ISSN 1415-4765

TEXTO PARA DISCUSSÃO Nº 1021

DO INFORMAL WORKERS QUEUE FOR FORMAL JOBS IN BRAZIL?

Fábio Veras Soares

Brasília, maio de 2004

ISSN 1415-4765

TEXTO PARA DISCUSSÃO Nº 1021

DO INFORMAL WORKERS QUEUE FOR FORMAL JOBS IN BRAZIL?*

Fábio Veras Soares**

Brasília, maio de 2004

* I would like to acknowledge Capes (Brasília, Brasil) for my scholarship.
O presente texto, em língua Inglesa, não foi objeto de revisão editorial.
** Técnico de Planejamento e Pesquisa do Ipea.

Governo Federal

Ministério do Planejamento, Orçamento e Gestão

Ministro – Guido Mantega
Secretário-Executivo – Nelson Machado

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada ao Ministério do Planejamento, Orçamento e Gestão, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

TEXTO PARA DISCUSSÃO

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

SUMÁRIO

SINOPSE

Este artigo investiga a existência de fila por trabalhos formais (com carteira) no mercado de trabalho brasileiro. Estima-se uma regressão com mudança endógena (*endogenous switching regression*) a fim de especificar corretamente o processo de alocação setorial dos trabalhadores na presença de filas, e desse modo, conseguir estimar o diferencial de salários entre trabalhadores do setor formal e do informal livre de viés de seletividade. Foram avaliados três diferentes especificações de probit bivariados com o intuito de avaliar a sensibilidade das estimativas das filas a diferentes hipóteses sobre o processo de alocação setorial.

Em particular, testou-se a sensibilidade das estimativas das filas por trabalho à hipótese de observação parcial dos modelos de fila por trabalho usando perguntas subjetivas de pesquisas (Pnad 1990) sobre o desejo dos trabalhadores do setor informal de mudar para um emprego formal. Os testes não rejeitam a hipótese de existência de fila por trabalho formal. Entre os trabalhadores com maiores dificuldades de serem escolhidos da fila, uma vez que estejam nela, estão os negros, as mulheres, os analfabetos, os jovens que estão entrando no mercado de trabalho e os trabalhadores que foram informais no último trabalho. Este último resultado sugere que um período no setor informal reduz bastante a chance de um trabalhador conseguir um posto de trabalho formal no futuro.

ABSTRACT

This paper investigates the existence of a job queue for formal (registered) jobs in the Brazilian labour market in an endogenous switching regression framework. This approach aims at correctly specifying the allocation process in the presence of queuing and getting unbiased wage equation estimates in order evaluate the role of wage differential between formal and informal sector in determining sector allocation. We estimate three types of bivariate probit specifications in order to evaluate the sensitivity of the results to different assumptions about the sector allocation process. In particular, we assess the sensitivity of the job queue estimates to the usual assumption of partial observability using subjective survey questions on the desire of informal (non-registered) workers to switch to formal job.

Our tests were not able to reject the hypothesis of job queue. Our estimates of the job queue "length" for selected groups show that non-white, female, illiterate, "new entrants" and former informal workers are the groups with the lower probability of being chosen from the queue conditional on being in the queue. This result is particularly strong for workers whose last job was in the informal sector, suggesting that a spell in the informal sector may jeopardize the worker's chance of getting a formal job.

1 INTRODUCTION

It is a well-established fact that there is a large wage differential between formal and informal sector workers in Brazil.[1] This differential may be even larger if one takes into account the total compensation package of a formal job contract. This package includes mandatory annual bonus (usually one extra monthly wage per year), paid holidays and access to public funds such as unemployment benefit and severance payment. However the wage gap between (observably) similar workers allocated into these two sectors is not enough to prove the existence of segmentation in the labour market.[2]

Maloney (1997) argues that the wage differential is not a good guide to assess the existence of segmentation and claims that tests on the existence of a job queue is a much clearer indicator of segmentation. If some workers would prefer to get a formal job, but do not get it and at the same time similar workers do get it, then one could claim the existence of some sort of segmentation. This evidence could be an indication of rigidity in the labour market that can be caused by the excess of regulation or simply by efficiency wage practices or a combination of both of them. The existence of a job queue for formal jobs, therefore, has implications for both economic efficiency and long-run income inequality. As for the first aspect, law reform turning the labour market more flexible has been listed as one of the important reforms that must be implemented by developing and developed countries that face slow growth and increasing unemployment or under-employment.[3] The existence of job queue can be considered an indicator of this lack of flexibility. As for income inequality, the major problem is that formal jobs come together with several benefits, including pensions, that are not readily available for informal sector workers. The longer the individual spends in the queue for formal jobs, the more likely is that he/she will lose these benefits or get a smaller proportion of them; perpetuating and increasing inequality. Differently, it has been argued that the possibility of tax evasion and more flexible hours were factors that would compensate for the lack of fringe benefits and access to social funds associated to the formal sector jobs (Maloney, 1997). However, if this were the case, then one should observe a high proportion of informal sector workers satisfied with their current occupational status. According to the 1990 Pnad, this can be considered the case for self-employed workers, but hardly can be considered consistent with the view that non-registered wage workers have of their own current status since 70% of them would like to switch to a formal job.

Most of the studies on labour market segmentation in developing countries have focused on the need to correct the wage equations for both formal and informal

1. Throughout this paper we use informal sector as the set of workers without registration (a signed work card).

2. One can argue that his/her definition of segmentation consists only of the existence of wage differential between otherwise similar workers employed in different sectors, but the existence of wage differential between similar workers is also compatible with other competitive explanations such as compensation differential, comparative advantages or returns to entrepreneurial ability (in the case of self-employment).

3. See Heckman and Pagés (2004) for a defence of labour market reform as a way to improve job creation in Latin American and Caribbean countries.

sector in order to get unbiased estimates of the wage differential between the two sectors. This procedure, however, does not limit itself to finding the unbiased wage differential between the formal and informal sectors, i.e., the differences in the wages that any individual drawn from the population would get in either sector. It also allows one to test whether or not individual select themselves into these two sectors according to their comparative advantages. The idea that individuals self-select into the occupation/sector where they are more productive dates back to Roy's (1951) model.[4] Roy shows how individuals choose to work in the occupation in which they are more productive and how the flow and the stock of workers in each occupation varies with demand shocks in the product of their industry. One of the key assumptions of the Roy model is that there are no barriers to entry in any of the sectors. Thus, there is no job queue in such framework.

It is interesting to note that despite all the controversy around the existence of rationing in the formal sector and the discussions on how to correct for selectivity bias, the literature on job queue has been scarcely applied to tackle this issue.[5] This literature, however, has been widely applied in the study of union/non-union wage differential, private and public wage differential, and queue for minimum wage jobs.[6]

According to the job queue approach the process of sector allocation cannot be correctly modelled by a univariate process that determines whether the individual prefer to work in one sector or other based on his/her comparative advantage and/or preferences. If jobs are rationed, one must take into account the employers' criteria to choose workers for rationed jobs. Individual characteristics and past employment history, for instance, may have different effects on the probability of desiring a formal job (in the queue status) and on the probability of being chosen (from the queue) by the employer for that job. In addition, the existence of a job queue for formal jobs would lead to a biased estimation of the wage equation, since the univariate specification of the selection equation would misrepresent the process of sector allocation (Farber, 1983). Abowd and Farber (1982) and Farber (1983) were the first to apply this idea to the allocation of workers into union/non-union sectors, but they did not go further to correct the wage equation using the double selectivity criteria. Mengistae (1998), following Maddala's (1983) suggestion, puts forward a methodology that not only incorporates the double selectivity criteria in the wage equation, but also enables the estimation of the proportion of workers queuing directly from the corrected wage equation for workers who were not chosen from the queue. The main weakness of this approach consists in the fact that, in general, one has to rely on the estimation of bivariate probit with severe partial observability, since the "in the queue" status is hardly observed. In such a situation the identification of the two selection equations relies heavily on nonlinearities in the functional form of the probability distribution (Farber, 1983). More worryingly, procedures based on Poirier's (1980) bivariate probit with partial observability are known to have a very

4. The major difference between Roy's model and other models of comparative advantage in the choice of sector/occupation is that in Roy's model only the incomes are compared in the decision rule, whereas in more general models of utility maximization, non-wage dimensions also play a role in the sector choice, see Heckman and Honore (1990).

5. We are only aware of Maloney (1997) attempt to test the existence of job queue for formal jobs in Mexico in a time series context.

6. See Abowd and Farber (1982) and Farber (1983) for union/non-union, Venti (1987), Heywood and Mohanty (1995), Mohanty (1998) and Mengistae (1998) for private/public sector and Holzer, Katz and Krueger (1991) for minimum wage applications.

bad performance in terms of convergence.[7] Thus, attempts to add more information on the severe partial observability of the "in the queue" status may help to identify the two equations.[8] We will use a special supplement of 1990 Brazilian Annual Household survey (Pnad) to identify workers in the informal sector who would be queuing for formal jobs in order to lessen the severe partial observability of Abowd and Faber (1982) and Poirier (1980) bivariate probit models.

This paper is structured as follows. First, we test the existence of a job queue for formal jobs and how its "length" varies for different groups of workers. This is important for public policy to target more vulnerable workers if one assumes that an informal sector job is a "second best" option. Second, we estimate selectivity-corrected wage equations for formal and informal workers in order to assess the role of the wage differential in determining whether or not a worker join the queue for a formal job (endogenous switching regression). Finally, we investigate the sensitivity of job queue estimates to different assumptions regarding the sector allocation process. Using subjective survey questions on whether or not an informal worker would accept a formal job we are able to estimate a bivariate probit with sample selection[9] and partially overcome the difficulties imposed by the severe partial observability on both Abowd and Faber (1982) and Poirier (1980) models.

2 ENDOGENOUS SWITCHING REGRESSION IN THE PRESENCE OF JOB QUEUE

Most endogenous switching regression models are based on a univariate probit (or logit) process that models the worker's preference for a specific type of job. The assumption behind this approach is that once a worker decides to get a job in a specific sector there is nothing preventing him/her to get hired. Therefore the probability of desiring a formal job is identical to the probability of having a job in that sector. Besides it is assumed that workers base their decision taking into account comparative advantages they would have in the chosen sector. Thus, the wage differential between what a worker would get in a sector *vis-à-vis* what he/she would get in the other sector should play a substantial role in determining the actual sector allocation. The common procedure in this case is to estimate in a first step a reduced form probit (or logit), then correct the wage equations for selectivity bias (due to non-random selection into the sector), get the wage differential based on the corrected wage equations and, finally, estimate a structural probit that incorporates the wage differential as an additional regressor.

To assess the existence of a job queue in an endogenous switching regression model we will relax the assumption that the probability of getting a formal job is equal to the probability of willing to be in the formal sector. The probability of being in the formal sector is the result of two independent decision processes: the decision to join the queue for formal jobs by the worker and the decision to hire a worker who

7. It is common to find papers that exclude some variables of the estimation process or redefine them only to guarantee convergence. See, for instance, Mohanty (1998).
8. Farber (1983) uses the intention to vote in favour of unionisation in order to help the identification of the workers who would be in the queue for union jobs.
9. This sort of bivariate probit is also known as Meng and Schmidt bivariate probit or censored bivariate probit.

is in the queue by the formal employers. A worker is only observed in the formal sector if he/she both had joined the queue and was chosen from the queue.

A worker decides to join the queue if the utility that he/she derives from this choice is higher than a specific threshold. In the case of the formal sector, he will choose this sector if the "advantages" related to a formal job (e.g. pensions, paid vacations) more than compensate possible shortcomings related to it (e.g. higher difficulty in evading income taxes, longer working hours). This decision can be modelled by the latent variable I_{1i}^{*} that summarizes the willingness to get a formal job:

$$\begin{aligned} d_{1i} &= 1 \text{ if } I_{1i}^{*} = X_{1i}^{'}\beta_1 + u_{1i} > 0 \\ d_{1i} &= 0 \text{ if } I_{1i}^{*} = X_{1i}^{'}\beta_1 + u_{1i} \leq 0 \end{aligned} \tag{1}$$

where d_{1i} is an indicator equal to one if the worker is in the queue and equal to zero if the worker is not in the queue, X'_{1i} is a vector of individual characteristics assumed to determine the individual decision to join the queue, β_1 is a vector of parameters and u_{1i} is an idiosyncratic component that captures unobserved heterogeneity in the preference for a formal job and omitted variables.

Similarly, the employer's decision can be modelled by a latent variable, I_{2i}^{*}, that captures his/her perception about the worker's productivity. Formal jobs are associated to higher firing and hiring costs and also to some mandatory fringe benefits. When deciding to hire a worker, a formal employer must evaluate whether or not the worker's productivity compensates the overall cost. The employer's decision, then, is based on the difference between the worker's productivity and his/her associated costs.

$$\begin{aligned} d_{2i} &= 1 \text{ if } I_{2i}^{*} = X_{2i}^{'}\beta_2 + u_{2i} > 0 \\ d_{2i} &= 0 \text{ if } I_{2i}^{*} = X_{2i}^{'}\beta_2 + u_{2i} \leq 0 \end{aligned} \tag{2}$$

where d_{2i} is an indicator equal to one if the worker is chosen from the queue and equal to zero if the worker is not chosen from the queue, X'_{2i} is a vector of observable individual characteristics assumed to determine the employer's decision to select a worker from the queue, β_2 is a vector of parameters and u_{2i} is an idiosyncratic component that captures unobserved heterogeneity in the employer's perception and/or omitted variables.

The endogeneity of the switching model resides in the fact that the difference between the worker's predicted wage in the formal sector and the worker's predicted wage in the informal sector enters the "in the queue" equation as an additional explanatory variable. Hence, sector allocation determines wage and wage differential determines sector allocation. The first problem to overcome in order to estimate a structural probit (or bivariate probit) like that is to get the correct equation to predict the wage differential. The wage equations for the formal and informal sectors can be estimated as:

$$W_{fi} = Z_{fi}\gamma_f + \varepsilon_{fi} \tag{3}$$

$$W_{ii} = Z_{ii}\gamma_i + \varepsilon_{ii} \quad (4)$$

where W_{fi} and W_{ii} are the log hourly wage rate for formal sector and informal workers respectively, Z_{fi} and Z_{ii} are the variables that determine wages including both individual and industry characteristics and γ_f and γ_i are vectors of parameters. However, it is well known that OLS estimates can be biased for not taking into account the sector allocation decision made by workers, and in a job queue approach also by the employers. OLS estimates assume that $E(\varepsilon_{fi} \mid X_{1i}) = 0$, but W_{fi} is only observed if the worker queued for a formal job and was chosen from the queue. This condition, in general, may lead to the violation of the assumption that the conditional expectation of the residuals is equal to zero, since the sample of formal workers is not randomly drawn from the population. The violation is due to the fact that there would be correlation between Z_{fi} and ε_{fi} operating through the relationship between ε_{fi} and the pair u_{1i} and u_{2i} in equations (1) and (2), respectively. More specifically estimates of equation (3) will be biased if $E(\varepsilon_{fi} \mid X_{1i}, I_{1i}^* > 0 \quad \text{and} \quad I_{2i}^* > 0) \neq 0$.

The wage equation for informal workers is also censored, but the rule is a bit more complicated than the one for the formal sector. This is due to the fact that there are, at least, two different types of informal workers. The ones who are in the queue and were not chosen from the queue – $E(\varepsilon_{ii} \mid X_{1i}, I_{1i}^* > 0 \quad \text{and} \quad I_{2i}^* < 0)$ - and the ones who did not join the queue – $E(\varepsilon_{ii} \mid X_{1i}, I_{1i}^* < 0)$. In the next section we will discuss this point more thoroughly since its estimation is dependent on the assumptions we make regarding the degree of observability of the "in the queue" status.[10]

Similarly to the "in the queue" equation, the wage that a worker would command should he/she be hired by a formal employer may also enter as an extra explanatory variable in the "chosen from the queue" equation. As mentioned earlier, formal employers try to minimize labour cost conditional on worker's productivity. Hence, given the productive characteristics of a worker, the wage he/she would get in the formal sector should enter the "chosen from the queue" with a negative sign. The more expensive the worker given his/her productivity potential, the lower the probability that he/she would be selected from the queue. Again, we face the task of estimating the wage that any worker in the population would get once in the formal sector.

In this context a worker would be found in the formal sector if:

$$I_{1i}^* = \alpha_1(W_{fi} - W_{ii}) + X'_{1i}\beta_1 + u_{1i} > 0 \quad (5)$$

and

$$I_{2i}^* = \alpha_2(E(W_{fi} \mid I_{1i}^* > 0)) + X'_{2i}\beta_2 + u_{2i} > 0 \quad (6)$$

where $E(W_{fi} \mid I_{1i}^* > 0)$ is the expected wage that a worker would get in the formal sector should he/she be in the queue; α_1 is the parameter of the wage differential and α_2 is the parameter of the predicted formal wage should the worker be in the queue.

10. If we had full observability we should also observe the ones who were not in the queue and would be chosen from it.

X'_{1i}, X'_{2i}, u_{1i}, u_{2i}, β_1, and β_2 are defined as before.[11] In the next section we discuss how to estimate this structural bivariate probit according to different assumptions about the degree of observability of the "in the queue" status.

3 ESTIMATION PROCEDURE

Endogenous switching regression models such as the one described in the previous section can be estimated either in two or three steps. The two-step procedure would imply the joint estimation of the "in the queue" and "chosen from the queue" reduced form equations and of the wage equations for the two sectors – formal and informal – through the maximization of a likelihood function and then the estimation of a structural bivariate probit with the predicted wage differential as an additional regressor.

The three-step procedure would imply the estimation of a reduced form bivariate probit whose residuals would be used to approximate the non-zero expectation of ε_{fi} and ε_{ii} conditional on u_{1i} and u_{2i}. These approximations, the so-called Inverse Mills ratio, would enter the wage regression for each sector in a second step in order to restore the assumption of zero expectation of the wage regression residuals. Finally, the structural bivariate probit would be estimated using as additional regressors both the (offered) wage differences estimated from the wage regression of the second step in the "in the queue equation" and the wage that a worker in the formal sector would get should he/she be in the queue in the "chosen from the queue" equation.

A major issue that affects both the maximum likelihood (two-step procedure) and the three-step procedure refers to their reliance on the trivariate normality of the residuals of the sector allocation equation and of the wage equations. The assumption of normality makes the correct specification of the model more important than in the usual linear regression framework. Recent researches have relaxed this assumption by estimating the sector allocation equation without assuming any particular distribution, but the conclusions of these semiparametric attempts to correct for selectivity in order to overcome the bias triggered by erroneously assuming a trivariate (or bivariate) normal distribution for the residuals are mixed at best.[12]

Another problem related to the strategies to correct for selectivity bias is the issue of identification. The need for an exclusion restriction in order to separately identify the selection and wage equations is not relaxed by semiparametric techniques; on the contrary, it makes it even more restrictive, since it requires that at least one of the excluded variables to be continuous (Heckman, 1990). Therefore, one has to find a variable that determines the "in the queue" status, but does not determine the "chosen from the queue" status (or vice-versa). Moreover an additional

11. In the next section we will address the issue of identification, when it will become clear that X'_{1i}, X'_{2i} in the structural bivariate probit are in fact a subset of X'_{1i}, X'_{2i} used in a first step in order to estimate the reduced form of the bivariate probit.

