0.514	-0.061	-0.032	0.414	-0.187
	(1.21)	(0.68)	(1.14)	(2.92)**	(0.41)	(0.21)	(2.25)*	(1.06)
Age 45-54	0.514	-0.127	0.353	-0.367	0.340	-0.118	0.600	0.008
	(1.93)	(0.54)	(1.36)	(1.38)	(1.26)	(0.50)	(2.34)*	(0.03)
Age 55-64	-0.566	0.199	0.299	-0.312	-0.957	0.354	0.663	0.025
	(1.22)	(0.57)	(0.85)	(0.76)	(1.98)*	(1.02)	(1.94)	(0.06)
Primary school	0.144	-0.238	0.435	0.404	0.613	-0.726	0.266	0.612
	(0.87)	(1.62)	(2.47)*	(2.04)*	(3.68)**	(4.94)**	(1.54)	(3.06)**
Secondary	-0.115	0.027	0.161	0.554	0.759	-0.858	-0.125	0.742
	(0.65)	(0.16)	(0.76)	(2.70)**	(4.10)**	(4.94)**	(0.58)	(3.57)**
Some tertiary	0.063	0.824	1.122	-0.096	0.677	0.172	1.027	0.217
	(0.22)	(2.93)**	(3.32)**	(0.34)	(2.34)*	(0.60)	(3.00)**	(0.77)
Occupation groups								
Professional/technical	-0.975	-1.578	-1.394	0.843	-2.352	-2.344	-2.043	2.257
	(3.32)**	(5.31)**	(3.66)**	(4.03)**	(6.91)**	(7.14)**	(4.96)**	(9.36)**

Table 5 continued:

	Wooldridge's estimator				***Two-stage estimator***			
	Current employment status=Y_t				Current employment status=Y_t			
	Unemployed	Informal sector worker	Self-employed	Public sector worker	Unemployed	Informal sector worker	Self-employed	Public sector worker
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Admin/clerical	-0.685	-1.049	-0.972	0.261	-2.002	-1.732	-1.723	1.451
	(2.96)**	(4.51)**	(3.10)**	(1.38)	(7.69)**	(6.76)**	(4.91)**	(7.15)**
Skilled production	-1.235	-0.652	-1.246	0.382	-2.388	-1.149	-2.051	1.250
	(4.62)**	(3.39)**	(3.76)**	(1.99)*	(8.12)**	(5.81)**	(5.71)**	(6.31)**
Mother' tongue:								
Amharic	0.037	0.043	0.036	-0.038	0.081	-0.070	-0.103	0.068
	(0.50)	(0.59)	(0.43)	(0.43)	(1.11)	(0.97)	(1.23)	(0.75)
Oromo	0.133	-0.215	-0.294	0.020	0.192	-0.305	-0.389	0.048
	(1.56)	(2.52)*	(2.87)**	(0.19)	(2.25)*	(3.61)**	(3.84)**	(0.46)
Tigrawai	0.028	0.149	0.030	-0.113	0.110	0.041	-0.073	-0.143
	(0.22)	(1.14)	(0.20)	(0.71)	(0.84)	(0.32)	(0.49)	(0.88)
Religion:								
Orthodox Christian	0.000	-0.227	-0.256	0.183	0.081	-0.321	-0.293	0.206
	(0.00)	(2.18)*	(2.15)*	(1.50)	(0.77)	(3.07)**	(2.48)*	(1.69)
Muslim	-0.080	0.001	-0.148	-0.007	0.007	-0.026	-0.120	-0.129
	(0.60)	(0.01)	(1.02)	(0.04)	(0.05)	(0.20)	(0.84)	(0.75)
Town of residence (reference=Addis):								
Awasa	-0.247	-0.053	-0.147	0.251	-0.237	-0.115	-0.101	0.409
	(1.93)	(0.40)	(0.98)	(1.75)	(1.86)	(0.88)	(0.68)	(2.81)**
Bahir Dar	-0.100	-0.149	-0.034	0.293	-0.210	-0.162	-0.019	0.538
	(0.67)	(1.00)	(0.20)	(1.86)	(1.37)	(1.10)	(0.12)	(3.46)**
Dessie	-0.080	-0.133	-0.065	-0.146	-0.168	-0.074	0.007	-0.133
	(0.58)	(0.97)	(0.42)	(0.81)	(1.24)	(0.55)	(0.05)	(0.72)
Dire Dawa	0.282	-0.179	-0.178	-0.257	0.252	-0.275	-0.168	-0.084
	(2.72)**	(1.66)	(1.36)	(1.89)	(2.44)*	(2.56)*	(1.31)	(0.62)
Jimma	-0.151	0.068	-0.059	0.232	-0.252	0.011	-0.076	0.432
	(1.22)	(0.61)	(0.46)	(1.75)	(2.03)*	(0.10)	(0.61)	(3.23)**
Mekele	-0.191	0.079	0.180	0.125	-0.309	0.211	0.227	0.185
	(1.00)	(0.43)	(0.94)	(0.56)	(1.61)	(1.17)	(1.21)	(0.81)
Year of observation (reference=2000):								
1995	-0.096	0.062	0.095	0.038	-0.089	0.057	0.092	0.037
	(1.49)	(0.93)	(1.21)	(0.48)	(1.40)	(0.87)	(1.19)	(0.46)
1997	-0.586	0.422	0.296	0.064	-0.553	0.407	0.292	0.055
	(7.67)**	(5.68)**	(3.38)**	(0.72)	(7.29)**	(5.48)**	(3.37)**	(0.61)
2004	-0.523	0.279	0.407	0.180	-0.464	0.259	0.386	0.171
	(5.71)**	(3.16)**	(4.04)**	(1.80)	(5.12)**	(2.94)**	(3.82)**	(1.71)
Constant	-1.782	-1.269	-1.978	-2.673	-1.447	-0.092	-1.563	-3.019
	(12.22)**	(8.76)**	(12.25)**	(15.18)**	(9.82)**	(0.58)	(9.73)**	(16.41)**
Observations	7796	7796	7796	7796	7796	7796	7796	7796
Individuals	1949	1949	1949	1949	1949	1949	1949	1949
Log likelihood	-1335	-1309	-915	-884	-1358	-1329	-940	-879
estimate of σ_a	$9.2*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$
[s.e.(σ_a)]	0.045	0.056	0.071	0.072	0.040	0.048	0.053	0.071

Absolute value of z statistics in parentheses

* Significant at 5%; ** significant at 1%

Table 6: Estimates of dynamic random effects probits of current employment status, Wooldridge's estimator, 1997-2004 and 1994-2000