12. See for instance Newey, Powell and Walker (1990) for an application of semiparametric methods to correct for selectivity that do not imply major differences in relation to the standard "normal-based" procedure. However, Lanot and Walker (1998), Schafgans (1998) and Martins (2001) find substantial differences in the results yielded by those two procedures.

exclusion restriction is necessary in order to identify the parameters of the wage equations. So, we have to find, at least, one variable that determines the wage, but does not affect either the "in the queue" or "the chosen from the queue" decisions.[13]

So far we have skipped the discussion of the implication of partial observability for the estimation of both reduced and structural forms bivariate probit of the job queue model. In general, survey data only bring information on the actual status of workers: formal or informal. We do not know whether informal workers wish to get a formal job or if they prefer to stay as informal workers for whatever reason. Thus, we do not observe either d_{1i} or d_{2i}, but only the product $d = d_{1i} * d_{2i}$. If this product is equal to one, the worker is in the formal sector, if it is equal to zero he/she is in the informal sector. Note that besides the lack of information on whether informal workers are in the queue for formal jobs, we also have no information on whether or not formal employers would like to hire workers who are not in the queue. This latter type of partial observability is modelled by Poirier (1980) and assumes that both decisions – join the queue and choose from the queue – are taken simultaneously. Abowd and Farber (1982) assume that the sector allocation process can be modelled as a bivariate probit with partial observability and sequential decision as described in the last section. Employers only hire workers from the pool of workers who are in the queue, thus the distribution of u_{2i} is defined only over the subpopulation for which $d_{1i} = 1$. In this case, one can only make conditional inferences, whereas in Poirier's procedure allows both conditional and marginal inferences (Maddala, 1983).

A shortcoming of the bivariate probit with partial observability discussed above is the severe degree of partial observability. The dependent variable in both the "in the queue" equation and the "chosen from the queue" equation is the same. This occurs because we do not observe the two types of informal workers: the ones who are not in the queue and the ones who are in the queue and are not chosen from it. Abowd and Farber approach aims to fulfil this lack of information by matching the characteristics of workers for which $d = 1$ with those of workers for which $d = 0$, after having distinguished the set of characteristics that determine the probability that $I_1^* > 0$ through their effect on the "in the queue" status from the set of characteristics that influence the same probability through their effect on the "chosen from the queue" status. Therefore, the exclusion restrictions play a fundamental role on this identification process.

One way to get more information is to rely on survey questions that try to measure the willingness to get a formal job.[14] If we get an information like that, e.g., if the informal worker would like to switch to a formal job, we could estimate a bivariate probit with sample selection.[15] It is important to note that despite having

13. Note that the reduced form bivariate probit includes the variables used to identify the wage equation, nevertheless, they are excluded from both the "in the queue" equation and the "chosen from the queue equation" in the estimation of the structural bivariate probit.

14. Farber (1983) adopts a similar strategy and adds the information of whether or not the non-union worker would vote in favor of unionization.

15. This approach is quite common in attempts to control for both labour market participation and employment decision. As we only observe the employment status of the workers who participate in the labour market, the bivariate probit is censored for the sample that does not participate in the labour market. See Mohanty (2002) and Co *et al.* (2002) for interesting applications.

two different dependent variables (defined over different populations), we still do not have full observability. This is so because we do not observe formal workers who would like to become informal workers. However, if we assume that there are no barriers to entry in the informal sector, we can say that all formal workers desire a formal job. Another feature of this approach is that similar to Abowd and Farber (1982) approach, it assumes a sequential decision process, implying that formal employers only hire workers from the pool in the queue.

Assuming that u_{1i} and u_{2i} have a bivariate normal distribution with means zero and unit variances and zero covariance,[16] the likelihood function to be maximized for Abowd-Farber bivariate probit with partial observability and sequential decision process can be written as:[17]

$$L_1 = \prod_{d=1} [\Phi(\tilde{X}'_{1i}\beta_1)\Phi(X'_{2i}\beta_2)] \prod_{d=0} [1-\Phi(X'_{1i}\beta_1)\Phi(\tilde{X}'_{2i}\beta_2)] \qquad (7)$$

where Φ is the normal cumulative distribution and $\tilde{X}_{ji}$ are the explanatory variables of the in the queue $(j=1)$ and the chosen from the queue $(j=2)$ equations, including the variables that are assumed to affect only wages, since this first step corresponds to the estimation of a reduced form bivariate probit.

If one assumes that the sector allocation process is based on a joint (simultaneous) decision as put forward by Poirier (1980), the log likelihood function to be maximized can be written as:

$$L_2 = \prod_{d=1} \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho) \prod_{d=0} [1-\Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)] \qquad (8)$$

where Φ_2 is the bivariate normal cumulative distribution and ρ is the covariance between u_{1i} and u_{2i}.[18]

Note that if we assume that there is no correlation between the residual of the two equations, $\rho = 0$ in equation (8), the joint decision model becomes identical to the sequential decision model in equation (7).[19]

Finally, the likelihood function of the bivariate probit with sample selection for the in the queue and the chosen from the queue equations is:

$$L_3 = \prod_{d_1=1, d_2=1} \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho) \prod_{d_1=1, d_2=0} \Phi_2(\tilde{X}'_{1i}\beta_1, -\tilde{X}'_{2i}\beta_2; -\rho) \prod_{d_1=0} \Phi(-\tilde{X}'_{1i}\beta_1) \qquad (9)$$

It is worth noting that in this latter case we clearly have two different dependent variables, one for each "selection" equation.

16. Note that the Abowd-Farber bivariate probit assumes that both decisions are uncorrelated.

17. Despite being called Abowd and Faber bivariate probit, this approach was not empirically implemented in their 1982 paper. Actually, in that paper they estimate a modified likelihood function that incorporates the concept of "job rights". They combine a simple probit for the ones with job right and a bivariate probit for those without "job rights". Workers with "job rights" are the ones who held a union job in the previous year, were not fired from that job and did no quit that job to take another union job.

18. This is so because $Var(u_{1i}) = 1$ and $Var(u_{2i}) = 1$ due to normalization.

19. Maddala (1983) argues that if the aim of the research were to find out which factors influence the employer's decision in hiring a specific type of worker, a simultaneous framework would be more appropriate.

A first test for the existence of a job queue for formal jobs can be implemented through the imposition of restrictions on the bivariate estimation. Abowd and Farber (1982) suggest that despite the univariate probit not being nested in the bivariate probit, the adequacy of the bivariate specification can be assessed through a Likelihood Ratio test that constrains all the coefficients of the chosen from the queue equation to zero $(\beta_2 = 0)$, with exception of the constant term that is arbitrarily fixed as a positive number large enough to ensure that all workers in the queue are chosen from the queue, i.e., $P(I_2^* > 0 \mid I_1^* > 0) = 1$. Mengistae (1998) extends this idea and puts forward the same test for the "in the queue" equation, so that one can test the hypothesis of a universal queue for formal jobs, *i.e.*, there is no worker who would prefer to stay in a informal job. The procedure is symmetrical to the one described above. All coefficients of the "in the queue" equation are constrained to zero, with exception of the constant that is fixed as a positive number large enough to ensure that all workers are in the queue for formal jobs.

Based on the results of the bivariate probit, one can estimate the probability that workers with a specific set of characteristics are a) found in the formal sector; b) in the queue and c) chosen from the queue conditional on being in the queue.[20] In the sequential case with partial observability, these probabilities can be calculated, respectively, as:

a1) $Prob(I_{1i}^* > 0 \text{ and } I_{2i}^* > 0) = \Phi(\tilde{X}'_{1i}\beta_1)\Phi(\tilde{X}'_{2i}\beta_2)$

b1) $Prob(I_{1i}^* > 0) = \Phi(\tilde{X}'_{1i}\beta_1)$

c1) $Prob(I_{2i}^* > 0 \mid I_{1i}^* > 0) = \Phi(\tilde{X}'_{2i}\beta_2)$

For the simultaneous case with partial observability and for the sample selection case we would have:[21]

a2) $Prob(I_{1i}^* > 0 \text{ and } I_{2i}^* > 0) = \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)$

b2) $Prob(I_{1i}^* > 0) = \Phi(\tilde{X}'_{1i}\beta_1)$

c2) $Prob(I_{2i}^* > 0 \mid I_{1i}^* > 0) = \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)/\Phi(\tilde{X}'_{1i}\beta_1)$

An estimate of the length of the queue, q, can be obtained by the inverse of the average probability of being chosen from the queue given that the worker joined the queue, $Prob(I_{2i}^* > 0 \mid I_{1i}^* > 0)$ (Farber, 1983 and Venti, 1987) for each one of the bivariate models:

20. One could also estimate the probability that a worker would be chosen from the queue, but this probability is only meaningful in the simultaneous case, since as discussed above, only in this case it is possible to make marginal inferences regarding the chosen from the queue equation.

21. This feature does not imply that both methods yield the same results since their likelihood functions are different.

$$q_1 = \frac{N}{\sum_{i=1}^{N} \Phi(\tilde{X}'_{1i}\beta_1)} \tag{10}$$

and

$$q_2 = \frac{N}{\sum_{i=1}^{N} \Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)/\Phi(\tilde{X}'_{1i}\beta_1)} \tag{11}$$

where N is the number of observations.

As seen in the last section, the OLS estimates of the wage equation for both formal and informal workers can be biased due to the censoring rules generated by the sector allocation process. Heckman (1979) puts forward a technique to approximate the non-zero expectation of the residuals in the univariate (probit) case. The extension for the bivariate case is quite straightforward. Assuming that the residual has a normal distribution the first moment of a truncated distribution from below is given by (Maddala, 1983):

$$E(\varepsilon_{fi} \mid I_1^* > 0, I_2^* > 0) = \sigma_{1f}\lambda_{1fi} + \sigma_{2f}\lambda_{2fi} \tag{12}$$

where

$$\lambda_{1fi} = \frac{\phi(\tilde{X}'_{1i}\beta_1)\Phi(\tilde{X}'_{2i}\beta_2 - \rho\tilde{X}'_{1i}\beta_1)/\sqrt{(1-\rho^2)}}{\Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)} \tag{13}$$

and

$$\lambda_{2fi} = \frac{\phi(\tilde{X}'_{2i}\beta_2)\Phi(\tilde{X}'_{1i}\beta_1 - \rho\tilde{X}'_{2i}\beta_2)/\sqrt{(1-\rho^2)}}{\Phi_2(\tilde{X}'_{1i}\beta_1, \tilde{X}'_{2i}\beta_2; \rho)} \tag{14}$$

and σ_{1f} and σ_{2f} are, respectively, the correlation between the residuals of the wage equation of the formal sector and the residuals of the "in the queue" equation and of the "chosen from the queue" equation. Then we can approximate the non-zero expectation by the inclusion of estimates of λ_{1fi} and λ_{2fi}, the so-called Inverse Mills ratios, in the OLS regression. This procedure yields consistent estimates of the wage equation parameters and also allows us to test for the presence of selectivity in the formal sector via a simple t-test of significance of the parameters σ_{1f} and σ_{2f}.[22]

However, the bivariate probit with partial observability and sequential decision implies that the correlation between the "in the queue" equation and "the chosen

22. Ahn (1992) shows that the F-statistic of a model with double selection criteria is asymptotically equivalent to the LM test statistic.

from the queue" equation is (assumed) to be zero $\rho = 0$. In this case (13) and (14) simplify to:

$$\lambda_{1fi} = \frac{\phi(\tilde{X}'_{1i}\beta_1)}{\Phi(\tilde{X}'_{1i}\beta_1)} \tag{15}$$

and

$$\lambda_{2fi} = \frac{\phi(\tilde{X}'_{2i}\beta_2)}{\Phi(\tilde{X}'_{2i}\beta_2)} \tag{16}$$

Therefore the wage equation for workers in the formal sector can be estimated as:

$$E(W_{fi} \mid d = 1) = Z_{fi}\gamma_f + \sigma_{1f}\lambda_{1fi} + \sigma_{2f}\lambda_{2fi} + \eta_{fi} \tag{17}$$

where $\eta_{fi} = \varepsilon_{fi} - \sigma_{1f}\lambda_{1fi} - \sigma_{2f}\lambda_{2fi}$ so that $E(\eta_{fi} \mid d = 1) = 0$.

The unknown Inverse Mills ratio λ_{1fi} and λ_{2fi} can be approximated by substituting the bivariate probit estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ in (13) and (14) or (15) and (16) according to the model in use, simultaneous or sequential.

The major problem with this approach resides in estimating a similar wage equation for workers in the informal sector, $(d = 0)$. In the simultaneous case, it is not possible to derive a correction based on the Inverse Mills ratio to estimate the wage equation for the subsample $(d = 0)$, since instead of two, we would have three types of informal sector workers, not clearly defined (Tunali, 1986). Nevertheless, such correction is readily available for the bivariate probit with sample selection. The expected wage for workers not in the queue $(I^*_{1i} < 0)$ can be written as:

$$E(W^1_{ii} \mid I^*_{1i} < 0) = Z_{ii}\gamma_i + \sigma_{1i}\lambda_{3ii} + \eta^1_{ii} \tag{18}$$

where, $\lambda_{3ii} = -\phi(\tilde{X}'_{1i}\beta_1)/[1 - \Phi(\tilde{X}'_{1i}\beta_1)]$ and $\eta^1_{ii} = \varepsilon_{ii} - \sigma_{1i}\lambda_{3ii}$. On the other hand, the expected wage for workers who are in the queue, but were not chosen from the queue can be written as:

$$E(W^2_{ii} \mid I^*_{1i} > 0 \quad \text{and} \quad I^*_{2i} < 0) = Z_{ii}\gamma_i + \sigma^*_{1i}\lambda_{1ii} + \sigma_{2i}\lambda_{4ii} + \eta^2_{ii} \tag{19}$$

where, $\lambda_{4ii} = -\phi(\tilde{X}'_{2i}\beta_2)/[1 - \Phi(\tilde{X}'_{2i}\beta_2)]$ and $\eta^2_{ii} = \varepsilon_{ii} - \sigma_{1i}\lambda_{1ii} - \sigma_{2i}\lambda_{4ii}$.

As for the Abowd and Farber bivariate probit, Mengistae (1998) shows that the expected wage for workers with $d = 0$ can be estimated as the weighted average between the expected wage for workers with $I^*_{1i} < 0$ (not in the queue workers) and

the expected wage for workers with $I_{1i}^{*} > 0$ and $I_{2i}^{*} < 0$ (in the queue but not chosen from the queue workers)

The overall expected wage for the informal sector can be written as $(1-\pi)E(W_{ii}^{1} \mid I_{1i}^{*} < 0) + \pi E(W_{ii}^{2} \mid I_{1i}^{*} > 0 \quad \text{and} \quad I_{2i}^{*} < 0)$ [23] or:

$$E(W_{ii}) = Z_{ii}\gamma_i + \sigma_{1i}\lambda_{3ii} + \delta_1\lambda_{1i}^{*} + \delta_2\lambda_{4ii} + \eta_{ii} \qquad (20)$$

where, $\lambda_{1i}^{*} = \lambda_{1ii} - \lambda_{3ii}$, $\delta_1 = \pi\sigma_{1i}$, $\delta_2 = \pi\sigma_{2i}$ and $\eta_{ii} = (1-\pi)\eta_{ii}^{1} + \pi\eta_{ii}^{2}$.

The bivariate probit estimates from the sequential model estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ are used in (15) and (16) to approximate the unknown Inverse Mills ratio.

The standard errors of the estimated parameters for both formal and informal workers need to be corrected to account for the heterogeneity introduced by the λ_i's and for the fact that they are first-step estimated variables and not an observed variable, we do that following the adaptation of Mengistae (1998) for Ham's (1982) methodology for the correction of the standard errors of estimated parameters from equations with two selectivity criteria.

An interesting feature of this approach for the estimation of the wage equation for the informal sector is that we can test directly the existence of a queue for formal jobs and get the proportion of (informal) workers queueing, π, for a formal job from this equation. Mengistae (1998) points out that a test for the absence of a job queue, $\pi = 0$, can be based on the asymptotic joint significance test of δ_1 and δ_2. Similarly, a test for a universal job queue, $\pi = 1$, can be done by testing the hypothesis that $\sigma_{1i}^{*} = \delta_1$. The rejection of both hypothesis means that there is a partial queue for formal jobs.

The third step of the endogenous switching regression model consists in estimating the "in the queue" and the "chosen from queue" equations through a structural bivariate probit. The procedure is based on the inclusion of an additional regressor - the wage difference between the expected (offered) wage that the worker would face in the formal sector and his/her expected (offered) wage in the informal sector – in the "in the queue" equation. We can also add the wage that he/she would get in the formal sector conditional on being in the queue in the "chosen from the queue" equation. Note that we also have to exclude the variables that only affect the wages from this equation. The structural bivariate probit can only be meaningfully estimated for the Abowd and Faber approach and for the bivariate probit with sample selection,[24] since there is no estimate for the informal sector wage equation available from the simultaneous model.

23. Where π corresponds to the proportion of informal workers who are in the queue and were not chosen from it and $(1-\pi)$ is the proportion of informal workers who are not in the queue. Note that π approximates the size of the queue in this approach.

24. In the case of the bivariate probit with sample selection the wage differences between the offered wage in the formal sector and the offered wage in the informal sector are based on different equations for the informal sector, depending on the desire to switch for a formal job.

A final point to be addressed refers to the choice of appropriate and believable exclusion restrictions. Most papers that deal with formal *versus* informal sector choice usually assume that variables such as the presence of children or elderly people in the household and other members of the household income play a crucial role in determining the choice of sector, but do not affect the individual's wage. It is argued that the informal sector offers some sort of flexibility in terms of hours and workplace that makes it more desirable for people who have children and cannot pay for childbearing services (Marcoullier *et al.*, 1997).

In order to identify the sector allocation equations in the case of Brazil, we opt to use a wide set of variables that should capture how an individual in a specific household would value a formal job *vis-à-vis* an informal one. In order to do that, we build the following set of variables that are used in the sector allocation equation(s), but not in the wage equations: 1) variables indicating the number of children and elder people per age group in the household; 2) variables indicating number of other household members by specific working status (e.g. number of registered workers, non-registered workers, self-employed, unemployed); 3) last occupational status;[25] 4) income of the other members of the household (*per capita*).