	Current employment status=Y_{it}							
	1997-2004				1994-2000			
	Unemployed	Informal sector worker	Self-employed	Public sector worker	Unemployed	Informal sector worker	Self-employed	Public sector worker
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Y_{t-1}	1.092	1.277	1.637	1.618	1.383	1.549	2.039	1.873
	(9.92)**	(12.08)**	(11.93)**	(11.29)**	(6.86)**	(7.46)**	(6.09)**	(5.78)**
Y_{1994}	1.202	1.340	1.256	1.357	1.925	1.906	1.632	1.945
	(10.38)**	(12.29)**	(8.86)**	(9.16)**	(9.39)**	(9.13)**	(4.85)**	(5.95)**
Female	-0.060	0.085	0.134	0.023	-0.074	0.098	0.161	0.050
	(0.97)	(1.34)	(1.75)	(0.30)	(1.17)	(1.51)	(2.04)*	(0.61)
Married	-0.531	0.240	0.125	0.110	-0.515	0.418	0.249	0.008
	(3.24)**	(1.60)	(0.76)	(0.67)	(2.40)*	(2.11)*	(1.11)	(0.03)
Age 30-44	-0.257	0.067	-0.090	0.558	-0.135	-0.045	-0.218	0.667
	(1.96)*	(0.51)	(0.56)	(3.76)**	(0.81)	(0.26)	(1.02)	(3.36)**
Age 45-54	-0.373	-0.165	-0.175	0.613	-0.277	0.004	-0.121	0.747
	(1.59)	(0.80)	(0.77)	(2.73)**	(0.88)	(0.02)	(0.40)	(2.44)*
Age 55-64	-0.003	-0.037	0.020	0.331	-0.098	0.067	0.101	0.759
	(0.01)	(0.13)	(0.07)	(0.94)	(0.20)	(0.18)	(0.25)	(1.70)
Completed schooling (reference group=no formal education)								
Primary	0.071	-0.026	-0.339	-0.143	-0.312	0.053	-0.360	0.144
	(0.50)	(0.21)	(2.29)*	(0.88)	(1.77)	(0.35)	(1.85)	(0.64)
Secondary	0.512	-0.485	-0.183	-0.150	0.487	-0.733	-0.291	-0.210
	(3.27)**	(3.31)**	(1.00)	(0.87)	(2.48)*	(3.88)**	(1.15)	(0.85)
Some tertiary	-0.037	-0.833	-0.719	0.465	0.547	-0.839	-1.078	0.108
	(0.15)	(3.47)**	(2.46)*	(2.05)*	(1.74)	(2.75)**	(2.84)**	(0.33)
Ethinicity/ religion dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time (or group) means of time varying characteristics								
Marital status	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age groups	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Schooling	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1.497	-0.855	-1.587	-2.497	-1.957	-1.435	-2.149	-2.659
	(9.25)**	(5.41)**	(9.13)**	(13.12)**	(11.74)**	(8.75)**	(11.84)**	(12.92)**
Observations	4484	4484	4484	4484	7125	7125	7125	7125
Individuals)	2242	2242	2242	2242	2375	2375	2375	2375
Log likelihood	-1132	-1130	-799	-775	-1005	-1003	-698	-609
estimate of σ_a	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$	$9.1*10^{-4}$
[s.e.(σ_a)]	0.052	0.070	0.087	0.089	0.063	0.065	0.083	0.095

Absolute value of z statistics in parentheses

* significant at 5%; ** significant at 1%

Table 6 : Maximum Likelihood Estimation of Probit Model of Unemployment

Age group 15-64 only

	Initial selection equation by year of observation		Dynamic selection equation by year of observation			
	1994	2000	1997	2000	2004	2004
Unemployed$_{1994}$			3.195	1.345	0.672	
			(8.53)**	(5.14)**	(2.01)*	
Unemployed$_{2000}$						1.349
						(4.06)**
Genresidualunemployed$_{1994}$			0.112	-0.029	-0.039	
			(0.50)	(0.18)	(0.17)	
Genresidualunemployed$_{2000}$						-0.429
						(2.13)*
Female	-0.039	-0.071	0.106	-0.207	-0.022	0.125
	(0.73)	(1.33)	(0.78)	(2.13)*	(0.18)	(1.37)
Married$_{1994}$	-0.761		-0.338	-0.282	-0.381	
	(10.47)**		(1.61)	(1.88)	(2.06)*	
Married$_{2000}$		-0.650				-0.380
		(8.86)**				(2.54)*
Age groups (reference group=Age 15-19):						
Age 30-44	-0.599	-0.547	-0.180	-0.238	-0.502	-0.354
	(9.50)**	(8.71)**	(1.00)	(2.14)*	(3.75)**	(3.46)**
Age 45-54	-0.759	-0.678	0.262	-0.276	-0.737	-0.729
	(6.64)**	(6.08)**	(1.00)	(1.52)	(3.37)**	(4.01)**
Age 55-64	-0.915	-0.837	-0.000	-0.414	-0.728	-0.828
	(4.87)**	(4.53)**	(0.00)	(1.44)	(2.23)*	(2.56)*
Education (reference group=no formal schooling or primary incomplete):						
Primary school completed	0.709	-0.265	0.062	0.138	0.048	-0.358
	(9.60)**	(2.18)*	(0.28)	(0.94)	(0.25)	(1.61)
Secondary school completed	1.187	-0.246	0.572	0.120	0.118	-0.137
	(16.60)**	(1.93)	(2.26)*	(0.70)	(0.54)	(0.46)
Some tertiary education	0.123	0.145	0.318	-0.842	0.037	-1.014
	(1.04)	(1.03)	(1.03)	(2.90)**	(0.14)	(2.88)**
City dummies	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity and religion dummies	Yes	Yes	Yes	Yes	Yes	Yes
Paternal education dummies	Yes	Yes	No	No	No	No
Paternal occupation dummies	Yes	Yes	No	No	No	No
Constant	-0.692	-0.577	-2.304	-1.197	-0.418	-0.980
	(5.21)**	(4.63)**	(6.10)**	(4.60)**	(1.30)	(3.74)**
Observations	3312	3138	1430	1228	726	1253
Log likelihood	-1540	-1561	-218	-474	-301	-536
Pseudo R-squared	0.26	0.18	0.76	0.29	0.18	0.21