The assumptions behind these variables are quite straightforward. The use of the number of children as a proxy for home workload tries to capture whether or not the presence of children makes people more willing to get an informal job that could give them more "flexibility" in terms of hours. This is particularly important for women who bear most of the childbearing cost. For men, particularly for the heads of the household, the presence of children could make an unemployment spell more costly, making them more likely to accept an informal job. The variables related to the distribution of occupational status among other household members intend to capture whether or not the presence of formal workers in the household makes other members of the household less demanding in terms of job characteristics and then less resistant to get an informal job. The presence of formal workers in the household would act as an insurance for the other members of the household. But it also could be the case, that once knowing the benefits related to the possession of a registered work-card, other members would become more reluctant in getting a job without registration. The variables related to the recent employment history of the individual, as measured by her last occupation before moving to the current job, try to capture some inertia factors that would make a worker more or less likely to get a formal job. A former holder of a formal job must have less incentive to get an informal job than a former informal worker with similar characteristics. Similarly, employers tend to look at the employment history of the job applicant when making decision whether to hire him/her. Finally, we also incorporate the *per capita* income of other members of the household in order to assess how others "disposable incomes" affect the willingness to join a formal job.

The variables listed above enter the univariate probit together with human capital variables and industry and regional characteristics. However, in order to identify separately the "in the queue" equation and the "chosen from the queue"

25. Last occupational status is only available for workers who had switched to the current job less than 5 yeas ago. For this reason we include "being in the current job for more than 5 years" as a variable in this category.

equation in a bivariate context, we exclude the following variables from the "chosen from the queue" equation: tenure, variables related to the number of children in the household and older people, occupational status of the other members of the household, and *per capita* income of other members of the household. Tenure, or seniority on the current job, has been traditionally included only in the "in the queue" equation.[26] In doing so, we assume that the worker takes into account the loss of specific human capital and of the benefits that accrue from seniority when joining the queue, but that this specific human capital is not useful for the decision of the employer. Experience as a proxy for general human capital should be the variable that employers rely on when taking the hiring decision.[27] We assume that the variables related to the household situation do not affect the decision of the employers with the exception of the marital status, that traditionally has been used as a signal for commitment to work in the literature of economics of personnel. Thus, the number of children, the number of elder people in the household, the working status of other members and the *per capita* income of the others members of the household only affect the decision to join the queue. Additionally, in order to identify the wage equation we assume that industry affiliation only affects wages, therefore, we exclude the variables related to industry affiliation from the structural bivariate probit.

4 DATA

The data used in this paper comes from the 1990 Brazilian Annual Household Survey (Pnad). We chose this specific year because its survey questionnaire had a special supplement where non-registered workers and self-employed workers were asked whether or not they would like to switch to a formal sector job. This question allows us to relax the severe partial observability of both Abowd and Farber (1982) and Poirier (1980) bivariate probit models. Moreover, this supplement also investigate the past employment history of the individuals yielding a much richer data set than the ones usually available.

In order to properly build variables related to the family composition (e.g. the number of other members of the household who are unemployed and the number of children per age group), we exclude from our sample all persons who are not related to the head of the household where they live. This procedure excludes, for instance, domestic servants who live in the house of their employers.

The sample used in the estimation is also restricted to workers who worked more than 20 hours per week and who were between 15 and 64 years old. We dropped from the sample all individuals for whom there were missing observations for any of the variables used in either the bivariate probit or the wage equations. The resultant sample has 60,138 observations from which 43,322 are formal (registered) workers and 16,816 are informal (non-registered) workers. Table 1 depicts the means and the standard deviation of all variables used in the analysis for the full sample and for formal workers and informal workers sub-samples.

26. See Abowd and Farber (1982) and Faber (1983).

27. We are aware that this assumption is a bit controversial and for this reason, we also check the robustness of the results to the inclusion of the tenure in the "chosen from the queue" equation.

The log hourly wage wages for informal workers is substantially lower than for formal workers in 1990. Informal workers have lower schooling, only 10% of them have completed the secondary education or more, whereas 33% of formal workers have at least the secondary education. Male, married and white workers are more likely to be formal as well as workers who live in metropolitan areas and in the South or in the Southeast regions. Formal workers have more than the double of the tenure of their informal counterparts (5.3 *versus* 2.3 years). Formal workers on average have more experience[28] than informal workers.

TABLE 1
Descriptive statistics

	Full sample		Formal sector		Informal sector	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
log hourly wage	0.43	0.999	0.68	0.921	-0.20	0.914
illiterate (base)	0.08	0.275	0.06	0.236	0.14	0.349
some elementary	0.14	0.352	0.12	0.322	0.22	0.411
complete elementary	0.36	0.480	0.33	0.472	0.42	0.494
complete primary	0.16	0.368	0.18	0.381	0.12	0.330
complete secondary	0.19	0.391	0.23	0.421	0.08	0.266
complete college	0.06	0.244	0.08	0.274	0.02	0.132
exp/10	1.86	1.259	1.92	1.234	1.70	1.306
exp/100	5.03	6.272	5.20	6.151	4.60	6.555
tenure	4.42	5.752	5.25	6.068	2.29	4.140
sex (male=1)	0.63	0.482	0.65	0.477	0.59	0.492
metropolitan	0.51	0.500	0.56	0.497	0.38	0.484
race (white=1)	0.53	0.499	0.57	0.495	0.42	0.494
new entrants (base)	0.24	0.428	0.21	0.410	0.31	0.463
more than 5 years	0.26	0.437	0.31	0.462	0.12	0.329
former formal worker	0.30	0.457	0.34	0.475	0.18	0.383
former informal worker	0.16	0.368	0.10	0.294	0.33	0.471
former public servant	0.00	0.069	0.01	0.073	0.00	0.061
former self-employed	0.04	0.192	0.03	0.177	0.05	0.225
#children 0-3 years	0.96	1.135	0.90	1.066	1.11	1.285
#children 4-6 years	0.28	0.542	0.27	0.525	0.32	0.582
#children7-10 years	0.43	0.691	0.39	0.658	0.52	0.763
#children 11-13 year	0.33	0.592	0.29	0.562	0.43	0.654
# elder	0.07	0.272	0.06	0.266	0.07	0.286
# formal	0.70	0.960	0.76	0.996	0.56	0.845
#self-employed	0.25	0.522	0.21	0.480	0.34	0.607
#informal	0.38	0.777	0.26	0.604	0.71	1.036
#unemployed	0.11	0.372	0.11	0.365	0.12	0.388
per capita income (others)	0.15	0.287	0.16	0.298	0.11	0.255
married	0.52	0.500	0.57	0.495	0.38	0.485
Northeast (base)	0.24	0.425	0.21	0.405	0.31	0.463
North	0.09	0.285	0.08	0.275	0.11	0.308
Southeast	0.39	0.487	0.42	0.493	0.31	0.462
South	0.17	0.373	0.19	0.392	0.11	0.313
Mid-West	0.12	0.326	0.10	0.306	0.16	0.370
Manufacturing (base)	0.27	0.444	0.31	0.462	0.17	0.377
constructing	0.08	0.277	0.06	0.246	0.13	0.340
retail	0.14	0.350	0.15	0.354	0.13	0.339
lodging, food and personal services	0.18	0.386	0.11	0.310	0.37	0.484
Productive Sector	0.21	0.409	0.25	0.432	0.12	0.326
Social Services	0.11	0.311	0.12	0.329	0.07	0.254
N	60138		43322		16816	

Formal workers are mainly concentrated on the manufacturing sector (31%) and productive services (25%) (banking, telecommunications, transport), whereas informal workers are prevalent in personal, food and lodging services (37%). As for

28. Experience is calculated as $exp = age - 6 - yearsofschooling$.

their last occupation within a 5 year-period, formal and informal workers present different patterns. Both groups show a high degree of inertia, 33% of informal workers used to be informal workers in their last occupation and 34% in the case of formal workers were also formal worker on their last job. However, whereas 31% of informal workers are "new entrants", i.e., they had no previous experience in the labour market, only 21% of the formal workers fall in this category. For formal workers 26% are in the current job for more than 5 years, in contrast, only 12% of informal workers are in the current job for more than 5 years.[29] Only 10% of formal workers were informal in the previous job, whereas 18% of the informal workers were formal. Finally, only 3% of formal workers and 5% of informal workers were self-employed before switching to their current job. These figures suggest that informal jobs seem to be the most common entry into the labour market for many young workers and also suggest that the last occupational status is a good predictor of the current one.

The number of children per age group in the household is higher for informal workers than for registered workers for all age groups. The number of people with more than 70 years is also higher for informal workers. The number of other members of the household who are formal workers is higher for registered workers than for informal workers. Similarly, informal workers are much more likely to live with other informal workers and with unemployed people. Formal workers also live in household where the *per capita* income of other members is higher in comparison to informal workers.

5 RESULTS

5.1 UNIVARIATE PROBIT REDUCED FORM

We will use the results provided by the standard endogenous switching regression model with just one selection equation as a benchmark to evaluate the results for the different specifications of bivariate probit discussed in Section 3.

The reduced-form probit of the first step in the estimation of the selectivity-corrected wage equations yields a good description of how the variables included in the (reduced form) model affect the sector choice made by workers in the absence of queue for formal jobs. The results on table 2[30] show that all education groups are more likely to be formal workers than the illiterate group (base category). In fact, the probability of choosing the formal sector increases almost monotonically with the educational level. It also increases with experience but with a decreasing rate. Tenure (or seniority in the current job) has a positive impact on the probability of being a formal worker. The coefficients on gender and on race, despite their positive sign, are not statistically significant, hence, it suggests that there is no "discrimination" on the sector choice (allocation) process. Workers who live in metropolitan areas are more likely to choose formal jobs. Workers who were formal workers in the previous job are more likely to be formal in the current one than "new entrants" (base category) as

29. This figure does not come as a surprise given the large difference in tenure between the two groups.
30. Notice that we report both the coefficient and the marginal effects.

well as workers who have been in the job for more than 5 years and former public servants and self-employed. However, former informal workers have a lower probability than "new entrants" to be formal in the current job. As for the effect of industry affiliation, all sectors seem to have a lower probability of having formal workers than the manufacturing sector (base category), particularly, lodging, food and personal services and the constructing industry.

TABLE 2

Univariate probit specification

	2-STEP				MLE	
	Coeff.	s.e.	Marg. Effect	s.e.	Coeff.	s.e.
Constant	-0.50	0.039	-0.15	0.012	-0.44	0.037
some elementary	0.08	0.026	0.02	0.007	0.08	0.026
complete elementary	0.26	0.025	0.07	0.007	0.24	0.025
complete primary	0.55	0.030	0.14	0.006	0.51	0.029
complete secondary	0.80	0.032	0.19	0.006	0.74	0.031
complete college	0.81	0.044	0.18	0.006	0.68	0.041
exp/10	0.36	0.021	0.11	0.006	0.34	0.020
exp/100	-0.06	0.004	-0.02	0.001	-0.06	0.004
tenure	0.05	0.002	0.01	0.001	0.05	0.001
sex (male=1)	0.00	0.015	0.00	0.004	0.00	0.014
metropolitan	0.19	0.013	0.06	0.004	0.17	0.013
race (white=1)	0.00	0.014	0.00	0.004	-0.01	0.014
more than 5 years	0.32	0.021	0.09	0.006	0.37	0.020
former formal worker	0.59	0.019	0.16	0.005	0.62	0.019
former informal worker	-0.16	0.020	-0.05	0.006	-0.13	0.019
former public servant	0.26	0.088	0.07	0.021	0.29	0.077
former self-employed	0.07	0.033	0.02	0.009	0.14	0.030
#children 0-3 years	-0.03	0.009	-0.01	0.003	-0.02	0.009
#children 4-6 years	0.01	0.016	0.00	0.005	0.01	0.015
#children7-10 years	-0.02	0.012	-0.01	0.003	-0.02	0.011
#children 11-13 year	-0.02	0.011	-0.01	0.003	-0.03	0.011
# elder	-0.05	0.023	-0.02	0.007	-0.10	0.023
# formal	0.18	0.007	0.05	0.002	0.15	0.007
#self-employed	-0.07	0.012	-0.02	0.004	-0.10	0.011
#informal	-0.19	0.008	-0.06	0.002	-0.21	0.008
#unemployed	0.02	0.017	0.01	0.005	0.01	0.015
per capita income (others)	-0.16	0.024	-0.05	0.007	0.15	0.016
married	0.13	0.016	0.04	0.005	0.11	0.016
North	0.08	0.024	0.02	0.007	0.08	0.024
Southeast	0.28	0.017	0.08	0.005	0.25	0.017
South	0.36	0.022	0.10	0.005	0.32	0.022
Mid-West	0.05	0.022	0.01	0.006	0.01	0.022
constructing	-0.60	0.024	-0.21	0.009	-0.57	0.024
retail	-0.24	0.021	-0.07	0.007	-0.21	0.020
lodging, food and personal serv.	-0.90	0.019	-0.31	0.007	-0.86	0.019
Productive Sector	-0.07	0.020	-0.02	0.006	-0.07	0.020
Social Services	-0.15	0.025	-0.05	0.008	-0.14	0.025
Log L	-26054.75					
N	60138				60138	

Base categories: education: illiterate, past experience: entrants; region: Northeast; sector: manufacturing.

Workers who live in the Northeast (base category) are less likely to be formal than workers who live in any other region. Married workers are more likely to be formal than single workers and the number of children and the presence of older people in the household seem to have a negative impact on the probability of being formal. The "children effect", however, is not significant for the age groups 4 to 6 and 10 to 13 years of age. The presence of formal workers in the household increases the probability of the individual being in the formal sector whereas the presence of informal workers and self-employed reduces it. The number of unemployed workers

does not have a significant effect, but it has a positive sign. As for the other member of the household income (*per capita*) the probit equation suggests a negative impact on the probability of being formal.

5.2 BIVARIATE PROBIT REDUCED FORM

Abowd-farber bivariate probit

The reduced form of the Abowd-Farber bivariate probit with partial observability and sequential decision reveals interesting different patterns for the "in the queue" equation and for the "chosen from the queue" equation (table 3, part A). The probability of being in the queue is not significantly different among different education levels, whereas the probability of being chosen from the queue is significantly higher for workers with a higher educational level, particularly for the groups with complete secondary school or more. The results of the univariate specification discussed in the last subsection disguise the different impact of schooling on the allocation process observed in the bivariate context.

TABLE 3
Bivariate probit specifications

	Abowd-farber (A)				Poirier (B)				BP with sample selection (C)			
	IQ		CFQ		IQ		CFQ		IQ		CFQ	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	0.32	0.080	-0.44	0.064	0.71	0.082	-0.42	0.062	1.44	0.046	-0.67	0.037
some elementary	-0.06	0.054	0.19	0.044	-0.10	0.055	0.19	0.043	-0.01	0.031	0.17	0.026
complete elementary	-0.01	0.052	0.46	0.042	-0.10	0.053	0.44	0.040	-0.05	0.030	0.45	0.025
complete primary	0.02	0.060	0.93	0.050	-0.13	0.062	0.89	0.048	0.00	0.036	0.86	0.030
complete secondary	0.08	0.063	1.30	0.052	-0.12	0.065	1.28	0.051	0.06	0.038	1.26	0.033
complete college	0.06	0.077	1.26	0.068	-0.10	0.079	1.26	0.066	0.05	0.050	1.51	0.052
exp/10	0.09	0.037	0.49	0.035	0.03	0.038	0.47	0.034	0.08	0.024	0.46	0.021
exp/100	-0.04	0.007	-0.06	0.006	-0.03	0.007	-0.06	0.006	-0.04	0.004	-0.05	0.004
tenure	0.40	0.012			0.39	0.011			0.03	0.002		
sex (male=1)	0.00	0.026	-0.01	0.027	0.00	0.028	0.03	0.025	-0.08	0.018	0.12	0.015
metropolitan	0.16	0.025	0.25	0.024	0.11	0.026	0.24	0.022	0.08	0.017	0.23	0.014
race (white=1)	-0.02	0.026	0.03	0.025	-0.02	0.027	0.03	0.024	-0.08	0.018	0.09	0.014
more than 5 years	-0.45	0.057	0.28	0.030	-0.46	0.055	0.27	0.029	0.20	0.025	0.41	0.024
former formal worker	0.27	0.044	1.02	0.048	0.16	0.044	0.89	0.041	0.51	0.025	0.30	0.020
former informal worker	-0.41	0.050	0.33	0.057	-0.45	0.053	0.36	0.052	0.21	0.024	-0.36	0.019
former public servant	-0.12	0.123	1.10	0.415	-0.24	0.129	0.99	0.314	0.13	0.111	0.19	0.097
former self-employed	-0.31	0.060	0.87	0.126	-0.42	0.063	0.78	0.092	0.13	0.039	-0.14	0.032
#children 0-3 years	-0.02	0.015			-0.02	0.014			-0.02	0.011		
#children 4-6 years	-0.01	0.026			-0.02	0.023			0.03	0.019		
#children7-10 years	-0.04	0.019			-0.04	0.017			-0.01	0.014		
#children 11-13 year	-0.01	0.018			-0.01	0.016			0.02	0.013		
# elder	-0.09	0.039			-0.09	0.036			0.00	0.030		
# formal	0.25	0.014			0.24	0.013			0.16	0.009		
#self-employed	-0.08	0.019			-0.08	0.017			-0.03	0.014		
#informal	-0.23	0.012			-0.20	0.011			-0.09	0.009		
#unemployed	0.03	0.026			0.03	0.024			0.08	0.022		
per capita income (others)	-0.21	0.030			-0.19	0.028			-0.27	0.020		
married	0.04	0.028	0.14	0.027	0.02	0.029	0.13	0.025	0.03	0.019	0.12	0.016
North	0.07	0.047	0.07	0.041	0.05	0.048	0.08	0.040	-0.22	0.030	0.20	0.025
Southeast	0.19	0.033	0.34	0.031	0.13	0.034	0.31	0.029	-0.09	0.023	0.41	0.018
South	0.32	0.041	0.37	0.038	0.26	0.043	0.34	0.036	-0.07	0.029	0.52	0.024
Mid-West	0.03	0.041	0.08	0.040	0.02	0.043	0.10	0.038	-0.21	0.026	0.22	0.022
constructing	0.24	0.068	-1.11	0.044	0.49	0.070	-1.08	0.043	-0.13	0.032	-0.67	0.024
retail	-0.08	0.036	-0.37	0.038	-0.03	0.038	-0.33	0.037	-0.21	0.027	-0.18	0.022
lodging, food and services	-0.48	0.040	-1.10	0.036	-0.18	0.046	-1.03	0.035	-0.44	0.024	-0.75	0.020
Productive Sector	-0.13	0.032	-0.07	0.037	-0.12	0.034	-0.05	0.036	0.00	0.026	-0.10	0.022
Social Services	-0.27	0.040	-0.07	0.045	-0.26	0.042	-0.05	0.044	-0.09	0.032	-0.14	0.028
$\rho_{1,2}$					-0.80	0.029			-0.96	0.011		
Log L	-25059				-24979				-35326			
N	60138				60138				60138			

Former formal sector workers have the highest probability of being in the queue, followed by the base category "new entrants".[31] In contrast, all others "last occupation status" have a higher probability of being chosen from the queue than the "new entrants". These probabilities are particularly high for former public servants and former formal workers. These results suggest that employers prefer workers with any type of experience rather than the inexperienced "new entrants".

The estimates for the effect of gender and race are not significant either in the in the queue equation or in the chosen from the queue equation.[32] Married workers are not more likely to be in the queue than single ones. However, they are more likely to be chosen from it. Workers who live in metropolitan areas are more likely to both be in the queue and be chosen from the queue. Workers who live in the Northeast (base category) seems to be both less likely to be in the queue and less likely to be chosen from the queue. As for industry affiliation, manufacturing workers (base category) are more likely to be in the queue in relation to all sectors except for the constructing sector, and also more likely to be chosen from the queue than the other sectors.[33]

As for the variables that are only included in the "in the queue" equation the results are not much different from the ones of the univariate specification. Tenure increases the probability of being in the queue for formal jobs. The variables related to number of children per age group, if anything, tend to reduce the probability of being in the queue.[34] The presence of older people has a statistically significant negative effect. The number of other members in the household who are formal workers increases the probability of being in the queue, whereas the number of other informal members decreases it. The presence of self-employed also decreases that probability, whereas the number of unemployed has a positive but not significant effect. Finally, *per capita* income of other members of the household has a negative effect on the probability of being in the queue.

Poirier bivariate probit

Table 3 – part B, displays the results for Poirier's bivariate probit with partial observability and simultaneous decision. As most of the estimates are quite similar to the sequential decision specification in terms of sign and magnitude, we will only comment on the differences between the two models.

Some intermediate groups of education show a lower probability of being in the queue than the illiterate group (base category), whereas the probability of being chosen from the queue continues to grow almost monotonically with education. Male workers are neither more likely to be chosen from the queue nor more likely than female workers to queue for formal jobs. However, the sign of the coefficient for male workers in the "chosen from the queue" is positive, whereas it is negative in the "in the queue" equation. Similarly, the coefficients on race (white=1) continue to be

31. Surprisingly, former informal workers have the lowest probability of being in the queue.
32. It is worth noting that the sign of the estimates indicate a positive effect of being male and white on the probability of being chosen from the queue, and a negative effect on the the probability of being in the queue
33. Actually the coefficient for the productive sector and for the social services are not statistically significant at 5% despite being negative.
34. The only statistically significant coefficient in this group of variables is the one for the number of children between 10 and 12 years old, for all other categories the sign is negative, but never significant.

not significant either in the "in the queue" equation (negative) or in the "chosen from the queue" equation (positive).

The correlation between the two equations, "in the queue" and "chosen from the queue", is negative (–0.80) and statistically significant. This means that unobservables that make a worker more likely to be in the queue make him/her less likely to be chosen from the queue and vice-versa.

Bivariate probit with sample selection

The 1990 special supplement of the Brazilian household survey (Pnad) asked self-employed and informal (non-registered) workers whether or not they would like to switch to a formal (registered) job. Whereas roughly 70% of non-registered workers answered that they would like to, only 30% of the self-employed gave a positive answer. This rough statistics suggests that the desire for a formal job is much more common among informal (non-registered) workers than among self-employed. This fact supports the argument that our definition of informal sector is more attached to the labour market strictu sensu than the definitions that include self-employed, small employers and non-remunerated workers.

Table 4 displays some descriptive statistics for the two types of informal sector workers: potential switchers and potential non-switchers. Among the interesting differences between these two groups we observe that those who want to switch to the formal sector earn less than the ones who do not want to. They are also less educated, less experienced and have a lower amount of seniority. They are predominantly non-white and were also informal workers in their previous job. They have more children and live with other informal workers. The *per capita* income of the other members of the household is lower for them than for non-switchers.

The 1990 supplement also asked non-switchers about the reasons for staying in the informal sector. Only 9% argued that they would earn more as informal workers. Another 9% argued that they prefer an informal job because of either domestic work or time flexibility. The great majority 65% argued that they simply were satisfied with their current job. This response may in reality represent a mix of satisfaction with their wage and/other advantages offered by informal jobs such as time flexibility. A tiny proportion of 7% revealed that they would not like to move to a registered job because they do not have the necessary requirements to get a job in that sector. This self-selection answer suggests that this group of workers lost their interest in joining the queue after spending some time on it.

TABLE 4

Descriptive statistics for informal sector workers

	Switchers		Non-switchers	
	Mean	Std. Dev.	Mean	Std. Dev.
log hourly wage	-0.37	0.81	0.28	1.02
illiterate (base)	0.15	0.35	0.12	0.33
some elementary	0.23	0.42	0.18	0.38
complete elementary	0.44	0.50	0.39	0.49
complete primary	0.12	0.32	0.14	0.35
complete secondary	0.06	0.24	0.13	0.33
complete college	0.01	0.09	0.04	0.21
exp/10	1.57	1.21	2.06	1.48
exp/100	3.93	5.79	6.45	8.01
tenure	1.78	3.31	3.68	5.62
sex (male=1)	0.58	0.49	0.62	0.48
metropolitan	0.36	0.48	0.42	0.49
race (white=1)	0.39	0.49	0.53	0.50
new entrants (base)	0.31	0.46	0.31	0.46
more than 5 years	0.09	0.29	0.22	0.41
former formal worker	0.18	0.38	0.18	0.39
former informal worker	0.37	0.48	0.22	0.42
former public servant	0.00	0.05	0.01	0.08
former self-employed	0.05	0.22	0.06	0.24
#children 0-3 years	1.17	1.32	0.95	1.16
#children 4-6 years	0.34	0.60	0.27	0.53
#children7-10 years	0.55	0.78	0.44	0.71
#children 11-13 year	0.46	0.67	0.33	0.59
# elder	0.08	0.29	0.07	0.26
# formal	0.57	0.86	0.53	0.81
#self-employed	0.36	0.62	0.28	0.56
#informal	0.77	1.08	0.53	0.86
#unemployed	0.14	0.40	0.09	0.34
per capita income (others)	0.09	0.14	0.18	0.43
married	0.34	0.47	0.49	0.50
Northeast (base)	0.35	0.48	0.21	0.40
North	0.10	0.30	0.12	0.32
Southeast	0.29	0.46	0.35	0.48
South	0.10	0.30	0.15	0.35
Mid-West	0.16	0.37	0.18	0.38
Manufacturing (base)	0.17	0.37	0.18	0.39
constructing	0.15	0.36	0.09	0.29
retail	0.12	0.33	0.15	0.36
lodging, food and personal services	0.39	0.49	0.33	0.47
Productive Sector	0.11	0.31	0.15	0.36
Social Services	0.06	0.24	0.09	0.29
N	12364		4478	

In order to assess our previous results of the reduced-form bivariate probit with partial observability, we estimate a bivariate probit with sample selection exploiting the additional information provided by the "would you like to switch to a formal job?" question. We assume that workers who answered positively this question and workers who are in the formal sector are in the queue. Workers who answered the question negatively are assumed not to be in the queue. In doing that we get rid of the severe partial observability and end up with a censored dependent variable.

The results of the bivariate probit with sample selection reported in table 3 – part C are not very different from the previous ones. This is reassuring because these specifications have different dependent variables. In fact, the most remarkable

differences refer to the significance of the estimated parameters for race and gender in both equations, and to the sign of the variables related to "last occupation status" in both equations "in the queue" and "chosen from the queue".

Female and non-white workers are more likely to be in the queue, but are less likely to be chosen from the queue. These results, differently from the previous specifications, are statistically significant and point to the existence of gender and race "discrimination" in hiring policies in the formal sector. Workers in the Northeast now are more likely to be in the queue than workers from any other region, but less likely to be chosen from the queue.

"New entrants" are the least likely category to be in the queue for formal jobs. However, they are not less likely to be chosen from the queue than informal workers. The latter seems to display some characteristics that make them less attractive to formal employers, even when compared to inexperienced workers. These results suggest that accepting an informal job may worsen the individual chances to get a formal job in the next period conditional on all the variables included in the model. Another different result is the fact that this group is no longer the least likely to be "in the queue", according to this specification, they are only less likely than former formal workers to be "in the queue".

Similarly to the bivariate probit with simultaneous decision, the correlation between the selection equation ("in the queue equation") and the chosen from the queue equation is negative, -0.96, meaning that unobservables that make workers more likely to be in the queue make them less likely to be chosen from the queue.

5.3 JOB QUEUE TESTS AND THE LENGTH OF THE QUEUE

Applying Abowd and Farber (1982) likelihood ratio test (LR) to the hypothesis of a universal queue and to the hypothesis of no queue, we are able to reject both models in favour of a complete bivariate specification in the three specifications that we adopt. As the univariate probit model is not nested in the bivariate specification, we follow the Abowd and Farber (1982) procedure and test those hypotheses by imposing restrictions on the bivariate specifications. For the non-queue model, we impose the restriction that all variables in the "chosen from the queue" equation are zero with the exception of a positive constant that is sufficiently high to make all workers eligible for a registered job. Similarly, when testing the universal queue hypothesis, we impose the restriction that all variables in the "in the queue" equation are zero with the exception of a positive constant. Table 5 shows that for all specifications we are able to reject both the universal queue hypothesis and the no queue hypothesis. These results suggest that the bivariate specification is the most appropriate procedure to describe the sector allocation process.

TABLE 5
LR test for universal queue and no queue hypothesis

	Abowd-farber		Poirier		Sample selection	
	Universal queue	No queue	Universal queue	No queue	Universal queue	No queue
LR	5081.34	3510.06	5243.27	3672	19470.04	35778.08
p-value	0	0	0	0	0	0
critical value	50.6	37.65	50.6	37.65	50.6	37.65

Table 6 reports the average probability of being in the formal sector (PF), of being in the queue for a formal job (PIQ) and of being chosen from the queue conditional on being in the queue (PCF) that we get from the three bivariate models described above for some selected characteristics. It also contains an estimate of the length of the queue (q) and the probability of being a formal worker given by the univariate probit. The first interesting thing to note is that the probability of being chosen from the queue, in general, is lower in the bivariate probit with sample selection (censored probit) than in both sequential and simultaneous bivariate probit with partial observability. The only exceptions are workers with at least college education and workers who are in the current job for more than 5 years. For these workers the probability of being chosen from the queue in the specification of the bivariate probit with sample selection is higher than in the other partial observability specifications. Another relevant difference among the results yielded by different specifications is the extremely low probability of being chosen from the queue (conditional on being in the queue) for former informal workers in the bivariate probit with sample selection. This category presents the highest estimate of the length of the queue, 2.11, meaning that for each worker in the formal sector with such characteristics there are 2.11 workers willing to get a formal job.[35]

Overall, the different specifications yield an estimate of the length of the queue that varies from 1.20 (Abowd-Farber bivariate probit) to 1.30 (bivariate probit with sample selection). These estimates mean that for each worker in the formal sector there are something around 1.20 or 1.30 workers queuing for formal jobs. However, the interesting point of this analysis is to evaluate how the probability of being chosen from the queue (conditional on being in the queue) varies according to different characteristics. In this regard, female, non-white, former informal workers, "new entrants" and workers with low schooling are the groups who once "in the queue" have the lowest probability of being chosen from it. As pointed out above, being a former informal worker seems to be the most damaging handicap that one can have in the labour market. A spell in an informal job acts as a "scar" for the workers who experience it, signalling some characteristics that are not valued by formal employers. The fact that the probability of being chosen from the queue is even lower than for "new entrants" is quite revealing.

35. Note that this result should be expected given the different signs of the variable "last occupation as informal worker" in the "in the queue" equation and in the "chosen from the queue" equation in the bivariate probit with sample selection when compared to the other specifications.

TABLE 6

Probabilites of a) being chosen form the queue conditional on being in the queue (PCF), b) being in the queue (PIQ), c) being in the formal sector; and length of the queue (Q)

	Abowd-farber		Poirier		Censored		Univariate	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
ALL								
PCF	0.84	0.20	0.80	0.24	0.77	0.24		
PIQ	0.84	0.20	0.87	0.16	0.92	0.06		
PF	0.72	0.26	0.72	0.26	0.72	0.24	0.72	0.25
Q	1.19		1.25		1.30			
MEN								
PCF	0.85	0.19	0.82	0.22	0.79	0.22		
PIQ	0.85	0.19	0.88	0.15	0.93	0.05		
PF	0.74	0.24	0.74	0.24	0.74	0.22	0.74	0.23
Q	1.18		1.22		1.26			
WOMEN								
PCF	0.82	0.21	0.77	0.27	0.73	0.27		
PIQ	0.81	0.23	0.85	0.17	0.92	0.07		
PF	0.68	0.28	0.68	0.29	0.68	0.27	0.69	0.27
Q	1.23		1.30		1.37			
WHITE								
PCF	0.88	0.17	0.85	0.20	0.83	0.20		
PIQ	0.87	0.18	0.89	0.14	0.92	0.06		
PF	0.78	0.23	0.78	0.23	0.78	0.21	0.78	0.22
Q	1.14		1.17		1.20			
NON-WHITE								
PCF	0.79	0.22	0.74	0.26	0.70	0.26		
PIQ	0.80	0.22	0.85	0.17	0.92	0.06		
PF	0.65	0.27	0.66	0.28	0.65	0.26	0.66	0.26
Q	1.27		1.34		1.43			
ENTRANTS								
PCF	0.72	0.24	0.68	0.27	0.69	0.25		
PIQ	0.86	0.18	0.91	0.12	0.90	0.07		
PF	0.64	0.27	0.64	0.28	0.63	0.24	0.64	0.25
Q	1.40		1.46		1.45			
+ 5 YEARS								
PCF	0.89	0.14	0.89	0.15	0.92	0.10		
PIQ	0.96	0.12	0.97	0.09	0.94	0.05		
PF	0.86	0.18	0.86	0.17	0.86	0.12	0.86	0.15
Q	1.12		1.13		1.09			
Former formal								
PCF	0.95	0.08	0.93	0.10	0.87	0.12		
PIQ	0.88	0.11	0.90	0.09	0.95	0.03		
PF	0.83	0.14	0.83	0.13	0.83	0.13	0.83	0.14
Q	1.06		1.08		1.15			
Former informal								
PCF	0.71	0.21	0.60	0.28	0.47	0.25		
PIQ	0.57	0.22	0.67	0.18	0.90	0.06		
PF	0.43	0.23	0.43	0.25	0.43	0.24	0.43	0.22
Q	1.42		1.67		2.11			
illiterate								
PCF	0.68	0.22	0.61	0.26	0.58	0.26		
PIQ	0.74	0.26	0.81	0.20	0.88	0.08		
PF	0.52	0.26	0.52	0.27	0.52	0.25	0.52	0.25
Q	1.46		1.63		1.74			
some elementary								
PCF	0.73	0.23	0.67	0.28	0.63	0.28		
PIQ	0.76	0.24	0.83	0.19	0.90	0.06		
PF	0.58	0.28	0.58	0.29	0.58	0.26	0.58	0.27
Q	1.36		1.48		1.58			

cont.

cont.

	Abowd-farber		Poirier		Censored		Univariate	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
complete elementary								
PCF	0.80	0.20	0.76	0.25	0.72	0.24		
PIQ	0.81	0.21	0.86	0.16	0.92	0.06		
PF	0.67	0.26	0.67	0.26	0.67	0.24	0.67	0.25
Q	1.25		1.32		1.39			
complete primary								
PCF	0.90	0.13	0.87	0.15	0.83	0.16		
PIQ	0.87	0.16	0.89	0.13	0.94	0.05		
PF	0.78	0.20	0.78	0.19	0.78	0.17	0.78	0.19
Q	1.11		1.14		1.20			
complete secondary								
PCF	0.96	0.06	0.95	0.07	0.93	0.08		
PIQ	0.92	0.12	0.93	0.11	0.95	0.04		
PF	0.88	0.14	0.89	0.13	0.88	0.09	0.88	0.12
Q	1.04		1.05		1.08			
complete college								
PCF	0.97	0.04	0.97	0.05	0.97	0.04		
PIQ	0.95	0.10	0.95	0.09	0.95	0.05		
PF	0.92	0.11	0.92	0.10	0.92	0.07	0.92	0.09
Q	1.03		1.03		1.03			

5.4 WAGE EQUATIONS FOR FORMAL AND INFORMAL WORKERS

Univariate-based corrections

The wage equations corrected for selectivity using the univariate reduced-form probit either by the two-step OLS procedure or by the maximum likelihood procedure yield very similar results.[36] Both results shown in table 7, part A and B – are quite robust in showing the existence of selectivity in both the formal sector and the informal sector and in indicating an overestimation of most parameters in the standard OLS procedure (table 7, part C). In the selectivity-corrected OLS procedure, the negative and statistically significant coefficient (-0.20) of the Inverse Mills ratio for the formal sector equation suggests that individuals who are less likely to be registered workers benefit more of this status than the ones who are more likely to be registered. There is a negative correlation (-0.32) between the residual of the probit specification and the residual of the wage equation. This result does not support the hypothesis that individuals allocate themselves into the sectors where they have comparative advantage. Similar results are obtained in the ML approach, since the correlation between the residuals of the selection equation and the wage equation is also negative and statistically significant, -0.53. Tannuri-Pianto and Pianto (2002) find similar results for Brazil using a semiparametric approach. They argue that the negative effect of the selectivity parameters on wages indicates that workers choose to join the formal sector due to non-wage benefits associated with this kind of job that are not captured by the wage equation. The shortcoming of this sort of explanation is that it assumes that every worker who prefers a formal to an informal job is able to get it, i.e, there is no queue for formal jobs. Another important fact that one should not overlook is

36. The selection equation in the ML estimation also do not differ much from the probit estimation of the two-step procedure (see table 2), only the coefficients on race and on per capita income display a different sign. However, only the latter is statistically significant in both equations.

that workers who are less likely to be formal workers are the ones who tend to earn relatively more once in the formal sector.