Absolute value of z statistics in parentheses

* significant at 5%; ** significant at 1%

Table 7 : Earnings and Selection into the Private and Public Sectors of Wage Employment: 2004

	(1)	(2)	(3)	(4)
	Public sector log monthly earnings in 2004	Private sector log monthly earnings in 2004	Log monthly earnings in 2004	Selection into public sector in 2004 (Probit)
Female	-0.214	-0.388	-0.302	-1.891
	(3.22)**	(4.49)**	(5.60)**	(7.21)**
Married				-0.020
				(0.12)
Log age	5.361	4.085	3.961	-9.877
	(1.66)	(1.89)	(2.34)*	(1.80)
Log age squared	-0.678	-0.485	-0.468	1.621
	(1.50)	(1.58)	(1.95)	(2.08)*
Education (reference=no formal schooling):				
Grades 1 to 10	0.338	0.157	0.254	-1.746
	(2.42)*	(1.21)	(2.74)**	(5.66)**
Preparatory school	0.687	0.281	0.457	-4.095
	(4.70)**	(2.01)*	(4.61)**	(7.49)**
Some tertiary education	1.077	0.769	0.886	-3.021
	(7.07)**	(4.31)**	(7.93)**	(6.38)**
Own occupation (reference=unskilled workers):				
Professional or technical	0.251	0.822	0.466	6.654
	(2.37)*	(4.95)**	(5.27)**	(10.39)**
Admin or clerical	0.207	0.334	0.286	0.967
	(1.86)	(2.63)**	(3.52)**	(5.15)**
Skilled production worker	0.292	0.482	0.426	1.777
	(2.72)**	(4.29)**	(5.55)**	(7.82)**
Public_{2004}			0.492	
			(4.64)**	
public sector wage premium_{2004}				12.174
				(9.72)**
$\text{Genresidualpublic}_{2004}$			-0.168	
			(2.44)*	
Selectivity term 1	-0.180			
	(2.27)*			
Selectivity term 2		0.088		
		(0.62)		
Constant	-4.669	-2.532	-2.630	18.126
	(0.81)	(0.68)	(0.89)	(1.85)
City dummies	Yes	Yes	Yes	Yes
Ethnicity dummies	Yes	Yes	Yes	Yes
Religion dummies	Yes	Yes	Yes	Yes
Observations	291	312	603	603
R-squared	0.49	0.45	0.51	
Log likelihood				-286
Pseudo R-squared				0.31

Absolute value of z statistics in parentheses

* significant at 5%; ** significant at 1%

Table 8: Earnings and Selection into the Private and Public Sectors of Wage Employment: 20000