As for the wage equation of the three-step OLS procedure for the informal sector the coefficient of the Inverse Mills ratio is negative and significant, –0.21. This result indicates the presence of selectivity in the informal sector. The workers who are less likely to be in the formal sector are the ones who command higher than expected wages in the informal sector. The residuals of the selection equation (probability of choosing the formal sector) and the residuals of the wage equation are negatively correlated, –0.31.[37] One can argue that this result suggests that informal workers choose to join the informal sector due to comparative advantages in that sector (Carneiro and Henley, 2001).[38] However, in order to conclude that workers select themselves in the sector where they are have comparative advantage, we should also find a similar result for the formal sector equation, which is not the case.[39]

TABLE 7

Wage equations (univariate-based corrections)

	Selectivity corrected – OLS (A)				OLS (C)			
	Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.96	0.025	-1.80	0.028	-1.14	0.019	-1.76	0.027
some elementary	0.14	0.015	0.15	0.019	0.16	0.015	0.17	0.018
complete elementary	0.32	0.015	0.36	0.019	0.36	0.014	0.42	0.018
complete primary	0.64	0.017	0.64	0.026	0.70	0.016	0.75	0.023
complete secondary	1.12	0.018	1.07	0.031	1.20	0.016	1.23	0.027
complete college	1.93	0.020	1.89	0.047	2.01	0.019	2.05	0.044
exp/10	0.30	0.010	0.36	0.017	0.33	0.010	0.43	0.016
exp/100	-0.05	0.002	-0.05	0.003	-0.06	0.002	-0.06	0.003
tenure	0.03	0.001	0.01	0.002	0.03	0.001	0.02	0.001
sex (male=1)	0.30	0.007	0.38	0.012	0.30	0.007	0.39	0.012
metropolitan	0.09	0.006	0.19	0.012	0.11	0.006	0.23	0.011
race (white=1)	0.16	0.007	0.11	0.012	0.16	0.007	0.11	0.012
married	0.17	0.007	0.21	0.014	0.18	0.007	0.23	0.013
North	0.42	0.012	0.48	0.019	0.43	0.012	0.49	0.019
Southeast	0.28	0.009	0.31	0.015	0.31	0.008	0.36	0.014
South	0.27	0.011	0.35	0.021	0.31	0.010	0.42	0.020
Mid-West	0.41	0.012	0.46	0.017	0.41	0.011	0.47	0.016
constructing	-0.01	0.014	0.14	0.021	-0.06	0.013	0.06	0.019
retail	-0.21	0.010	-0.02	0.020	-0.23	0.010	-0.05	0.019
lodging, food and services	-0.22	0.013	-0.13	0.020	-0.31	0.011	-0.25	0.016
Productive Sector	0.02	0.008	0.10	0.020	0.02	0.008	0.10	0.020
Social Services	-0.16	0.011	-0.06	0.025	-0.18	0.011	-0.08	0.025
Inverse Mills ratio (lambda)	-0.20	0.018	-0.21	0.022				
R2	0.56		0.46		0.55		0.46	

cont.

37. The ML procedure also gives similar results for the correlation between the residual of the selection equation and the residual of the wage equation for the informal sector, the correlation is negative and statistically significant, –0.33.

38. Carneiro and Henley (2001) estimate an endogenous switching regression model for formal and informal sector using the Brazilian annual household survey for 1999 and find similar results in terms of sign and significance of the selectivity terms. However, they focus their analysis on the coefficient of the Inverse Mills ratio for the informal sector wage equation, and overlook the implication of the sign of the coefficient of the Inverse Mills ratio for the formal sector.

39. See Yamada (1996) and Maddala (1983) for a discussion of this point.

cont.

	MLE (B)			
	Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.83	0.023	-1.80	0.027
some elementary	0.13	0.017	0.15	0.019
complete elementary	0.30	0.016	0.36	0.019
complete primary	0.60	0.018	0.63	0.025
complete secondary	1.06	0.018	1.05	0.030
complete college	1.88	0.020	1.86	0.040
exp/10	0.27	0.010	0.36	0.017
exp/100	-0.05	0.002	-0.05	0.003
tenure	0.02	0.001	0.01	0.001
sex (male=1)	0.30	0.007	0.38	0.012
metropolitan	0.08	0.006	0.18	0.012
race (white=1)	0.16	0.007	0.11	0.012
married	0.16	0.007	0.21	0.013
North	0.42	0.011	0.47	0.018
Southeast	0.26	0.009	0.30	0.015
South	0.25	0.011	0.35	0.022
Mid-West	0.41	0.011	0.46	0.017
constructing	0.03	0.014	0.14	0.022
retail	-0.20	0.010	-0.03	0.019
lodging, food and services	-0.16	0.012	-0.13	0.020
Productive Sector	0.03	0.008	0.10	0.019
Social Services	-0.15	0.011	-0.06	0.024
$\rho_{1,u}$	-0.53	0.015	-0.33	0.028

The estimates of the effect of human capital variables – groups of education, experience, experience squared, tenure – in the formal sector equation are overestimated in the non-corrected OLS. The ML equation estimates show a somewhat higher bias in the OLS procedure than the two-step procedure. The estimates for the informal sector are also overestimated, but now in some cases, the two-step procedure estimates are even lower than their ML correspondents.

It is worth noting that even the selectivity-corrected estimates show strong returns to education for both formal and informal workers. Thus, it is not correct to treat the 'informal' sector as a secondary sector in which there is no return to education.[40] However, it seems that tenure is much more rewarded in the formal than in the informal sector. The results also show the existence of a wage premium for married workers.

Bivariate-based corrections

The wage equations corrected by the double selectivity criteria, for both formal and informal workers also show that the non-corrected OLS procedure renders overestimated estimates for most coefficients in the case of sequential decision models, but not in the simultaneous decision model. In the case where the correction is based on the Abowd-Farber bivariate probit with partial observability and sequential decision (table 8, part A) all estimates of selectivity correction parameters are negative and significantly different from zero.[41] Therefore, both the workers less

40. Similar results are found for different Latin American countries. See Yamada (1996) and Saavedra and Chong (1999).

41. $\sigma_{1f} = -0.19$ and $\sigma_{2f} = -0.07$ in the formal sector wage equation and $\sigma_{1i} = -0.18$ and $\sigma_{2i} = -0.18$ in the informal sector wage equation.

likely to be in the queue and the workers less likely to be chosen from the queue are the ones who tend to earn more than would be expected, given their productive characteristics, in the formal sector. Differently, the informal workers who are less likely to be in the queue for formal jobs and those who are less likely to be chosen from the queue are the ones who benefit more from the informal job. As most results for the estimated wage equations remain unchanged we avoid repetition and analyse only the result of the job queue test suggested in Mengistae (1998). The hypothesis of no job queue ($\delta_1 = \delta_2 = 0$) is rejected at a standard 5% level of significance. The calculated F statistics is F (1, 16791) = 77.27 with a p-value of 0.000. On the other hand, we cannot reject the hypothesis of a universal queue for formal jobs $\pi = 1$. The F statistic for the restriction that $\sigma_{1i} = \delta_1$ is 0.009 with a p-value of 0.9754.[42]

As for the correction based on the simultaneous model (Poirier bivariate probit) we can only estimate a meaningful equation for the formal sector (see Section 3). The estimates of the human capital variables (table 8, part B), if anything, seem to be underestimated in the non-corrected OLS regression. The selectivity coefficient for the "chosen from the queue equation", indicates that workers more likely to be chosen from the queue are the ones who earn more in the formal sector, whereas the selectivity coefficient for the "in the queue" equation indicates that workers more likely to be in the queue earn less than would be expected given their productive characteristics. Note that this result is quite different from the one yielded by the Abowd-Farber bivariate probit with partial observability and sequential decision model. The fact that the difference occurs in the "chosen from the queue" equation is not surprising since this probability is defined over different populations in these two models as shown in Section 3.

TABLE 8

Wage equations (bivariate-based corrections)

	Abowd-faber (A)				Poirier (B)	
	Formal		Informal		Formal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-1.02	0.025	-1.69	0.037	-1.14	0.024
some elementary	0.15	0.015	0.14	0.019	0.16	0.016
complete elementary	0.34	0.015	0.33	0.020	0.37	0.015
complete primary	0.67	0.018	0.57	0.029	0.72	0.018
complete secondary	1.16	0.019	0.97	0.037	1.22	0.019
complete college	1.97	0.021	1.81	0.051	2.03	0.021
exp/10	0.31	0.010	0.32	0.019	0.34	0.010
exp/100	-0.05	0.002	-0.05	0.003	-0.06	0.002
tenure	0.02	0.001	0.01	0.004	0.02	0.001
sex (male=1)	0.30	0.007	0.36	0.012	0.30	0.007
metropolitan	0.10	0.006	0.16	0.012	0.11	0.006
race (white=1)	0.16	0.007	0.11	0.012	0.16	0.007
married	0.17	0.007	0.20	0.014	0.18	0.007
North	0.42	0.012	0.47	0.019	0.43	0.012
Southeast	0.29	0.009	0.28	0.016	0.31	0.009
South	0.28	0.011	0.32	0.021	0.30	0.011

cont.

42. As Mengistae (1998) points out, his methodology for testing the job queue using the unconditional wage equation for workers who were not chosen from the queue or were not in the queue is completely independent of Abowd and Farber (1982) test for job queue. In our case, it seems that both methods point to the rejection of the "in the queue" univariate specification for the correction of the wage equations. However, they yield different results in relation to the existence of universal queue.

cont.

	Abowd-faber (A)				Poirier (B)	
	Formal		Informal		Formal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Mid-West	0.41	0.012	0.45	0.017	0.42	0.012
constructing	-0.04	0.015	0.19	0.026	-0.09	0.016
retail	-0.22	0.010	0.00	0.020	-0.23	0.010
lodging, food and services	-0.25	0.014	-0.06	0.023	-0.32	0.014
Productive Sector	0.02	0.008	0.11	0.020	0.02	0.008
Social Services	-0.17	0.011	-0.05	0.024	-0.17	0.011
σ1f	-0.19	0.016			-0.15	0.017
σ2f	-0.07	0.025			0.08	0.023
σ1i			-0.18	0.018		
δ_1			-0.18	0.021		
δ_2			-0.18	0.021		
R2	0.56		0.47		0.56	
	BP with sample selection (C)					
	Formal		Informal (IQ)		Informal (NIQ)	
Constant	-0.99	0.029	-1.62	0.031	-1.62	0.115
some elementary	0.14	0.016	0.11	0.021	0.21	0.042
complete elementary	0.32	0.016	0.26	0.023	0.50	0.040
complete primary	0.64	0.019	0.47	0.033	0.92	0.049
complete secondary	1.11	0.020	0.81	0.044	1.37	0.052
complete college	1.91	0.023	1.64	0.076	2.06	0.069
exp/10	0.31	0.010	0.28	0.022	0.47	0.032
exp/100	-0.06	0.002	-0.04	0.004	-0.07	0.006
tenure	0.03	0.001	0.01	0.002	0.01	0.002
sex (male=1)	0.29	0.007	0.35	0.014	0.35	0.025
metropolitan	0.10	0.007	0.19	0.014	0.18	0.025
race (white=1)	0.15	0.007	0.07	0.013	0.16	0.025
married	0.17	0.007	0.17	0.015	0.30	0.026
North	0.40	0.013	0.40	0.021	0.53	0.042
Southeast	0.28	0.009	0.30	0.018	0.23	0.033
South	0.27	0.011	0.34	0.025	0.28	0.041
Mid-West	0.38	0.012	0.47	0.019	0.33	0.038
constructing	-0.02	0.015	0.15	0.025	0.13	0.046
retail	-0.23	0.010	-0.08	0.021	0.04	0.040
lodging, food and services	-0.29	0.013	-0.15	0.024	-0.18	0.038
Productive Sector	0.02	0.009	0.06	0.022	0.20	0.039
Social Services	-0.18	0.011	-0.02	0.027	-0.15	0.047
σ1f	0.31	0.037				
σ2f	-0.24	0.030				
σ1i*			-0.89	0.123		
σ1i					-0.01	0.054
σ2i			-0.17	0.030		
R2	0.56		0.40		0.48	

The results for the Bivariate Probit with sample selection (table 8, part C) indicates that the results from the non-corrected OLS regressions are overestimated. The three different wage regressions derived from this approach reveal a somewhat different pattern than the ones seen so far. In the case of the formal sector, the positive and significant coefficient on the Inverse Mills ratio from the "in the queue equation", σ_{1f}, suggests that workers more likely to be in the queue have higher wages than expected in the formal sector. Differently, the coefficient on the Inverse Mills ratio from the "chosen from the queue equation", σ_{2f}, suggests that workers more likely to be chosen from the queue earn less than expected. The result for the correction based on the "in the queue" equation differs from the one based on the

Abowd-Farber bivariate probit and the result for the correction based on the "in the queue" equation differs from the Poirier bivariate probit.

As for the informal workers we have two different equations. One for the workers who are in the queue and were not chosen from it and one for the workers who were not in the queue. The results for the former group suggest that workers more likely to be in the queue earn lower wages in the informal sector,[43] whereas workers less likely to be chosen from the queue are the ones who earn more in the informal sector.[44] This result suggests that informal workers who are "in the queue" for formal jobs tend to have lower wages than expected.

As for the equation for the workers who are not in the queue there is no evidence of the presence of selectivity, the sign of the coefficient on the Inverse Mills ratio from the "in the queue equation" is negative, but not statistically significant, $\sigma_{1i} = -0.01)$.

In short, as can be seen in table 9,[45] different specifications of the sector allocation process imply different results in terms of the effect of selectivity on earnings. Assuming that the correct model is one with sequential decision (either Abowd-Farber bivariate probit or bivariate probit with sample selection) we can conclude that workers more likely to be "chosen from the queue" are the ones who will benefit less (relatively) from a formal job. This may be indicating some cost minimizing behavior of the employers. The results regarding the "in the queue" equation, however, differ according to the sequential specification adopted.

For the formal sector, both sequential specifications indicate a negative effect of the "chosen from the queue" status on the wage equation. The "in the queue" equation, however, displays a negative effect in the Abowd-Farber specification, and a positive one in the bivariate probit with partial observability.

As for the informal sector, the Abowd-Farber specification, indicates that the "in the queue" status has a positive impact on earnings, whereas the bivariate probit with sample selection shows a negative impact on the wage for informal workers who are "in the queue" and a positive, but not statistically significant effect for informal workers who are not it the queue.[46]

TABLE 9

Effect of the correlation between unobservables in the choice equations and in the outcome equation

	Predicted wages					
	Abowd-faber		Poirier	Bivariate probit with sample selection		
	Formal	Informal	Formal	Formal	Informal in the queue	Informal not in the queue
More likely to be in the queue	lower	lower	lower	higher	lower	not significant
More likely to be chosen from the queue	lower	lower	higher	lower	higher	

43. $\sigma_{1i}^{*} = -0.89$

44. $\sigma_{2i} = -0.17$

45. Table 8 highlights how unobservables of the two sector-selection equations affect wages of the formal and informal sector.

46. Note that we can only make this distinction between informal workers in the queue and informal workers not in the queue when using the correction based on the bivariate probit with sample selection.

5.5 THE ROLE OF WAGE DIFFERENTIAL ON SECTOR ALLOCATION

Univariate specification

The structural probit is estimated including the wage differential as an additional regressor and excluding the variables that are assumed to determine only the wage equation (industry dummies). The wage differential is computed as the difference between the offered wage in the formal sector minus the offered wage in the informal sector. The offered wage differs from the conditional wage because it does not include the Inverse Mills ratio in the calculation of the predicted (log) wage.[47] This is so, because we are interested in the wage that any member of the population of employees would get, regardless if he/she is actually working in the informal or in the formal sector. Table 9 shows both the results based on the three-step OLS and the result based on the maximum likelihood procedure for the univariate case (probit). The results of the OLS-based procedure suggest that the wage differential is the most important variable determining the "choice" of formal sector by workers. The coefficient is positive and statistically significant. The marginal effect is also positive and statistically significant $0.55\ (0.030)$. The marginal effect indicates that the wage differential is the most important variable determining sector allocation in the univariate model.[48] Nevertheless the other variables of the model continue to be significant and the only noticeable change relative to the reduced form specification is that the coefficient on race is negative and significant, whereas in the reduced form it was positive and statistically insignificant.[49]

TABLE 10
Structural univariate probit

	OLS		MLE	
	Coeff.	Std. Err.	Coeff.	Std. Err.
$(W_f - W_i)$	1.83	0.100	0.12	0.107
Constant	-2.33	0.083	-1.07	0.104
some elementary	0.12	0.025	0.11	0.025
complete elementary	0.38	0.024	0.33	0.025
complete primary	0.71	0.029	0.70	0.029
complete secondary	0.93	0.031	1.01	0.031
complete college	0.97	0.043	1.05	0.043
exp/10	0.45	0.021	0.35	0.022
exp/100	-0.06	0.004	-0.06	0.004
tenure	0.03	0.002	0.05	0.002
sex (male=1)	0.26	0.016	0.13	0.016
metropolitan	0.33	0.016	0.17	0.017

cont.

47. The unconditional mean formal sector wage premium varies a lot according to the specification adopted: 22% in the OLS specification, 73% in the univariate probit-corrected specification, 105% in the Abowd-Farber bivariate probit-corrected specification; and 93% in relation to informal workers who want to switch to formal sector and -0.3% in relation to informal workers who do not want to switch, according to the bivariate probit with sample selection-corrected specifications. Thus it seems that the OLS specification underestimates the mean wage premium for formal sector workers. However, this mean wage premium is different for different type of informal workers as revealed by the bivariate probit with sample selection in the expected direction.

48. If instead of using offered wages, we use conditional wages to calculate the wage differential we get very similar results, but the effect of the wage differential becomes even larger. These results are available upon request.

49. The coefficient on gender remains positive and turns out to be significant, indicating that male workers are more likely to be found in the formal sector.

cont.

	OLS		MLE	
	Coeff.	Std. Err.	Coeff.	Std. Err.
race (white=1)	-0.07	0.015	0.03	0.015
more than 5 years	0.33	0.021	0.33	0.021
former formal worker	0.59	0.019	0.59	0.019
former informal worker	-0.20	0.019	-0.20	0.019
former public servant	0.27	0.087	0.27	0.087
former self-employed	0.06	0.032	0.05	0.032
#children 0-3 years	-0.04	0.009	-0.04	0.009
#children 4-6 years	0.02	0.016	0.02	0.016
#children7-10 years	-0.02	0.011	-0.02	0.011
#children 11-13 year	-0.02	0.011	-0.02	0.011
# elder	-0.04	0.023	-0.04	0.023
# formal	0.18	0.007	0.19	0.007
#self-employed	-0.07	0.011	-0.07	0.011
#informal	-0.19	0.008	-0.19	0.008
#unemployed	0.02	0.016	0.02	0.016
per capita income (others)	-0.13	0.024	-0.13	0.024
married	0.23	0.016	0.17	0.017
North	0.18	0.024	0.09	0.024
Southeast	0.31	0.017	0.27	0.017
South	0.49	0.023	0.36	0.024
Mid-West	0.11	0.022	0.02	0.022
Log L	-27347.97		-27515.92	
N	60138		60138	

The results based on the ML procedure, however, show a smaller and statistically non-significant positive effect of the wage differential on the probability of choosing the formal sector. The marginal effect is also small and not significant, $0.04(0.032)$ As for the other variables, only the coefficient on race shows a different sign from the OLS-based structural probit. In this specification, white workers have a higher probability of being found in the formal sector than non-white workers, but neither the coefficient nor the marginal effects are statistically significant.

Bivariate specifications

In the case of the structural bivariate probit we exclude the variables exclusive to the wage equation (i.e. industry dummies) from the "in the queue" equation, and include the "wage in the formal sector conditional on being in the queue", $(W_f \mid IQ = 1)$, as an additional regressor in the "chosen from the queue" equation. The results based on the Abowd-Farber structural bivariate probit with partial observability in table 10 (panel A) reveal that for the "in the queue" equation the wage differential plays the most important role in determining the "in the queue" status, its coefficient is positive and significant.[50] The variables related to education that were statistically insignificant in the reduced form, turned out to be negative and significant in some cases. The coefficients on experience and experience squared turn out to be statistically insignificant. The other differences refer to the fact that male workers are now more likely to be "in the queue". The other coefficients did not change much in relation to the previous reduced form results.