	(1)	(2)	(3)	(4)
	Public sector log earnings in 2000	Private sector log monthly earnings in 2000	Log monthly earnings in 2000	Selection into public sector in 2000 (Probit)
Female	-0.180	-0.465	-0.321	-1.745
	(2.21)*	(4.44)**	(4.78)**	(2.01)*
Married				0.200
				(1.38)
Log age	-0.978	-0.023	-0.194	2.798
	(1.46)	(0.08)	(0.73)	(0.95)
Log age squared	7.745	0.497	1.872	-22.031
	(1.60)	(0.23)	(0.98)	(0.99)
Education (reference=no formal schooling):				
Grades 1 to 10	-0.304	0.217	0.047	4.217
	(2.22)*	(1.67)	(0.49)	(2.68)**
Preparatory school	0.089	0.411	0.343	2.952
	(0.64)	(2.65)**	(3.21)**	(2.96)**
Some tertiary education	0.350	1.344	0.763	8.257
	(2.35)*	(6.27)**	(6.11)**	(2.74)**
Own occupation				
(reference=unskilled worker):	0.755	0.862	0.842	1.572
Professional or technical	(5.90)**	(4.03)**	(7.46)**	(4.13)**
	0.334	0.601	0.461	2.564
Admin or clerical	(2.83)**	(3.76)**	(4.68)**	(3.15)**
	0.319	0.499	0.434	1.249
Skilled production worker	(2.57)*	(4.56)**	(5.20)**	(2.29)*
Public$_{2000}$			0.401	
			(3.76)**	
public sector wage premium$_{2000}$				7.618
				(2.54)*
Genresidualpublic$_{2000}$			-0.135	
			(1.91)	
Selectivity term 1	-0.070			
	(0.83)			
Selectivity term 2		0.072		
		(0.55)		
Constant	-9.554	3.470	0.877	38.011
	(1.09)	(0.92)	(0.26)	(0.94)
city dummies	Yes	Yes	Yes	Yes
ethnicity dummies	Yes	Yes	Yes	Yes
religion dummies	Yes	Yes	Yes	Yes
Observations	291	289	580	580
R-squared	0.46	0.51	0.47	
Log likelihood				-297
Pseudo R-squared				0.26

Absolute value of z statistics in parentheses

* significant at 5%; ** significant at 1%

Table 9 : Earnings and Selection into the Private and Public Sectors of Wage Employment: 1997

	(3)	(4)	(5)	(6)
	Public sector log monthly earnings in 1997	Private sector log monthly earnings in 1997	Log monthly earnings in 1997	Selection into public sector in 1997 (Probit)
Female	-0.245	-0.508	-0.375	0.492
	(3.50)**	(4.31)**	(5.82)**	(2.49)*
Married				0.367
				(2.96)**
Log age	6.035	8.117	8.072	16.900
	(2.04)*	(2.89)**	(4.12)**	(4.39)**
Log age squared	-0.767	-1.089	-1.069	-2.251
	(1.85)	(2.73)**	(3.86)**	(4.12)**
Education (reference=no formal schooling):				
Grades 1 to 10	-0.033	0.141	0.096	0.402
	(0.25)	(0.88)	(0.94)	(1.95)
Preparatory school	0.322	0.597	0.518	0.564
	(2.22)*	(3.12)**	(4.40)**	(2.11)*
Some tertiary education	0.585	1.084	0.782	1.024
	(3.89)**	(3.77)**	(5.77)**	(2.69)**
Own occupation (reference=unskilled workers):				
Professional or technical	0.530	0.799	0.584	0.519
	(5.27)**	(3.69)**	(5.91)**	(2.07)*
Admin or clerical	0.384	0.439	0.430	0.088
	(3.72)**	(2.58)*	(4.55)**	(0.53)
Skilled production worker	0.322	0.461	0.409	-0.130
	(2.98)**	(3.16)**	(4.56)**	(0.73)
Public_{1997}			0.320	
			(4.46)**	
$\text{Genresidualpublic}_{1997}$			-0.228	
			(2.41)*	
$\text{public sector wage premium}_{1997}$				-0.268
				(0.44)
Selectivity term 1	-0.200			
	(2.09)*			
Selectivity term 2		0.164		
		(0.80)		
Constant	-6.299	-10.006		-32.604
	(1.20)	(2.05)*		(4.85)**
city dummies	Yes	Yes	Yes	Yes
ethnicity dummies	Yes	Yes	Yes	Yes
religion dummies	Yes	Yes	Yes	Yes
Observations	402	328	730	730
R-squared	0.35	0.37	0.39	
Log likelihood				-402
Pseudo R-squared				0.2