As for the "chosen from the queue" equation the results change a lot in relation to the reduced form. The coefficient on the "formal sector wage" conditional on

50. Its marginal effect is also positive and significant, 0.45 (0.01).

being in the queue is positive and significant.[51] However, one should expect that if formal sector employers were cost minimizers, the sign of this coefficient should be negative. The higher the wage of a worker conditional on other productive attributes, the less likely the employer would be to hire him/her. Besides this awkward result, there were major changes in the sign of other variables such as education groups, sex and race, in most cases in unexpected ways.[52] This fact led us to believe that the "formal sector wage" variable is picking up the effect of human capital-related variables, instead of isolating the effect that a cost-minimization decision would have in the hiring process. Then, we re-estimate the bivariate probit excluding the "formal sector wage" from the "chosen from the queue" equation. The results in table 10 (panel B) are now much more similar to the ones we get from the reduced form of the "chosen from the queue" equation. Moreover, the "in the queue" equation does not change much. The coefficient of the wage differential is positive and significant,[53] and despite small changes in magnitudes, the sign of all parameters remain the same as in the previous reduced-form specification.

TABLE 11

Abowd-farber bivariate probit specification

	Abowd-farber (A)				Abowd-farber (B)			
	IQ		CFQ		IQ		CFQ	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
$(W_f - W_i)$	1.67	0.185			2.72	0.110		
W formal			1.68	0.043				
Constant	-0.95	0.154	1.27	0.067	-1.60	0.079	-0.89	0.069
some elementary	-0.17	0.090	-0.07	0.034	-0.14	0.046	0.34	0.049
complete elementary	-0.15	0.087	-0.17	0.035	-0.10	0.047	0.71	0.046
complete primary	-0.19	0.098	-0.35	0.049	-0.12	0.060	1.22	0.055
complete secondary	-0.29	0.107	-0.89	0.066	-0.22	0.070	1.65	0.056
complete college	-0.19	0.137	-2.31	0.103	-0.20	0.080	1.65	0.070
exp/10	0.06	0.062	-0.30	0.031	0.17	0.031	0.40	0.040
exp/100	0.00	0.013	0.05	0.005	-0.03	0.006	-0.05	0.007
tenure	1.31	0.063			0.27	0.008		
sex (male=1)	0.37	0.041	-0.60	0.025	0.33	0.022	-0.03	0.028
metropolitan	0.27	0.040	0.03	0.018	0.36	0.021	0.18	0.027
race (white=1)	-0.12	0.042	-0.24	0.020	-0.17	0.022	0.07	0.029
more than 5 years	-0.25	0.110	0.29	0.023	-0.40	0.048	0.34	0.033
former formal worker	0.26	0.065	0.82	0.026	0.31	0.038	1.37	0.123
former informal worker	-0.41	0.069	0.20	0.031	-0.43	0.045	0.60	0.106
former public servant	0.57	0.405	0.33	0.105	-0.01	0.123	0.98	0.494
former self-employed	0.15	0.111	0.26	0.046	-0.15	0.059	0.77	0.176
#children 0-3 years	-0.01	0.024			-0.03	0.012		
#children 4-6 years	-0.02	0.041			0.00	0.021		
#children7-10 years	-0.06	0.029			-0.03	0.016		
#children 11-13 year	-0.01	0.028			-0.02	0.015		
# elder	-0.12	0.060			-0.05	0.033		
# formal	0.27	0.022			0.21	0.011		
#self-employed	-0.05	0.030			-0.07	0.016		
#informal	-0.24	0.017			-0.22	0.011		
#unemployed	0.04	0.041			0.03	0.022		

cont.

51. Its marginal effect is also positive and significant, 0.423 (0.001).

52. According to the results of this specification, female and non-white workers would be more likely to be chosen from the queue than male and white workers. Similarly, workers with complete secondary education or more would be less likely to be chosen from the queue.

53. Its marginal effect now is still positive and significant, 0.30 (0.016).

cont.

	Abowd-farber (A)				Abowd-farber (B)			
	IQ		CFQ		IQ		CFQ	
per capita income (others)	-0.20	0.045			-0.17	0.029		
married	0.08	0.045	-0.15	0.021	0.15	0.023	0.18	0.030
	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.	Coeff.	Std.Err.
North	0.17	0.079	-0.64	0.035	0.15	0.038	0.14	0.046
Southeast	0.03	0.052	-0.16	0.026	0.13	0.028	0.38	0.035
South	0.28	0.065	-0.10	0.031	0.36	0.033	0.43	0.044
Mid-West	0.12	0.069	-0.66	0.033	0.13	0.035	0.01	0.045
Log L	-26018				-26229			
N	60138							

The results of the structural bivariate probit with sample selection[54] (table 11) show a strong and statistically significant effect of the wage differential on the probability of being in the queue, and also a small, but significant effect of the "formal sector wage" on the probability of being chosen from the queue. The marginal effect of wage differential in the "in the queue" equation is positive (0.018) and significant $(z = 2.00)$, the marginal effect of the wage in the formal sector in the "chosen from the queue" equation is also positive and significant, but very small (0.003).[55] In this latter case, the coefficients on the other variables do not seem to be much affected by the inclusion of this variable. Education, for instance, continues to have a positive effect on the probability of being chosen from the queue and a negative effect on the probability of queuing for formal jobs. The major differences are in the "in the queue" equation, where the coefficient on tenure turns out to be negative and significant,[56] and the coefficient on sex becomes positive and significant.

TABLE 12

Bivariate probit with sample selection structural bivariate probit specification

	BP with sample selection (A)				BP with sample selection (B)			
	IQ		CFQ		IQ		CFQ	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
$(W_f - W_i)$	19.38	0.467			20.52	0.493		
W formal			0.08	0.039				
Constant	-8.95	0.259	-1.18	0.047	-9.28	0.268	-1.24	0.036
some elementary	-0.09	0.110	0.19	0.028	-0.16	0.111	0.19	0.028
complete elementary	0.03	0.136	0.52	0.028	-0.14	0.141	0.54	0.026
complete primary	-1.64	0.215	1.03	0.037	-1.99	0.220	1.07	0.031
complete secondary	-3.52	0.370	1.48	0.052	-4.36	0.457	1.56	0.033
complete college	-3.37	0.350	1.68	0.090	-3.41	0.363	1.82	0.052
exp/10	0.65	0.113	0.45	0.025	0.53	0.116	0.47	0.022
exp/100	0.02	0.018	-0.05	0.004	0.04	0.018	-0.05	0.004
tenure	-0.32	0.011			-0.33	0.011		
sex (male=1)	1.25	0.080	0.18	0.020	1.25	0.081	0.21	0.014
metropolitan	1.78	0.091	0.22	0.014	1.83	0.095	0.23	0.014
race (white=1)	-1.25	0.101	0.09	0.016	-1.46	0.105	0.10	0.015
more than 5 years	0.06	0.136	0.47	0.024	0.16	0.139	0.47	0.024
former formal worker	0.50	0.137	0.33	0.021	0.47	0.143	0.32	0.020
former informal worker	0.22	0.092	-0.47	0.020	0.30	0.093	-0.47	0.020

cont.

54. As we cannot estimate a "selectivity-corrected" wage equation for informal workers in the case of a bivariate probit with simultaneous decision, we do not estimate the structural bivariate probit for this model.

55. Notice that the marginal effects of the wage differential in the "in the queue" equation and of the formal sector wage in the "chosen from the queue" equation is smaller in this case than in the Abowd-Faber bivariate probit.

56. Notice that only in this specification we could find a negative sign for seniority as expected according to the theory. Workers with seniority should be less likely to queue since they would lose the benefits that accrue from it.

cont.

	BP with sample selection (A)				BP with sample selection (B)			
	IQ		CFQ		IQ		CFQ	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
former public servant	0.37	0.502	0.16	0.102	0.20	0.530	0.22	0.101
former self-employed	0.11	0.180	-0.22	0.035	0.16	0.182	-0.23	0.034
#children 0-3 years	0.05	0.048			0.03	0.049		
#children 4-6 years	-0.03	0.085			0.01	0.086		
#children7-10 years	-0.17	0.062			-0.13	0.062		
#children 11-13 year	0.07	0.052			0.06	0.052		
# elder	0.01	0.108			-0.03	0.112		
# formal	0.21	0.041			0.17	0.042		
#self-employed	-0.02	0.070			-0.03	0.071		
#informal	-0.09	0.033			-0.10	0.033		
#unemployed	0.11	0.127			0.12	0.131		
per capita income (others)	-0.33	0.127			-0.22	0.209		
married	0.86	0.115	0.15	0.018	0.78	0.118	0.16	0.017
North	1.18	0.846	0.15	0.030	1.51	0.720	0.19	0.026
Southeast	0.14	0.127	0.42	0.021	0.13	0.133	0.44	0.018
South	1.09	0.191	0.54	0.026	1.17	0.198	0.56	0.024
Mid-West	0.79	0.109	0.12	0.028	0.89	0.111	0.16	0.024
Log L	-22924				-22917			
N	60138							

Overall, the estimation of different specifications for the structural bivariate probit seems to indicate that the wage differential is, indeed, a major contributor for the decision to queue for formal jobs. Similarly, schooling seems to be the main screening device used by formal employers when selecting workers from the queue. The most remarkable difference between the sequential decision specifications refers to the effect of have been an informal worker in the last job on both equations. The univariate probit specification indicates a negative impact of being a "former informal worker", i.e, workers who worked in the informal sector in their last job are less likely to be in the formal sector in the current one. However, whereas the structural Abowd-Farber bivariate probit shows that this negative effect is due to the fact that former informal sector workers are less likely to join the queue for formal jobs, the structural bivariate probit with sample selection indicates that this negative effect is due to employers "discriminating" against former informal workers.

5.6 ROBUSTNESS CHECKS

Tenure as an exclusion restriction in the bivariate probit models

Following the literature[57] we have excluded tenure (seniority) from the "chosen from the queue" equation of the bivariate probit models estimated in this paper. The argument is that as seniority is basically "specific human capital", the new employer should not look at it when deciding to hire or not an employee. Much more important for the employer's decision would be the worker's experience in the labour market[58]. Besides, as we included in both equations the employment history of the worker, we are already assessing the effect of the kind of experience that the

57. See Farber (1983) and Abowd and Farber (1982).

58. Notice that the variable experience is not "true experience", but the traditional proxy: $age - yearsofeducation - 6$.

individual had in the labour market,[59] an information that could be used as a screening device by the employers.

TABLE 13

Senstitivity of the abowd-faber bivariate probit to the inclusion of tenure in the CFQ

	Without tenure in the CFQ				With tenure in the CFQ			
	IQ		CFQ		IQ		CFQ	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	0.32	0.080	-0.44	0.064	-0.60	0.061	0.71	0.091
some elementary	-0.06	0.054	0.19	0.044	0.15	0.041	-0.04	0.062
complete elementary	-0.01	0.052	0.46	0.042	0.41	0.039	-0.03	0.059
complete primary	0.02	0.060	0.93	0.050	0.88	0.047	-0.08	0.067
complete secondary	0.08	0.063	1.30	0.052	1.23	0.050	-0.04	0.069
complete college	0.06	0.077	1.26	0.068	1.28	0.066	-0.14	0.084
exp/10	0.09	0.037	0.49	0.035	0.62	0.034	-0.11	0.042
exp/100	-0.04	0.007	-0.06	0.006	-0.08	0.006	-0.01	0.007
tenure	0.40	0.012			-0.003	0.002	0.51	0.016
sex (male=1)	0.00	0.026	-0.01	0.027	0.01	0.024	-0.01	0.030
metropolitan	0.16	0.025	0.25	0.024	0.22	0.022	0.16	0.028
race (white=1)	-0.02	0.026	0.03	0.025	0.03	0.023	-0.03	0.029
more than 5 years	-0.45	0.057	0.28	0.030	0.28	0.028	-0.53	0.065
former formal worker	0.27	0.044	1.02	0.048	0.86	0.039	0.24	0.052
former informal worker	-0.41	0.050	0.33	0.057	0.17	0.043	-0.45	0.058
former public servant	-0.12	0.123	1.10	0.415	9.62	1.4E+12	-0.39	0.124
former self-employed	-0.31	0.060	0.87	0.126	0.53	0.082	-0.31	0.070
#children 0-3 years	-0.02	0.015			-0.04	0.013		
#children 4-6 years	-0.01	0.026			0.03	0.023		
#children7-10 years	-0.04	0.019			-0.02	0.016		
#children 11-13 year	-0.01	0.018			-0.02	0.015		
# elder	-0.09	0.039			-0.06	0.033		
# formal	0.25	0.014			0.25	0.011		
#self-employed	-0.08	0.019			-0.09	0.016		
#informal	-0.23	0.012			-0.23	0.011		
#unemployed	0.03	0.026			0.07	0.025		
per capita income (others)	-0.21	0.030			-0.17	0.029		
married	0.04	0.028	0.14	0.027	0.17	0.025	-0.01	0.030
North	0.07	0.047	0.07	0.041	0.04	0.038	0.09	0.055
Southeast	0.19	0.033	0.34	0.031	0.29	0.028	0.19	0.037
South	0.32	0.041	0.37	0.038	0.32	0.035	0.34	0.046
Mid-West	0.03	0.041	0.08	0.040	0.08	0.036	0.02	0.047
constructing	0.24	0.068	-1.11	0.044	-0.96	0.040	0.27	0.075
retail	-0.08	0.036	-0.37	0.038	-0.30	0.034	-0.12	0.041
lodging, food and services	-0.48	0.040	-1.10	0.036	-1.02	0.032	-0.48	0.044
Productive Sector	-0.13	0.032	-0.07	0.037	-0.03	0.033	-0.20	0.037
Social Services	-0.27	0.040	-0.07	0.045	-0.02	0.041	-0.37	0.045
$\rho_{1,2}$								
Log L	-25059				-24856			
N	60138				60138			

59. Actually the dummy for "more than 5 years in the current job" - one of the variables related to the employment history - is already a proxy for seniority.

TABLE 14

Senstitivity of the poirier bivariate probit to the inclusion of tenure in the CFQ

	Without tenure in the CFQ				With tenure in the CFQ			
	IQ		CFQ		IQ		CFQ	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
some elementary	-0.01	0.031	0.17	0.026	-0.01	0.031	0.15	0.026
complete elementary	-0.05	0.030	0.45	0.025	-0.04	0.030	0.40	0.025
complete primary	0.00	0.036	0.86	0.030	0.01	0.036	0.77	0.030
complete secondary	0.06	0.038	1.26	0.033	0.08	0.038	1.13	0.033
complete college	0.05	0.050	1.51	0.052	0.08	0.050	1.32	0.053
exp/10	0.08	0.024	0.46	0.021	0.10	0.024	0.35	0.022
exp/100	-0.04	0.004	-0.05	0.004	-0.04	0.004	-0.04	0.004
tenure	0.03	0.002			0.02	0.002	0.06	0.002
sex (male=1)	-0.08	0.018	0.12	0.015	-0.07	0.018	0.09	0.016
metropolitan	0.08	0.017	0.23	0.014	0.07	0.017	0.23	0.014
race (white=1)	-0.08	0.018	0.09	0.014	-0.08	0.018	0.09	0.015
more than 5 years	0.20	0.025	0.41	0.024	0.19	0.025	0.26	0.024
former formal worker	0.51	0.025	0.30	0.020	0.47	0.025	0.49	0.021
former informal worker	0.21	0.024	-0.36	0.019	0.16	0.024	-0.19	0.020
former public servant	0.13	0.111	0.19	0.097	0.08	0.110	0.36	0.100
former self-employed	0.13	0.039	-0.14	0.032	0.07	0.039	0.08	0.033
#children 0-3 years	-0.02	0.011			-0.02	0.011		
#children 4-6 years	0.03	0.019			0.03	0.019		
#children7-10 years	-0.01	0.014			-0.01	0.014		
#children 11-13 year	0.02	0.013			0.02	0.013		
# elder	0.00	0.030			0.00	0.030		
# formal	0.16	0.009			0.16	0.009		
#self-employed	-0.03	0.014			-0.03	0.014		
#informal	-0.09	0.009			-0.09	0.009		
#unemployed	0.08	0.022			0.08	0.022		
per capita income (others)	-0.27	0.020			-0.27	0.019		
married	0.03	0.019	0.12	0.016	0.03	0.019	0.12	0.016
North	-0.22	0.030	0.20	0.025	-0.23	0.029	0.22	0.025
Southeast	-0.09	0.023	0.41	0.018	-0.09	0.023	0.41	0.018
South	-0.07	0.029	0.52	0.024	-0.07	0.029	0.52	0.024
Mid-West	-0.21	0.026	0.22	0.022	-0.21	0.026	0.22	0.022
constructing	-0.13	0.032	-0.67	0.024	-0.12	0.032	-0.64	0.024
retail	-0.21	0.027	-0.18	0.022	-0.21	0.027	-0.17	0.022
lodging, food and services	-0.44	0.024	-0.75	0.020	-0.44	0.024	-0.73	0.020
Productive Sector	0.00	0.026	-0.10	0.022	0.00	0.026	-0.11	0.023
Social Services	-0.09	0.032	-0.14	0.028	-0.09	0.032	-0.14	0.028
$\rho_{1,2}$	-0.96	0.011			-0.96178	0.011657		
Log L	-35326				-34949.8			
N	60138				60138			

TABLE 15

Sensitivity of the bivariate probit with sample selection to the inclusion of tenure in the CFQ

	Without tenure in the CFQ				With tenure in the CFQ			
	IQ		CFQ		IQ		CFQ	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	0.71	0.082	-0.42	0.062	-0.53	0.060	0.99	0.091
some elementary	-0.10	0.055	0.19	0.043	0.16	0.040	-0.09	0.061
complete elementary	-0.10	0.053	0.44	0.040	0.41	0.038	-0.13	0.059
complete primary	-0.13	0.062	0.89	0.048	0.85	0.046	-0.22	0.068
complete secondary	-0.12	0.065	1.28	0.051	1.20	0.048	-0.21	0.070
complete college	-0.10	0.079	1.26	0.066	1.26	0.065	-0.29	0.085
exp/10	0.03	0.038	0.47	0.034	0.57	0.033	-0.15	0.042
exp/100	-0.03	0.007	-0.06	0.006	-0.07	0.006	0.00	0.007
tenure	0.39	0.011			0.00	0.002	0.51	0.016
sex (male=1)	0.00	0.028	0.03	0.025	0.02	0.024	0.01	0.030
metropolitan	0.11	0.026	0.24	0.022	0.21	0.021	0.12	0.029
race (white=1)	-0.02	0.027	0.03	0.024	0.03	0.022	-0.03	0.029
more than 5 years	-0.46	0.055	0.27	0.029	0.27	0.028	-0.50	0.062
former formal worker	0.16	0.044	0.89	0.041	0.77	0.036	0.18	0.051
former informal worker	-0.45	0.053	0.36	0.052	0.22	0.044	-0.43	0.058
former public servant	-0.24	0.129	0.99	0.314	0.76	0.218	-0.25	0.140
former self-employed	-0.42	0.063	0.78	0.092	0.52	0.073	-0.35	0.072
#children 0-3 years	-0.02	0.014			-0.03	0.011		
#children 4-6 years	-0.02	0.023			0.02	0.020		
#children7-10 years	-0.04	0.017			-0.02	0.014		
#children 11-13 year	-0.01	0.016			-0.02	0.013		
# elder	-0.09	0.036			-0.07	0.029		
# formal	0.24	0.013			0.22	0.010		
#self-employed	-0.08	0.017			-0.08	0.014		
#informal	-0.20	0.011			-0.19	0.009		
#unemployed	0.03	0.024			0.06	0.021		
per capita income (others)	-0.19	0.028			-0.15	0.026		
married	0.02	0.029	0.13	0.025	0.16	0.024	-0.03	0.031
North	0.05	0.048	0.08	0.040	0.03	0.037	0.10	0.055
Southeast	0.13	0.034	0.31	0.029	0.26	0.028	0.15	0.038
South	0.26	0.043	0.34	0.036	0.29	0.034	0.30	0.047
Mid-West	0.02	0.043	0.10	0.038	0.07	0.035	0.03	0.048
constructing	0.49	0.070	-1.08	0.043	-0.95	0.040	0.52	0.078
retail	-0.03	0.038	-0.33	0.037	-0.30	0.033	-0.05	0.043
lodging, food and services	-0.18	0.046	-1.03	0.035	-0.96	0.032	-0.20	0.048
Productive Sector	-0.12	0.034	-0.05	0.036	-0.02	0.033	-0.19	0.038
Social Services	-0.26	0.042	-0.05	0.044	-0.02	0.040	-0.34	0.046
$\rho_{1,2}$	-0.80	0.029			-0.79505	0.031069		
Log L	-24979				-24773.5			
N	60138				60138			

Nevertheless, unlike the other exclusion restrictions to identify the "chosen from the queue" equation that are related to the situation of the household where the individual resides, tenure is related to the labour market and as such can also determine the employer's decision. For that reason we re-estimate the bivariate probit models including tenure as an additional regressor in the "chosen from the queue" equation.

The results of the Abowd-Farber bivariate probit with partial observability and of the Poirier biraviate probit with partial observability are very different when we include tenure (see table 12 and table 13). The sign of the variables related to the past experience change in both equations and so do the coefficients of the variables related to education.[60] According to the new results, education seems to increase the probability of being in the queue and does not have any effect on the probability of being chosen from the queue. The change in the variables related to the individual labour market history does not come as a surprise since tenure must be correlated with them, but the change in the signal of the education dummies are quite surprising. One possible explanation for that is the fact that both Abowd-Faber bivariate probit and Poirier bivariate probit with partial observability rely heavily on the exclusion restrictions and on the non-linearities of the model in order to identify separately the two equations. As for the impact of tenure on the two equations, it has no effect on the probability of being in the queue and increases the probability of being chosen from the queue.

Table 14 shows the results for the bivariate probit with sample selection. Unlike the bivariate probit models with partial observability, the bivariate probit with sample selection shows much more stability on their estimates regardless of the inclusion of tenure as an explanatory variable in the "chosen from the queue" equation. The only variables whose coefficients magnitude change, but not the sign, are those related to the labour market history of the individual as should be expected given their correlation with tenure. As for its effect, tenure seems to affect positively both the probability of being in the queue and the probability of being "chosen from the queue".

These results mean that the bivariate probit with sample selection seems to be the most suitable model to treat the "job queue" issue, not only because it offers a richer specification than the other two bivariate probit with partial observability, but also because it is much less sensitive to changes in the exclusion restrictions.[61]

The stability of the parameters of the wage equations

Tables 16 to table 19 compare the different wage equation estimates that we get if we exclude one set of exclusion restriction per time from our previous estimations. We report the full model, and then we report the results with the variables omitted. These variables are: 1) the variables related to the number of children and older people in the household; 2) the past employment variables; 3) other member of household income (*per capita*); 4) variable related to the occupational distribution of the other members of the household.

60. But only the coefficients of the variables related to labour market history are statistically significant.

61. As for the impact in the wage equations corrected for selectivity, we do not observe any major changes in the estimated parameters due to the inclusion of tenure in the "chosen from the queue". The tables with these results are available upon request.

TABLE 16

OLS wage equations based on the univariate probit with different exclusions restrictions

	OLS - 2STEP											
	Full model				Excluding children and elder				Excluding past employment			
	Formal		Informal		Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.96	0.03	-1.80	0.03	-0.96	0.03	-1.80	0.03	-1.10	0.03	-1.78	0.03
some elementary	0.14	0.02	0.15	0.02	0.14	0.02	0.15	0.02	0.15	0.02	0.16	0.02
complete elementary	0.32	0.01	0.36	0.02	0.33	0.01	0.36	0.02	0.35	0.02	0.38	0.02
complete primary	0.64	0.02	0.64	0.03	0.64	0.02	0.65	0.03	0.69	0.02	0.69	0.03
complete secondary	1.12	0.02	1.07	0.03	1.12	0.02	1.07	0.03	1.18	0.02	1.13	0.04
complete college	1.93	0.02	1.89	0.05	1.93	0.02	1.89	0.05	1.99	0.02	1.96	0.05
exp/10	0.30	0.01	0.36	0.02	0.30	0.01	0.37	0.02	0.32	0.01	0.39	0.02
exp/100	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.06	0.00	-0.05	0.00	-0.06	0.00
tenure	0.03	0.00	0.01	0.00	0.03	0.00	0.01	0.00	0.03	0.00	0.01	0.00
sex (male=1)	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01
metropolitan	0.09	0.01	0.19	0.01	0.09	0.01	0.19	0.01	0.10	0.01	0.20	0.01
race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
married	0.17	0.01	0.21	0.01	0.17	0.01	0.21	0.01	0.18	0.01	0.22	0.01
North	0.42	0.01	0.48	0.02	0.42	0.01	0.48	0.02	0.43	0.01	0.48	0.02
Southeast	0.28	0.01	0.31	0.02	0.28	0.01	0.31	0.02	0.30	0.01	0.33	0.02
South	0.27	0.01	0.35	0.02	0.27	0.01	0.36	0.02	0.30	0.01	0.38	0.02
Mid-West	0.41	0.01	0.46	0.02	0.41	0.01	0.46	0.02	0.41	0.01	0.46	0.02
constructing	-0.01	0.01	0.14	0.02	-0.01	0.01	0.14	0.02	-0.05	0.01	0.10	0.02
retail	-0.21	0.01	-0.02	0.02	-0.21	0.01	-0.03	0.02	-0.23	0.01	-0.04	0.02
lodging, food and services	-0.22	0.01	-0.13	0.02	-0.22	0.01	-0.14	0.02	-0.28	0.02	-0.18	0.02
Productive Sector	0.02	0.01	0.10	0.02	0.02	0.01	0.10	0.02	0.02	0.01	0.10	0.02
Social Services	-0.16	0.01	-0.06	0.02	-0.16	0.01	-0.06	0.02	-0.17	0.01	-0.07	0.02
Inverse Mills ratio (lambda)	-0.20	0.02	-0.21	0.02	-0.20	0.02	-0.20	0.02	-0.05	0.02	-0.12	0.03

	Excluding other's income				Excluding other's ocupation			
	Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.90	0.03	-1.81	0.03	-0.98	0.03	-1.82	0.03
some elementary	0.14	0.02	0.14	0.02	0.14	0.02	0.15	0.02
complete elementary	0.31	0.01	0.35	0.02	0.33	0.02	0.36	0.02
complete primary	0.62	0.02	0.62	0.03	0.65	0.02	0.64	0.03
complete secondary	1.10	0.02	1.02	0.03	1.13	0.02	1.05	0.03
complete college	1.91	0.02	1.84	0.05	1.94	0.02	1.88	0.05
exp/10	0.29	0.01	0.35	0.02	0.30	0.01	0.36	0.02
exp/100	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00
tenure	0.02	0.00	0.01	0.00	0.03	0.00	0.01	0.00
sex (male=1)	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01
metropolitan	0.08	0.01	0.18	0.01	0.09	0.01	0.18	0.01
race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
married	0.16	0.01	0.20	0.01	0.17	0.01	0.21	0.01
North	0.42	0.01	0.48	0.02	0.42	0.01	0.48	0.02
Southeast	0.27	0.01	0.29	0.02	0.28	0.01	0.30	0.02
South	0.26	0.01	0.34	0.02	0.28	0.01	0.35	0.02
Mid-West	0.41	0.01	0.45	0.02	0.41	0.01	0.46	0.02
constructing	0.01	0.01	0.16	0.02	-0.02	0.01	0.15	0.02
retail	-0.21	0.01	-0.02	0.02	-0.22	0.01	-0.02	0.02
lodging, food and services	-0.19	0.01	-0.10	0.02	-0.24	0.01	-0.12	0.02
Productive Sector	0.03	0.01	0.10	0.02	0.02	0.01	0.11	0.02
Social Services	-0.16	0.01	-0.06	0.02	-0.17	0.01	-0.05	0.02
Inverse Mills ratio (lambda)	-0.26	0.02	-0.27	0.02	-0.16	0.02	-0.23	0.03

TABLE 17

MLE wage equations based on the univariate probit with different exclusions restrictions

	Maximum likelihood estimation											
	Full model				Excluding children and elder				Excluding past employment			
	Formal		Informal		Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.83	0.023	-1.80	0.027	-0.83	0.02	-1.80	0.03	-0.82	0.02	-1.80	0.03
some elementary	0.13	0.017	0.15	0.019	0.13	0.02	0.15	0.02	0.13	0.02	0.15	0.02
complete elementary	0.30	0.016	0.36	0.019	0.30	0.02	0.36	0.02	0.30	0.02	0.37	0.02
complete primary	0.60	0.018	0.63	0.025	0.60	0.02	0.64	0.03	0.59	0.02	0.65	0.03
complete secondary	1.06	0.018	1.05	0.030	1.06	0.02	1.05	0.03	1.06	0.02	1.08	0.03
complete college	1.88	0.020	1.86	0.040	1.88	0.02	1.87	0.04	1.88	0.02	1.89	0.04
exp/10	0.27	0.010	0.36	0.017	0.27	0.01	0.36	0.02	0.28	0.01	0.37	0.02
exp/100	-0.05	0.002	-0.05	0.003	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.06	0.00
tenure	0.02	0.001	0.01	0.001	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00
sex (male=1)	0.30	0.007	0.38	0.012	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01
metropolitan	0.08	0.006	0.18	0.012	0.08	0.01	0.18	0.01	0.07	0.01	0.19	0.01
race (white=1)	0.16	0.007	0.11	0.012	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
married	0.16	0.007	0.21	0.013	0.16	0.01	0.21	0.01	0.16	0.01	0.21	0.01
North	0.42	0.011	0.47	0.018	0.42	0.01	0.47	0.02	0.42	0.01	0.48	0.02
Southeast	0.26	0.009	0.30	0.015	0.26	0.01	0.31	0.02	0.26	0.01	0.31	0.02
South	0.25	0.011	0.35	0.022	0.24	0.01	0.35	0.02	0.24	0.01	0.36	0.02
Mid-West	0.41	0.011	0.46	0.017	0.41	0.01	0.46	0.02	0.41	0.01	0.46	0.02
constructing	0.03	0.014	0.14	0.022	0.03	0.01	0.14	0.02	0.03	0.01	0.13	0.02
retail	-0.20	0.010	-0.03	0.019	-0.20	0.01	-0.03	0.02	-0.20	0.01	-0.03	0.02
lodging, food and services	-0.16	0.012	-0.13	0.020	-0.16	0.01	-0.13	0.02	-0.15	0.01	-0.15	0.02
Productive Sector	0.03	0.008	0.10	0.019	0.03	0.01	0.10	0.02	0.03	0.01	0.10	0.02
Social Services	-0.15	0.011	-0.06	0.024	-0.15	0.01	-0.06	0.02	-0.16	0.01	-0.06	0.02
$\rho_{1,u}$	-0.53	0.015	-0.33	0.028	-0.53		-0.32		-0.52		-0.27	

	Excluding other's income				Excluding other's ocupation			
	Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.85	0.02	-1.80	0.03	-0.81	0.02	-1.82	0.03
some elementary	0.13	0.02	0.15	0.02	0.13	0.02	0.15	0.02
complete elementary	0.30	0.02	0.36	0.02	0.30	0.02	0.35	0.02
complete primary	0.61	0.02	0.65	0.03	0.59	0.02	0.62	0.03
complete secondary	1.07	0.02	1.07	0.03	1.06	0.02	1.04	0.03
complete college	1.89	0.02	1.89	0.04	1.87	0.02	1.85	0.04
exp/10	0.28	0.01	0.37	0.02	0.27	0.01	0.36	0.02
exp/100	-0.05	0.00	-0.06	0.00	-0.05	0.00	-0.05	0.00
tenure	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00
sex (male=1)	0.30	0.01	0.38	0.01	0.30	0.01	0.38	0.01
metropolitan	0.08	0.01	0.19	0.01	0.07	0.01	0.18	0.01
race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
married	0.16	0.01	0.21	0.01	0.16	0.01	0.21	0.01
North	0.42	0.01	0.48	0.02	0.42	0.01	0.47	0.02
Southeast	0.26	0.01	0.31	0.02	0.26	0.01	0.30	0.02
South	0.25	0.01	0.36	0.02	0.24	0.01	0.35	0.02
Mid-West	0.41	0.01	0.46	0.02	0.40	0.01	0.46	0.02
constructing	0.02	0.01	0.13	0.02	0.03	0.01	0.16	0.02
retail	-0.20	0.01	-0.03	0.02	-0.20	0.01	-0.02	0.02
lodging, food and services	-0.17	0.01	-0.14	0.02	-0.15	0.01	-0.11	0.02
Productive Sector	0.03	0.01	0.10	0.02	0.03	0.01	0.11	0.02
Social Services	-0.16	0.01	-0.07	0.02	-0.15	0.01	-0.06	0.02
$\rho_{1,u}$	-0.49		-0.29		-0.35	0.03	-0.53	0.02

TABLE 18

Wage equations based on the abowd-farber biivariate probit with different exclusions restrictions

	Abowd-faber (A)											
	Full model				Excluding children and elder				Excluding past employment			
	Formal		Informal		Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-1.02	0.03	-1.69	0.04	-1.02	0.03	-1.69	0.04	-1.55	0.04	-1.68	0.09
some elementary	0.15	0.02	0.14	0.02	0.15	0.02	0.14	0.02	0.20	0.02	0.17	0.02
complete elementary	0.34	0.02	0.33	0.02	0.34	0.01	0.33	0.02	0.46	0.02	0.38	0.04
complete primary	0.67	0.02	0.57	0.03	0.67	0.02	0.58	0.03	0.87	0.02	0.69	0.08
complete secondary	1.16	0.02	0.97	0.04	1.16	0.02	0.97	0.04	1.42	0.03	1.15	0.12
complete college	1.97	0.02	1.81	0.05	1.97	0.02	1.81	0.05	2.21	0.03	1.98	0.12
exp/10	0.31	0.01	0.32	0.02	0.31	0.01	0.32	0.02	0.42	0.01	0.39	0.05
exp/100	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.07	0.00	-0.06	0.01
tenure	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00	0.02	0.00	0.00	0.00
sex (male=1)	0.30	0.01	0.36	0.01	0.30	0.01	0.37	0.01	0.30	0.01	0.37	0.01
metropolitan	0.10	0.01	0.16	0.01	0.10	0.01	0.17	0.01	0.15	0.01	0.20	0.03
race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
married	0.17	0.01	0.20	0.01	0.17	0.01	0.20	0.01	0.21	0.01	0.22	0.02
North	0.42	0.01	0.47	0.02	0.42	0.01	0.47	0.02	0.42	0.01	0.47	0.02
Southeast	0.29	0.01	0.28	0.02	0.29	0.01	0.28	0.02	0.37	0.01	0.33	0.04
South	0.28	0.01	0.32	0.02	0.28	0.01	0.33	0.02	0.38	0.01	0.38	0.04
Mid-West	0.41	0.01	0.45	0.02	0.41	0.01	0.45	0.02	0.42	0.01	0.46	0.02
constructing	-0.04	0.02	0.19	0.03	-0.04	0.02	0.19	0.03	-0.28	0.02	0.06	0.09
retail	-0.22	0.01	0.00	0.02	-0.22	0.01	0.00	0.02	-0.29	0.01	-0.05	0.04
lodging, food and services	-0.25	0.01	-0.06	0.02	-0.25	0.01	-0.06	0.02	-0.55	0.03	-0.20	0.10
Productive Sector	0.02	0.01	0.11	0.02	0.02	0.01	0.11	0.02	0.00	0.01	0.11	0.02
Social Services	-0.17	0.01	-0.05	0.02	-0.17	0.01	-0.05	0.02	-0.20	0.01	-0.06	0.03
σ_{1f}	-0.19	0.02			-0.18	0.02			-0.13	0.02		
σ_{2f}	-0.07	0.03			-0.07	0.03			0.63	0.06		
σ_{1i}			-0.18	0.02			-0.17	0.02			-0.14	0.02
δ_1			-0.18	0.02			-0.18	0.02			-0.11	0.02
δ_2			-0.18	0.02			-0.18	0.02			-0.04	0.13
R_2	0.56		0.47		0.56		0.47		0.56		0.46	

	Excluding other's income				Excluding other's ocupation			
	Formal		Informal		Formal		Informal	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-1.01	0.03	-1.69	0.04	-1.03	0.03	-1.73	0.04
some elementary	0.15	0.02	0.14	0.02	0.15	0.02	0.14	0.02
complete elementary	0.34	0.01	0.33	0.02	0.34	0.02	0.33	0.02
complete primary	0.67	0.02	0.57	0.03	0.67	0.02	0.58	0.03
complete secondary	1.16	0.02	0.97	0.04	1.16	0.02	0.98	0.04
complete college	1.97	0.02	1.81	0.05	1.97	0.02	1.81	0.05
exp/10	0.31	0.01	0.32	0.02	0.31	0.01	0.33	0.02
exp/100	-0.05	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00
tenure	0.02	0.00	0.01	0.00	0.02	0.00	0.01	0.00
sex (male=1)	0.30	0.01	0.36	0.01	0.30	0.01	0.37	0.01
metropolitan	0.10	0.01	0.16	0.01	0.10	0.01	0.17	0.01
race (white=1)	0.16	0.01	0.11	0.01	0.16	0.01	0.11	0.01
married	0.17	0.01	0.20	0.01	0.17	0.01	0.20	0.01
North	0.42	0.01	0.47	0.02	0.42	0.01	0.47	0.02
Southeast	0.29	0.01	0.28	0.02	0.29	0.01	0.29	0.02
South	0.28	0.01	0.32	0.02	0.28	0.01	0.33	0.02
Mid-West	0.41	0.01	0.45	0.02	0.41	0.01	0.45	0.02
constructing	-0.04	0.02	0.18	0.03	-0.04	0.01	0.19	0.03
retail	-0.22	0.01	0.00	0.02	-0.22	0.01	0.00	0.02
lodging, food and services	-0.25	0.01	-0.06	0.02	-0.26	0.01	-0.07	0.02
Productive Sector	0.02	0.01	0.11	0.02	0.02	0.01	0.11	0.02
Social Services	-0.17	0.01	-0.05	0.02	-0.17	0.01	-0.05	0.02
σ_{1f}	-0.22	0.02			-0.17	0.02		
σ_{2f}	-0.07	0.03			-0.08	0.02		
σ_{1i}			-0.21	0.02			-0.15	0.02
δ_1			-0.19	0.02			-0.15	0.03
δ_2			-0.18	0.02			-0.19	0.02
R_2	0.56		0.47		0.56		0.46	