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

Table 10: Earnings and Selection into the Wage Employment and Self Employment: 1997, 2000, and 2004

	Log annual sales in year			Log annual wages in year			Structural (simple) probit of selection into self-employment in year		
	1997	2000	2004	1997	2000	2004	1997	2000	2004
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	0.748	0.061	-0.129	-0.354	-0.294	-0.184	-4.009	-0.425	-0.009
	(2.37)*	(0.12)	(0.22)	(5.33)**	(3.66)**	(1.87)	(12.11)**	(2.91)**	(0.05)
Married							-0.038	-0.074	-0.146
							(0.17)	(0.49)	(0.83)
Log age	0.939	-7.286	-20.522	9.818	3.017	9.216	33.008	13.515	76.826
	(0.12)	(0.68)	(1.29)	(4.73)**	(1.32)	(3.27)*	(7.33)**	(3.64)**	(12.44)**
Log age squared	-0.023	1.235	2.781	-1.291	-0.326	-1.171	-4.689	-1.993	-10.063
	(0.02)	(0.87)	(1.30)	(4.41)**	(1.02)	(2.96)*	(7.28)**	(3.89)**	(12.15)**
Education (reference=no formal schooling):									
Grades 1 to 10	0.398	0.281	1.412	0.187	-0.001	0.319	-0.818	-0.338	-2.853
	(1.36)	(0.68)	(1.93)	(1.78)	(0.01)	(2.28)*	(2.89)**	(2.22)*	(10.07)**
Preparatory school	1.722	1.560	1.871	0.890	0.575	0.508	-3.078	-1.200	-3.440
	(3.71)**	(2.46)*	(2.21)*	(7.97)**	(4.48)**	(3.14)*	(8.65)**	(5.90)**	(10.01)**
Some tertiary education	1.652	2.613	2.906	1.254	1.186	0.897	-1.633	-1.800	-5.526
	(2.40)*	(3.28)**	(2.58)*	(10.52)**	(8.70)**	(4.15)*	(3.50)**	(7.48)**	(10.71)**
Self-employment earnings permium							3.601	1.271	2.734
							(19.44)**	(14.64)**	(13.38)**
Selectivity term 1				-0.390	-0.569	0.060			
				(3.17)**	(4.34)**	(0.46)			
Selectivity term 2	-0.111	-0.692	-0.630						
	(0.37)	(1.66)	(1.33)						
Constant	5.092	18.032	47.048	-10.699	1.703	-9.869	-60.109	-21.993	-147.524
	(0.36)	(0.90)	(1.60)	(2.93)**	(0.42)	(1.99)*	(7.72)**	(3.28)**	(12.71)**
city dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ethnicity dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Father's education	No	No	No	No	No	No	No	No	No
Mother's education	No	No	No	No	No	No	No	No	No
Father's occupation	No	No	No	No	No	No	No	No	No
Observations	177	99	115	628	452	212	1748	750	507
R-squared	0.14	0.29	0.19	0.36	0.37	0.30			
Log likelihood							-80	-259	-183
Pseudo R-squared							0.93	0.45	0.47

Absolute value of z statistics in parentheses

* significant at 5%; ** significant at 1%