TABLE 19

Wage equations based on the bivariate probit with sample selection with different exclusions restrictions

	Full model						Excluding #children and #elder						Excluding past employment					
	Formal		Informal (IQ)		Informal (NIQ)		Formal		Informal (IQ)		Informal (NIQ)		Formal		Informal (IQ)		Informal (NIQ)	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.99	0.029	-1.62	0.031	-1.62	0.115	-0.99	0.028	-1.62	0.031	-1.62	0.116	-1.23	0.037	-1.55	0.035	-1.01	0.141
some elementary	0.14	0.016	0.11	0.021	0.21	0.042	0.14	0.015	0.11	0.020	0.21	0.042	0.17	0.016	0.19	0.028	0.22	0.042
complete elementary	0.32	0.016	0.26	0.023	0.50	0.040	0.32	0.015	0.26	0.023	0.50	0.040	0.38	0.016	0.47	0.053	0.51	0.040
complete primary	0.64	0.019	0.47	0.033	0.92	0.049	0.64	0.018	0.47	0.033	0.92	0.050	0.74	0.021	0.85	0.096	0.94	0.050
complete secondary	1.11	0.020	0.81	0.044	1.37	0.052	1.11	0.020	0.81	0.044	1.37	0.052	1.24	0.024	1.37	0.141	1.41	0.052
complete college	1.91	0.023	1.64	0.076	2.06	0.069	1.91	0.022	1.64	0.077	2.06	0.069	2.04	0.026	2.33	0.182	2.07	0.069
exp/10	0.31	0.010	0.28	0.022	0.47	0.032	0.31	0.010	0.28	0.022	0.47	0.032	0.35	0.011	0.50	0.057	0.52	0.033
exp/100	-0.06	0.002	-0.04	0.004	-0.07	0.006	-0.06	0.002	-0.04	0.004	-0.07	0.006	-0.06	0.002	-0.07	0.007	-0.09	0.006
tenure	0.03	0.001	0.01	0.002	0.01	0.002	0.03	0.001	0.01	0.002	0.01	0.002	0.03	0.001	0.01	0.002	0.02	0.002
sex (male=1)	0.29	0.007	0.35	0.014	0.35	0.025	0.29	0.007	0.35	0.014	0.35	0.025	0.30	0.007	0.41	0.020	0.35	0.025
metropolitan	0.10	0.007	0.19	0.014	0.18	0.025	0.10	0.006	0.19	0.014	0.18	0.025	0.12	0.007	0.29	0.027	0.21	0.025
race (white=1)	0.15	0.007	0.07	0.013	0.16	0.025	0.15	0.007	0.07	0.013	0.16	0.025	0.16	0.007	0.10	0.015	0.13	0.025
married	0.17	0.007	0.17	0.015	0.30	0.026	0.17	0.007	0.17	0.015	0.30	0.026	0.18	0.007	0.23	0.020	0.31	0.026
North	0.40	0.013	0.40	0.021	0.53	0.042	0.40	0.013	0.40	0.021	0.53	0.042	0.42	0.013	0.47	0.026	0.46	0.043
Southeast	0.28	0.009	0.30	0.018	0.23	0.033	0.28	0.009	0.30	0.017	0.23	0.033	0.32	0.010	0.46	0.040	0.23	0.033
South	0.27	0.011	0.34	0.025	0.28	0.041	0.27	0.011	0.34	0.025	0.28	0.041	0.32	0.012	0.55	0.055	0.28	0.041
Mid-West	0.38	0.012	0.47	0.019	0.33	0.038	0.38	0.012	0.47	0.018	0.33	0.038	0.41	0.012	0.53	0.024	0.27	0.039
constructing	-0.02	0.015	0.15	0.025	0.13	0.046	-0.02	0.014	0.15	0.025	0.13	0.046	-0.09	0.016	-0.10	0.066	0.09	0.046
retail	-0.23	0.010	-0.08	0.021	0.04	0.040	-0.23	0.010	-0.08	0.021	0.03	0.040	-0.24	0.010	-0.15	0.027	-0.02	0.040
lodging, food and services	-0.29	0.013	-0.15	0.024	-0.18	0.038	-0.29	0.012	-0.15	0.024	-0.18	0.038	-0.37	0.014	-0.44	0.072	-0.31	0.042
Productive Sector	0.02	0.009	0.06	0.022	0.20	0.039	0.02	0.008	0.06	0.022	0.20	0.039	0.01	0.008	0.03	0.024	0.20	0.039
Social Services	-0.18	0.011	-0.02	0.027	-0.15	0.047	-0.18	0.011	-0.02	0.028	-0.15	0.047	-0.18	0.011	-0.07	0.031	-0.18	0.047
σ_{1f}	0.31	0.037					0.30	0.037					0.13	0.044				
σ_{2f}	-0.24	0.030					-0.24	0.030					0.03	0.040				
σ_{1i*}			-0.89	0.123					-0.89	0.114					-1.03	0.116		
σ_{1i}					-0.01	0.054					-0.01	0.054					0.32	0.070
σ_{2i}			-0.17	0.030					-0.17	0.030					0.34	0.124		
R_s	0.56		0.40		0.48		0.56		0.40		0.48		0.56		0.40		0.48	

	Excluding other's income						Excluding other's ocupation					
	Formal		Informal (IQ)		Informal (NIQ)		Formal		Informal (IQ)		Informal (NIQ)	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-1.10	0.03	-1.62	0.03	-2.21	0.12	-0.91	0.03	-1.63	0.03	-1.46	0.15
some elementary	0.15	0.02	0.11	0.02	0.21	0.04	0.14	0.02	0.11	0.02	0.22	0.04
complete elementary	0.35	0.02	0.26	0.02	0.50	0.04	0.30	0.02	0.27	0.02	0.50	0.04
complete primary	0.69	0.02	0.47	0.03	0.89	0.05	0.61	0.02	0.48	0.03	0.92	0.05
complete secondary	1.18	0.02	0.80	0.04	1.33	0.05	1.07	0.02	0.82	0.04	1.38	0.05
complete college	1.99	0.02	1.64	0.08	2.05	0.07	1.86	0.02	1.66	0.08	2.06	0.07
exp/10	0.32	0.01	0.28	0.02	0.43	0.03	0.30	0.01	0.29	0.02	0.48	0.03
exp/100	-0.05	0.00	-0.04	0.00	-0.06	0.01	-0.06	0.00	-0.05	0.00	-0.08	0.01
tenure	0.03	0.00	0.01	0.00	0.01	0.00	0.03	0.00	0.01	0.00	0.01	0.00
sex (male=1)	0.30	0.01	0.35	0.01	0.36	0.02	0.28	0.01	0.35	0.01	0.35	0.03
metropolitan	0.11	0.01	0.19	0.01	0.15	0.02	0.10	0.01	0.20	0.01	0.19	0.03
race (white=1)	0.16	0.01	0.07	0.01	0.18	0.03	0.14	0.01	0.07	0.01	0.15	0.03
married	0.18	0.01	0.17	0.02	0.28	0.03	0.16	0.01	0.17	0.02	0.30	0.03
North	0.42	0.01	0.40	0.02	0.58	0.04	0.37	0.01	0.40	0.02	0.51	0.04
Southeast	0.30	0.01	0.30	0.02	0.23	0.03	0.26	0.01	0.30	0.02	0.23	0.03
South	0.30	0.01	0.34	0.02	0.28	0.04	0.25	0.01	0.35	0.02	0.28	0.04
Mid-West	0.41	0.01	0.47	0.02	0.38	0.04	0.36	0.01	0.46	0.02	0.31	0.04
constructing	-0.05	0.01	0.15	0.02	0.16	0.05	0.01	0.01	0.15	0.02	0.12	0.05
retail	-0.23	0.01	-0.08	0.02	0.09	0.04	-0.24	0.01	-0.09	0.02	0.02	0.04
lodging, food and services	-0.29	0.01	-0.14	0.02	-0.07	0.04	-0.30	0.01	-0.16	0.02	-0.22	0.04
Productive Sector	0.02	0.01	0.06	0.02	0.20	0.04	0.02	0.01	0.06	0.02	0.20	0.04
Social Services	-0.18	0.01	-0.02	0.03	-0.13	0.05	-0.18	0.01	-0.02	0.03	-0.16	0.05
σ_{1f}	0.00	0.04					0.58	0.05				
σ_{2f}	-0.04	0.03					-0.42	0.03				
σ_{1i*}			-0.92	0.11					-1.09	0.20		
σ_{1i}					-0.33	0.06					0.08	0.07
σ_{2i}			-0.17	0.03					-0.16	0.03		
R_s	0.55		0.40		0.48		0.56		0.40		0.48	

The results indicate a high degree of stability of the parameters of the wage equations. The estimates of the human capital variables do not vary much with changes in the exclusion restrictions of the sector allocation equations. The only exception refers to the variables related to past employment history. The estimated parameters are quite sensitive to their exclusion, and tend to be higher than the ones we get from the full model. As these variables are jointly significant in all estimates of the bivariate probit and display sensible effects in both the "in the queue" and the "chosen from the queue" equation, the models that exclude these variables may not have captured the precise nature of the sector allocation process.

6 CONCLUSION

This paper tackled the issue of segmentation between formal and informal sector in Brazil. Unlike most studies on the literature that emphasize the wage differential between the two types of workers or different wage structures, we concentrate our attention on the relationship between unobservables that determine the sector allocation process and unobservables that determine wages. This approach has been used to test the hypothesis of comparative advantage in the sector allocation process based on the worker's choice. However, we depart from this approach, as we argue that this univariate approach would hide the true nature of the allocation process, which according to the evidence presented here is best depicted as a two-decision process: the decision of a worker to join the queue for a formal job and the decision of the employer to pick him/her up from the pool of workers in the queue.

We applied both Abowd and Farber (1982) and Mengistae (1998) tests for the existence of a job queue for formal jobs in Brazil and none of them were able to reject this hypothesis. We also present estimates of the job queue "length"[62] for selected groups and show that non-white (1.43), female (1.37), illiterate (1.74), "new entrants"(1.45) and former informal workers (2.11) are the groups with the lower probability of being chosen from the queue conditional on being in the queue.[63] This result is particularly strong for workers whose last job was in the informal sector, suggesting that a spell in the informal sector may jeopardize the worker's chance of getting a formal job. Assuming that workers really would prefer and would be better off if they would get a formal job, then these groups should receive special attention of public policies to encourage formal employers to hire them.

The estimation of the "in the queue" and of the "chosen from the queue" equation separately also allowed us to uncover some relationships that were hidden by the univariate procedure in the sector allocation process. The different impact of education levels in the "in the queue" and "chosen from the queue" equations is a good example: whereas workers from low education groups seem to "join the queue" with a higher probability than more educated workers, the latter are much more likely to be chosen from it.

62. Job length is this case is defined as the inverse of the probability of getting a formal job once in the queue.
63. The numbers between parenthesis correspond to the length of the queue.

The results for the wage equation show the existence of different types of selectivity. But the way selectivity affects wages differs according to the assumed model. In the case of sequential models, unobservables that make workers more likely to be chosen from the queue make them earn less than expected wages in the formal sector. Thus it seems that workers less likely to be chosen to work in the formal sector are the ones who benefit more from this condition. On the other hand, workers less likely to be chosen from the queue are the ones who command higher wages in the informal sector.

As for the structural bivariate probit, the results indicate that wage differential plays the most important role in the decision to join the queue. However, the coefficient on the wage in the formal sector in the "chosen from the queue" equation displays an unexpected positive sign.[64] We attribute this puzzling result to the fact that the wage regressor seems to be capturing the effect of the human-capital related variables.

In terms of the methodology to estimating bivariate probit with partial observability, robustness checks indicated that the bivariate probit with sample selection is much less sensitive to minor changes in the specification than the other types of bivariate probit.[65] Therefore, the availability of the information about the worker's willingness to change to a formal job was a fundamental piece of information that allowed us to relax the severe partial observability of the Abowd-Farber and Poirier's bivariate probits, yielding much more information on the process of allocation of workers between sectors.

64. Note that the fact that the higher the probability of being chosen from the queue the lower the wage indicates that employers somehow manage to minimize the cost in the hiring process.
65. The robustness checks also revealed that the wage equation results are quite stable as long as the past employment history variables are included among the exclusion restrictions that identify the wage equation separately.

REFERENCES

Abowd, J. M. and Farber, H. S (1982) "Job Queues and the Union Status of Workers". Industrial and Labor Relations Review, v. 35, n. 3, p. 354-367.

Ahn, S. C.(1992) "The Lagrangian multiplier test for a model with two selectivity criteria". Economics Letters, v. 38, p. 9-15.

Carneiro, F. G. and Henley, A. (2001) "Modelling Formal vs. Informal employment and earnings: microeconometric evidence for Brazil", Annals of the XXIX National Meeting of Economics of ANPEC.

Co, C., Gang, Ira N. and Yun, M. (2002) "Self-Employment and wage earning in Hungary during transition. Mimeo.

Farber, H. S. (1983) The determination of the union status of workers". Econometrica, v. 51, p. 1417-1437.

Ham, J. C. (1982) "Estimation of a Labour Supply Model with Censoring Due to Unemployment and Underemployment".The Review of Economic Studies, v. 49, n. 3, p. 335-354.

Heckman, James J. (1979) "Sample Selection Bias as a Specification Error". Econometrica, v. 47, n. 1, p. 153-161.

Heckman, James J., and Honore, Bo., (1990) "The Empirical Content of the Roy Model", Econometrica, v. 58, n. 5, p. 1121-1149.

Heckman, James J. and Pagés, C. (2004) Law and Employment: Lessons from Latin America and the Caribbean. University of Chicago Press.

Heywood, John S. and Mohanty, Madhu S. (1995) "Estimation of the US Federal Job Queue in the Presence of an Endogenous Union Queue". Economica, v. 62, p. 479 - 493.

Holtzer H. J., Katz, L. F and Krueger, A. B., (1991) "Job Queues and Wages". Quarterly Journal of Economics, p. 739-768.

Lanot, G. and Walker, I. (1998) "The union/non-union wage differential: An application of semi-parametric methods". Journal of Econometrics, v. 84, p. 327-349.

Maddala, G. S. (1983) Limited Dependent and Qualitative Variables in Econonometrics. Cambridge University Press.

Maloney, W. F. (1997) "The Structure of Labor Market in Developing Countries: Time Series Evidence on Competing Views". Mimeo. World Bank.

Marcoullier, D., V. Ruiz de Casilla, and C. Woodruff (1997) "Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru", Economic Development and Cultural Change, v. 45, n. 2, p. 367-392.

Martins, M. F. O. (2001) "Parametric and Semiparametric Estimation of Sample Selection Models: An Empirical Application to the Female Labour Force in Portugal". Journal of Applied Econometrics. v. 16, p. 23-39.

Mengistae, Taye (1998) "Wage Rates and Job Queues: Does the Public Sector Overpay in Ethiopia?" St. Antony's College and CSAE, University of Oxford. Working Paper Series/98-20.

Mohanty, Madhu S. (2002) "Bivariate Probit Approach to the Determination of Teen Employment: A Study of Employment Discrimination in Los Angeles County." Applied Economics, v. 34, p. 143-56.

Mohanty, Madhu S. (1998)"The Role of the Desire for Union Status in the Decision to Enter Local Government Job Queues: The US Evidence, Applied Economics, v. 30, p. 1369-1378.

Newey, W, Powell, J. and Walker, J. (1990) "Semiparametric Estimation of Selection Models: Some Empirical Results", textitAmerican Economic Review, Papers and Proceedings, 80(2), p. 324-328.

Poirier, D. J. (1980) "Partial observability in bivariate probit models. Journal of Econometrics, v. 12, p. 209-217.

Roy, A. D. (1951) "Some Thought on the Distribution of Earnings", Oxford Economic Papers (New Series), v. 3, n. 2, p. 135-146.

Saavedra, J. and Chong, A. (1999), "Structural reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru", Journal of Development Studies, v. 35, n. 4, p. 95-116.

Schafgans, M. M. A. (1998) "Ethnic Wage Differences in Malaysia: Parametric and Semiparametric Estimation of the Chinese-Malay Wage Gap". Journal of Applied Econometrics, v. 13, p. 481-504.

Tannuri-Pianto, M. and Pianto, D.M. (2002) "Informal Employment in Brazil – a choice at the top and segmentation at the bottom: A Quantile Regression Approach. Mimeo.

Tunali, Insan (1986) "A General Structure for Models of Double-Selection and an Application to Joint Migration/Earning Process with Remigration". In R. Ehrenberg (ed.), Research in Labor Economics. v. 8, Part B, p. 235-282. Greenwich, MA.: JAI Press.

Venti, S. F. (1987) "Wages in the federal and private sectors, in Public Sector Payrolls". In Public Sector Payrolls, ed. D. A. Wise, University of Chicago Press, Chicago, p. 147-182.

Yamada, Gustavo (1996) "Urban Informal Employment and Self-Employment in Developing Countries: Theory and Evidence". Economic Development and Cultural Change, v. 44, n. 2, p. 289-314.

EDITORIAL

Coordenação
Silvânia de Araujo Carvalho

Supervisão
Iranilde Rego

Revisão
Gisela Viana Avancini
Sarah Ribeiro Pontes
Allisson Pereira Souza (estagiário)
Constança de Almeida Lazarin (estagiária)

Editoração
Aeromilson Mesquita
Elidiane Bezerra Borges
Roberto Astorino

Brasília
SBS – Quadra 1 – Bloco J – Ed. BNDES,
10º andar – 70076-900 – Brasília – DF
Fone: (61) 315-5336
Fax: (61) 315-5314
Correio eletrônico: editbsb@ipea.gov.br

Rio de Janeiro
Av. Presidente Antônio Carlos, 51,
14º andar – 20020-010 – Rio de Janeiro – RJ
Fone: (21) 3804-8118
Fax: (21) 2220-5533
Correio eletrônico: editrj@ipea.gov.br

URL: http://www.ipea.gov.br

ISSN 1415-4765

Tiragem: 130 exemplares

Livros Grátis

(http://www.livrosgratis.com.br)

Milhares de Livros para Download:

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo