Alternative Ways of Measuring and Interpreting Worker Flows

Corseuil, Carlos H. L.

July 2009

ALTERNATIVE WAYS OF MEASURING AND INTERPRETING WORKER FLOWS*

Carlos Henrique L. Corseuil

Abstract

The present paper provides empirical evidence compatible with a proposed theoretical framework to explain the joint determination of two components of worker flows: worker replacement and job creation. We show that a negative correlation between job creation and replacement across firms emerges from such a framework. An empirical model is specified and its parameters are estimated taking into account two serious problems: measurement error and endogenous regressor. We take advantage of a matched employer-employee longitudinal database with detailed information on job and worker characteristics to tackle both issues. Our estimates confirm the negative correlation predicted by the theory.

*This material is based on a chapter of my PhD dissertation developed at University College London. I would like to thank Cláudio Ferraz, Sérgio Firpo, Miguel Foguel, Steffen Pischke, Francis Kramarz, Giuseppe Moscarini, Jean-Marc Robin and Eduardo P. Ribeiro for valuable comments on previous versions of this paper. I would also like to thank CAPES for the financial support.

1 INTRODUCTION

The present paper investigates the relationship between worker replacement and job creation at plant level. The link between different components of worker flows have been analyzed in previous studies, such as Davis and Haltiwanger (1998); Albaek and Sorensen (1998); Abowd et al. (1999); and Burgess et al. (2000), which are widely cited papers on this topic. Usually they estimate reduced form equations relating total worker flows to job flows. The relationship between replacement and job creation is explored in the last paper. Nevertheless the present paper brings some important contributions. The first one is to link our empirical analysis to a formal theoretical framework on worker flows, which was not done by any of the papers mentioned above.[1]

The second contribution is to propose alternative measures for worker flows and their components. We explore a recently available matched employer-employee database which contains detailed information on the jobs filled by each worker in order to enhance the standard measurement procedure used in previous papers. Contrasting those figures with the figures based on our new measurement procedure, we show that worker flows levels were underestimated as was the relative size of gross job flows (the sum of job creation and job destruction) as a percentage of worker flows. Moreover we demonstrate that the use of the standard measurement procedure may produce biased results in the estimation of our empirical model, even when the model is

[1]For instance, when Burgess et al. (2000) try to justify the specification for their empirical model they say: "In the absence of a formal model, we simply highlight two issues...".(Burgess et al., 2000)[p. 480]

correctly specified.

The theoretical framework considered in the present paper relies on firms learning the worker $\times$ job match quality and deciding simultaneously about job creation and worker replacement. The basic version of the framework is developed in a companion paper,[2] and briefly outlined in the next section. We argue that a testable hypothesis emerges from such a framework involving a negative relationship between the levels of job creation and replacement across firms. An empirical model is specified to perform this test, in which the specification is guided by the theoretical framework. Our findings confirm that replacement is in fact negatively correlated with job creation, which support the proposed view on the determinants of worker flows and their components.[3] This findings relied on an identification hypothesis which may be considered controversial. It says that there are no time varying unobservable determinants affecting job creation and worker replacement simultaneously.

The final contribution of the present paper is to generalize the theoretical framework to encompass the influence of idiosyncratic productivity shocks on the two dimensions of worker flows analyzed in this paper. In developing this generalized version of the theory, we show that the wage offered to recently hired workers may be used to control for such idiosyncratic shocks. Estimations are then repeated using an employment model that includes this control variable and uses a lighter identification hypothesis. The negative relationship between job creation and replacement is confirmed for this version of the empirical model.

[2] Corseuil (2009)

[3] This evidence adds to other successful predictions derived from the same theoretical framework, most of them on firm dynamics.

The next section summarizes the basic version of the theoretical model being tested and presents the identification problem to be faced by the empirical model. The third section is devoted to the exposition of concepts and measurement procedure to be applied, comparing with the standard measurement procedure. The fourth section introduces the data and comments on some relevant descriptive results. The fifth section comprises the specification and the results of the empirical models grounded on the basic version of the theoretical framework. The sixth section presents the extended version of the theoretical model, as well as the counterpart empirical model with an additional control variable which allows to relax the identification assumption invoked in the fifth section. The last section summarizes our conclusions. We incorporate relevant comparisons with related papers throughout this article.

2 BACKGROUND

In a companion paper, a labor market model is developed where firms decide simultaneously about job creation and worker replacement.[4] The main feature of the model is imperfect information about workers' productivity, which is revealed only after production.

This framework predicts that, conditioned on firm size in the previous period, the higher the number of worker replacements the smaller the number of jobs created.

The mechanism may be summarized as follows. Due to constant returns to scale, firms will always try to hire workers. The decision on how many

[4]The distinction between firm and establishment (or plant) is not relevant to the theoretical model.

vacancies will be posted, and how to distribute them among replacement and new job positions will depend on the revealed quality of matches previously formed. Firms with bad match composition will search for (and tend to have) a higher number of new matches, which mostly will be used to replace workers. Those with a good match composition will search for a lower number of new matches, which will, conversely, be used to fill newly created jobs. The theoretical framework suggests the following empirical model:

$$JC_{p,t} = \beta_1 \cdot x_{p,t-1} + \delta_1 \cdot \xi_{p,t} + \epsilon_{p,t} \tag{1}$$

$$REP_{p,t} = \beta_2 \cdot x_{p,t-1} + \delta_2 \cdot \xi_{p,t} + \eta_{p,t}, \tag{2}$$

where $JC_{p,t}$ and $REP_{p,t}$ stand for job creation and replacement respectively in establishment p at time t. On the right side of both equations, $x_{p,t-1}$ represents the size of the establishment in the previous period and $\xi_{p,t}$ denotes the proportion of high quality matches. So δ_1 and δ_2 are the parameters which describe the relationship between match composition, job creation, and replacement respectively.

The causal relationship predicted between the quality of the matches on the one hand and either replacement or job creation on the other hand can be summarized as:

$$\delta_1 > 0$$

$$\delta_2 < 0.$$

Since quality is not observed, we can derive another prediction based on the observed variables for the purpose of empirical investigations. If we assume that

$$cov(\epsilon_{p,t}, \eta_{p,t} \mid x_{p,t-1}) = 0; \tag{3}$$

then the following result can be derived:[5]

$$sign\{cov(REP_{p,t}, JC_{p,t}) \mid x_{p,t-1}\} = sign\{\frac{\delta_1}{\delta_2}\} < 0. \tag{4}$$

Note that the predicted relationship between replacement and job creation is not a causal one. It arises only through the influence that quality exerts on each of these variables. This result means that strait interpretation of the relationship between these variables without any theoretical background could be misleading. Another source of misleading interpretation of this relationship is measurement error. In the next section we claim that the standard measurement procedure may produce misleading results in our testing procedure. Therefore we propose alternative measurement procedures before turning to the specification of our empirical model and testing procedure that relies on equation 4.

[5]Just isolate $\xi_{p,t}$ in equation 2 and plug it into the equation 1.

3 CONCEPTS AND MEASURES

We quantify worker flows as the number of job positions where worker $\times$ job matches were either formed or dissolved within a time period. Some of these flows correspond to changes in firms' employment structure, both in terms of quantity of workers employed as well as their allocation across occupations. This is the case when a worker fills a newly created job, or where a separation occurrs, and the job closes down. The literature refers to this component of worker flows as gross job flows, or to be more precise, job creation (JC) in the case of match formation and job destruction (JD) in the case of match dissolution.[6] Some other flows correspond to changes in firms' labor force given a fixed employment structure. This is the case when a worker moves to a job position previously filled by another worker. This component of worker flows will be referred to in this paper as replacement (REP).[7]

Note for future reference that worker flows as defined above have a positive dimension (WF^{+}), in jobs where matches were formed, and a negative dimension (WF^{-}) in jobs where matches were dissolved. The following expressions, defined at plant level (p), may be used to clarify the conceptual framework discussed so far:

$$WF^{+}_{p,t} = JC_{p,t} + REP_{p,t}$$

[6] This component of worker flows represents transitions for both workers and jobs. In the case of jobs, the transitions happen from inactive to active or vice-versa.

[7] Note that it represents transitions only for workers, since jobs remain active by definition.

and

$$WF^{-}_{p,t} = JD_{p,t} + REP_{p,t}.$$

Previous papers have proposed alternative measurement procedures of worker flows that could be disaggregated into gross job flows and replacement. The standard measurement procedure is to consider the positive dimension of worker flows (WF^{+}) as the total number of hirings (H) within a time period (usually one year or one quarter), while separations (S) usually correspond to the negative dimension of worker flows (WF^{-}). Concerning the job creation and the job destruction components, one of them corresponds to the net employment growth (ΔN), depending on the sign, and the other should be null by definition. This methodology was used by Albaek and Sorensen (1998) and Burgess et al. (2000) for instance. Their measurement procedure considers replacement (REP) as the difference between either total hires and job creation or total separations and job destruction.[8]

The standard measures for worker flows and the respective components (job creation, job destruction and replacement) will be denoted with a superscript "s" and can be described as:[9]

$$WF^{s+}_{p,t} = H_{p,t},$$

$$WF^{s-}_{p,t} = S_{p,t},$$

[8]They actually use the term labor churning instead of worker replacement.
[9]This notation was chosen arbitrarily as there is no consensus in the literature.

$$JC^s_{p,t} = \Delta N_{p,t} \cdot I(\Delta N_{p,t} > 0),$$

$$JD^s_{p,t} = -\Delta N_{p,t} \cdot I(\Delta N_{p,t} \leq 0),$$

and

$$REP^s_{p,t} = H_{p,t} - JC^s_{p,t} = S_{p,t} - JD^s_{p,t}.$$

Note that there are two drawbacks associated with this procedure. First, the gross job flow measures, and consequently the replacement measures, implicitly require the assumption that occupations are homogeneous within the firm. So if the firm creates a new job position in a given occupation and destroys another position in a distinct occupation in the same time period, this process will not be computed either as job creation or as job destruction, but as a replacement. Another drawback of this procedure is the fact that if multiple match formations and dissolutions take place in a single job position, either as job flow or replacement, then they all will be counted as replacement. The example below illustrates the two drawbacks mentioned above. Table 1 summarizes the evolution of the employment structure in a hypothetical establishment between Decembers of two consecutive years. The two columns shown for each month included in the table correspond to occupational categories, labeled y and z.

Comparing the first and last months of the table we see in reality that one job position was destroyed in occupation y, one was created in z and no worker was replaced. However applying the standard measures described above one would conclude that no jobs were created or destroyed and only two workers were replaced.[10]

We propose an alternative procedure to measure worker flows and their components tackling the two drawbacks of the standard measures discussed above. According to our procedure, worker flows will be measured as *the number of worker × job matches that either existed in t but not in t-1 or the other way around*. We will refer to this measure as adjusted worker flows which can be represented as:

$$WF^{a+}_{p,t} = \sum_j [NM_{p,j,t} - NM^*_{p,j,t}]$$

and

$$WF^{a-}_{p,t} = \sum_j [BM_{p,j,t} - BM^*_{p,j,t}],$$

where $NM_{p,j,t}$ ($BM_{p,j,t}$) denotes all matches formed (dissolved) in t in occupational category j at establishment (plant) p, and $NM^*_{p,j,t}$ ($BM^*_{p,j,t}$) represents those which were dissolved (formed) before the end (after the start) of time t. Note that workers coming from other occupational categories within the same establishment are included in our definitions of $WF^+_{p,t}$ and $WF^-_{p,t}$ but not in $WF^{s+}_{p,t}$ and $WF^{s-}_{p,t}$.

Concerning job creation and job destruction, we use the following definitions, denoted by $JC^a_{p,t}$ and $JD^a_{p,t}$:

$$JC^a_{p,t} = \sum_j (\Delta n_{p,j,t}) \cdot I(\Delta n_{p,j,t} > 0)$$

[10] The conclusion comes from the fact that $\Delta N_{p,t} = 0$ and $H_{p,t} = S_{p,t} = 2$.

and

$$JD^a_{p,t} = \sum_j [-\Delta n_{p,j,t}] \cdot I(\Delta n_{p,j,t} \leq 0).$$

According to this measurement procedure, the definition for replacement becomes:

$$REP^a_{p,t} = WF^{a+}_{p,t} - JC^a_{p,t} = WF^{a-}_{p,t} - JD^a_{p,t}.$$

When applied to our example in Table 1, these measures confirm our understanding that one job was created, one destroyed and no worker was replaced.

Some papers deal with the two amendments described above for the standard measures individually. On the one hand Davis and Haltiwanger (1998) and Abowd et al. (1999) dealt with the time consistency of the measures, implicitly defining worker flows as *the number of worker × establishment matches that either exist in t but not in t-1 or the other way around.*[11] On the other hand, Hamermesh et al. (1996) and Lagarde et al. (1996) dealt with heterogeneous jobs when defining job flow measures.

An important point to bear in mind is that all these measures assume that a job position must be occupied to exist, which is the same as saying that they share the assumption that there is no vacant position at the time the stocks are computed. This is a strong assumption guided by the lack of

[11] Davis and Haltiwanger (1998) explicitly formulated the following concept for worker flows: the total number of workers "whose place of employment or employment status differs between t-1 and t." They refer to this concept as gross worker reallocation. They were able to compute the measures only at an aggregate level due to data constraints. Abowd et al. (1999) were able to implement worker flows measures at the establishment level, which they denoted as "entry (rate) excluding within year entries" and "exit (rate) excluding within year entries".

information about vacant positions.[12]

Finally, our measure for replacement is lower while the ones we use for job flows (job creation and destruction) are higher than the standard ones. That is, we may specify the relationship between our measure for a worker flows component and the respective counterpart following the standard procedure as follows:

$$JC^a_{p,t} = JC^s_{p,t} - \varepsilon^{jc}_{p,t}$$

and

$$REP^a_{p,t} = REP^s_{p,t} + \varepsilon^{r}_{p,t},$$

where $\varepsilon^{jc}_{p,t}$ and $\varepsilon^{r}_{p,t}$ are non-negative quantities capturing measurement errors. Moreover we claim that the standard measurement procedure may produce misleading results in our testing procedure. In order to see this, first consider the following relationship between the REP and JC covariance in the two alternative measurement procedures:

$$\begin{aligned} cov(REP^s_{p,t}; JC^s_{p,t} \mid x_{p,t-1}) = cov(REP^a_{p,t}; JC^a_{p,t} \mid x_{p,t-1}) + \\ cov(REP^a_{p,t}; -\varepsilon^{jc}_{p,t} \mid x_{p,t-1}) + cov(\varepsilon^{r}_{p,t}; JC^a_{p,t} \mid x_{p,t-1}) + \\ cov(\varepsilon^{r}_{p,t}; -\varepsilon^{jc}_{p,t} \mid x_{p,t-1}). \end{aligned} \tag{5}$$

In Appendix B we show that on the one hand both terms in the second line are positive, while on the other hand the last term is negative. Therefore,

[12] The paper of Yashiv (2000) for Israel is the only one that I am aware of having this information available.

whatever the sign of the first term on the right side, the opposite sign can show up on the left side. This is to say that the sign of the estimated covariance between worker replacement and job creation may differ from the real one if one uses the standard measurement procedure. Therefore this measurement procedure should be avoided in testing procedures based on the sign of the covariance between job creation and replacement, as the one we will specify later.

4 DATA AND DESCRIPTIVE RESULTS

4.1 Data

Our data come from a Brazilian administrative file (Relação Anual de Informações Sociais - RAIS) maintained by the Brazilian Ministry of Employment and Labor (Ministério do Trabalho e Emprego - MTE). All registered tax paying establishments must send to the Ministry information on employees who worked anytime during the reference year.[13]

The RAIS information provides a matched employer-employee longitudinal database, similar to those available in developed countries.[14] The information available in this database includes that specific to workers (such as gender, age and schooling), to establishments (such as location, industry category and type of ownership), and to the contracted relationships (such as wage, hours, dates and reason for hirings and separations). The distin-

[13] The absence of tax evaders prevents us from claiming that the data refer to the universe of Brazilian establishments.

[14] See Abowd and Kramarz (1999) for a description of the countries where this type of database was available then and how research on labor economics has benefited from such database.

guishing characteristic of these data comes in this last dimension, where we can find detailed information on workers' occupations.

This database allows us to list workers' identification code in each of the establishments' occupational categories for consecutive years. We then are able to identify match formation and dissolution at this level of observation, which is actually the procedure employed in our codes.[15] We take full advantage of the extra information available in the RAIS, but not available in other data sets used in the measurement of worker flows.[16]

We use 3-digit occupational categories within each establishment as the empirical counterpart of jobs.[17] We share the view expressed in Moscarini and Thomsson (2007) that "...measuring occupations at this level corresponds most closely to the notion of labour technology, with labour input being differentiated by the tasks involved...."

The use of such disaggregate categorization may raise concerns regarding the presence of measurement errors. It might happen that in consecutive years the same establishment classifies a worker performing the same tasks in different but closely related job categories. To address these concerns we analyze the robustness of our main results using two alternative procedures. The first excludes any movements within establishments; the other uses the 2-digit occupational category as the job classification. It worth is worth men-

[15]The procedure is equivalent to, although not the same as, the one described in the expressions for $WF^{+}_{p,t}$ and $WF^{-}_{p,t}$.

[16]Some data sets have the matched employer-employee structure but do not register occupation information; others, while presenting information on occupation, lack either worker or establishment identification code.

[17]This categorization is closely related to the 3-digit version of the International Standard Classification for Occupations (ISCO-88). See Muendler et al. (2004) for more details on the categorization used in this paper (CBO-94) and the ISCO-88.

tioning the fact that the Brazilian MTE has been working to assure that the collected RAIS data are accurate, if for no other reason than because certain labour and pension regulations in Brazil are linked directly to employee characteristics and occupation.[18]

We will use information from 1994 to 2001. Although information is available from 1986, there are clear and specific reasons not to include the whole period. First, there was an upward trend in coverage in the late 1980s. Moreover, the recent availability of some variables and changes in the definition of others in 1994 hinders comparisons.[19] The sample size in the original data set is about 2 to 2.5 million registered establishments per year, but we exclude some industries from the sample. We analyzed approximately 6.5 million year $\times$ establishments observations from the 3 industries for which we have solid information: manufacturing, services and trade.[20]

4.2 Aggregate Results

Table 2 shows the figures for each of the above defined components of worker flows using the alternative measurement procedure proposed in this paper and the standard one.

The first two columns report the numbers for the new measurement procedure as percentages of the average employment level. According to this

[18]In 1994, when introducing the CBO-94 into the RAIS, the MTE promoted an information campaign on this classification system among employers.

[19]It is also possible to claim that yet another reason is provided by some structural changes in the Brazilian economy in the early 1990s.

[20]Mining, utilities, health, education, public sector and social services were excluded due to a massive concentration of state operated companies, while agricultural establishments were excluded due to coverage problems and construction was excluded due to its idiosyncratic worker flows and labor relations. Additional screening procedures applied in the original data set are described in Appendix A.

method, new matches corresponded to 45% of average employment level, 17% due to replacement and 28% due to newly created jobs. Results are similar for dissolved matches, which amount to 43% of average employment level, with a job destruction rate of 26%. From these numbers we can calculate that 38% (40%) of new (dissolved) matches come from replacement.[21] These figures are considerably different from their counterparts computed using the standard measurement procedure, shown in the last two columns of Table 2. As mentioned at the end of Section 3, job flow percentages are lower and replacement percentages higher when the standard measurement procedure is implemented.

It is interesting to note that the relative importance of replacement is around 70% when using the standard measure, which is similar to the magnitude reported in other papers employing this measurement procedure.[22] For example, Burgess et al. (2000) applied the standard measurement procedure to quarterly data and calculated the replacement share to be 70% of the employment level, while Albaek and Sorensen (1998) identified a 60% replacement share using annual data. [23]

We had argued that the application of our measurement procedure should define a lower boundary for the replacement rate among those mentioned in the previous section, but the huge differences between our measurement procedure and the standard one show that improving the measurement of

[21] The 38% (40%) comes from the ratio 17/45 (17/43).

[22] 70% approximates both ratios: 39/57 and 39/53.

[23] Orellano and Pazello (2006) also apply the standard methodology to the same data source we use. Due to differences in spatial and temporal coverage, their results are not the same as the ones we report above. In their calculations, replacement corresponds approximately to 62% of worker flows.

worker flows and their components is not a minor point. The test to be performed later could be affected severely by measurement inconsistencies of this magnitude. We will come back to this point later.

4.3 Job Flows and Replacement Frequencies

An important feature of the data for our empirical analysis in the next section is the frequency of simultaneous job flows and replacement at the establishment level. This information can be retrieved from figure 1, which reports the cross-establishment distribution of all possible combinations of worker flows components.

Three relevant points are to be noted from this figure. The first is that 51% of establishments combine at least two of the three possible actions: replacement, job creation, and job destruction. Moreover 20% combine all the three actions simultaneously. This fact corroborates the idea that current theoretical models on worker flows addressing only one action might be missing important features.

The second relevant point is that 34.4% of establishments combine job creation and job destruction simultaneously. This is not captured in the standard measurement procedure for worker flows, which corroborates the idea of getting misleading results when employing that procedure.

The last relevant point is the relatively low share of establishments that combine replacement with job destruction. This is less frequent than combining replacement with job creation. In fact, this is the least frequent combination involving any two actions. This is relevant since the theoretical framework that grounds our analysis does not consider this combination.

5 THE BENCHMARK EMPIRICAL MODEL

5.1 Specification and Testing Procedure

Combining the information provided by the theoretical model with the considerations above allows us to build the following empirical model for establishment level observations:

$$JC_{p,t} = \gamma_1' \cdot X_{p,t} + \mu_{p,t}$$

and

$$REP_{p,t} = \gamma_2' \cdot X_{p,t} + \zeta_{p,t}.$$

$X_{p,t}$ represents the set of observable variables to be defined later, and $\mu_{p,t}$ and $\zeta_{p,t}$ represent the non-observed components which encompass the following terms already defined in Section 2:

$$\mu_{p,t} = \delta_1 \cdot \xi_{p,t} + \epsilon_{p,t}$$

$$\zeta_{p,t} = \delta_2 \cdot \xi_{p,t} + \eta_{p,t}.$$

As shown in equation 4, one testing procedure to confirm the main prediction of our theoretical framework is to check the sign of $cov(JC_{p,t}, REP_{p,t} \mid X_{p,t})$. In order to estimate the sign of the aforementioned covariance it is more convenient to re-arrange the empirical model in the following way:

$$JC_{p,t} = \delta \cdot REP_{p,t} + \gamma' \cdot X_{p,t} + \nu_{p,t}. \tag{6}$$

Our strategy consists of using the estimated sign for δ to test our null hypothesis, which is:

$$Ho: \frac{\delta_1}{\delta_2} < 0.$$

The link between these parameters is established below:

$$sign\{\delta\} = sign\{cov(JC_{p,t}, REP_{p,t} \mid X_{p,t})\} = sign\{\frac{\delta_1}{\delta_2}\}.$$

The second equality replicates equation 4 in the context of the empirical model. It is important to note that this second equality holds under an identification assumption related to some non-observable components described in equation 3, replicated below in terms of our empirical model as

$$cov(\epsilon_{p,t}, \eta_{p,t} \mid X_{p,t}) = 0.$$

The validity of this identification assumption depends on an appropriate set of control variables. To define this set, we have relied not only on the theory but also on the related empirical literature. Among empirical papers dealing with worker flows, as far as our survey could establish, the one with the closest related empirical analysis is Burgess et al. (2000). An important point raised by their paper is the relevance of establishment fixed effects as a determinant for both job flows and replacement. The authors interpret this variable as capturing idiosyncratic personal policies. Albaek and Sorensen

(1998) also make this point when estimating a model where the relationship between job flows and replacement is specified with replacement (or churning as they call it) as the dependent variable.

Other empirical models investigate the determinants of job flows[24] not considering the relationship with replacement. Although not comparable to our model, these analyses point to other potentially relevant determinants for job flows. They usually include plant characteristics among explanatory variables, such as age, size and industry category, with the last two variables being used as proxies for technology. Some establishments may be taking advantage of positive shocks related to technology, such as technological progress, and would tend to create a relatively high number of jobs anyway. Concerning age, some models in the industrial organization literature claim that younger firms are still learning about their optimal scale or their capabilities and therefore may be less reluctant to create new jobs. Results from these studies highlight the importance of age and size as the most relevant plant characteristics, whereas industry category and time effects also contribute to explaining job flows.

Based on these facts, we define the set of control variables as

$$X_{p,t} = \{x_{p,\bar{t}}, AGE_{p,t}, \alpha_p, i * t\},$$

where $x_{p,\bar{t}}$ represents the establishment average size across t and $t-1$, $AGE_{p,t}$ is a binary variable for establishments older than 3 years, α_p captures establishment-specific fixed effects, and $i * t$ denotes interactions between

[24]See Davis and Haltiwanger (1999) for a survey of these papers.

dummy variables for year (t) and industry (i) categories at the two-digit level.[25]

Going back to our identification assumption, it would be valid if, for instance, there is no other non-observable time-varying establishment effect that is simultaneously correlated with both replacement and job creation decisions, apart from the quality of the matches. However, one can argue that this is too restrictive, since firms may be vulnerable to idiosyncratic profitability/productivity shocks, which constitute a non-observable time-varying establishment effect. Although the time dimension of our panel data is short, we will consider explicitly the influence of such shocks in our results later.

It is worth mentioning that the empirical models specified in other papers, cited above, differ from ours on some specification issues. First, their dependant variable is usually defined as a rate relative to the establishment employment level, however, the theoretical predictions to be tested with our model refer to the replacement level and the job creation level. The size measure itself is another delicate point and deserves additional comment. Davis and Haltiwanger (1999) claim that the size effect on job flows is very sensitive to whether it corresponds to the initial size or the average size between two consecutive years. They argue in favor of the latter specification to avoid bias due to the "regression to the mean" effect. We follow their recommendation.

[25]A continuous age variable is constructed based on establishment first appearance from 1992 to 2001. If this first appearance happens to be in 1992 then this variable is coded as censored. As the first period used to measure worker flows is from 1994 to 1995 there are uncensored values up to 3 years old for this period. This was the main reason why 3 was chosen as the limit to split a binary age variable.

5.2 Results from Parameters Estimation

Table 3 presents the estimated coefficients for equation 6 employing the within or fixed effect estimator for the samples defined by each of the three broad industry categories shown in the table. Apart from the exclusion of some industry categories from our sample, as mentioned before, we also exclude birth and death episodes in the above analysis. The reason is obvious in the case of births: These establishments are unable to choose to replace any worker, since job creation corresponds to total employment by definition. The exclusion of death episodes is motivated by the theoretical framework, which emphasizes that the choice of job creation and replacement levels is conditioned on the survival of the firm.

According to the model prediction, we would expect a negative and significant coefficient for $REP_{p,t}$. It should be noted that results using the proposed measurement procedure confirm the prediction of the theoretical framework since the replacement coefficient is always negative. As shown in table 3 each replacement is associated with 0.28 fewer jobs created in the service industry, 0.41 in manufacturing and 0.60 in trade.

The analysis of the coefficients related to size and age shows that older establishments tend to create fewer jobs, which is a standard result in both empirical labour economics and empirical industrial organization literature. Comparisons involving effect of establishment size are not so straightforward since the other papers apply alternative specifications, as previously mentioned. Our specification is comparable to the one employed by Davis and Haltiwanger (1999). As in their paper, we find a positive effect for this variable.

It is worth mentioning that the negative signs for the replacement coefficients contrast with those obtained by Burgess et al. (2000) when they regressed job flows rates on establishment fixed effects as well as contemporaneous and lag transformations of churning (replacement) flow rates. The coefficient of the contemporaneous replacement is positive. It should be stressed that they use the standard measurement procedure for job flows and replacement, which not only differ from our measures but also may induce misleading results for this estimation procedure as previously explained. In fact, when using the standard measurement procedure with our own data we also obtain positive, although small, estimated values for the replacement coefficient as shown in Table 4.

In Appendix C we show that the qualitative results are robust to refinements in our proposed measurement procedure related to worker transitions across jobs within the same establishments.

5.3 Where the Data Better Fit the Theory

The mechanism of job creation under investigation may co-exist with other determinants of this variable. For instance, theoretical frameworks from the field of industrial organization point to the relevance of successful investments either in innovative activities [Klette and Kortum (2004)] or quality improvements [Ericson and Pakes (1995)]. In this section we will evaluate whether our data better fit the sub-samples that tend to be relatively less affected by these alternative driving forces of firm dynamics.

The first evidence comes from table 3, where it is possible to see that the worst goodness-of-fit indicator (lowest R^2) appears for the manufacturing sub-sample. Apart from the industry classification, we also expect that the prevalence of the proposed mechanism may vary according to the size and age of the establishment. The reason for the influence of establishment size is similar to the rationale for the influence of industry classification and consists of the assumption that the more complex the production process the more important is the role of innovation and quality improvement efforts. Therefore, we would expect that the data would fit better a sub-sample of small establishments, which is confirmed by the numbers in table 5. According to this table, R^2 increases at least 10 percentage points for service and manufacturing, compared with estimations using the whole sample.

Concerning establishment age, since the learning process about match quality drives the decisions regarding worker flows, we would expect the data to better fit the theory in a sub-sample of young establishments. Table 6 shows the estimated values for a sub-sample of establishments in either their second or third year of life. The increase in the goodness-of-fit indicator is striking for all three industry categories when compared to results for the whole sample. Also striking is the increase in the absolute value of the replacement coefficient which used to vary between -0.28 and -0.60, and now between 0.52 and 1.82.

In Appendix D we show that the negative coefficient for replacement does not depend on having either small or young establishments in the sample. Sub-samples excluding either young or small establishments still provide negative estimated values for this coefficient.

6 INTRODUCING IDIOSYNCRATIC PRODUCTIVITY SHOCKS

The results in the previous section relied on the assumption that no other unobserved component of worker flows that varies over time at firm level, apart from the quality of the matches, should induce a negative correlation between job creation and replacement. To be more precise, we relied on the following assumption: $cov(\epsilon_{p,t}, \eta_{p,t} \mid X_{p,t}) = 0$.

One of these unobserved components may be idiosyncratic shocks to firms' profitability. In what follows we will specify a new version of the theoretical framework to allow for such shocks, then estimate the counterpart version for the empirical model where a proxy variable for these shocks is added. We will justify the choice of such a variable along with the presentation of the theoretical framework.

6.1 Extending the Theoretical Framework

In this section we adapt the wage bargaining version of the theoretical framework developed in Corseuil (2009) to encompass idiosyncratic productivity shocks. Since the empirical model captures the hiring decision faced by firms in the beginning of the second period, we will focus the analysis on this part of the theoretical model. First, we need to introduce a firm component of the match productivity, denoted by $A_{i,t}$. This component evolves according to a transition rule given by $G(A_{i,t+1} \mid A_{i,t})$, where the value is revealed at the beginning of the period. We assume that $A_{i,t}$ influences revenue in a standard fashion, pre-multiplying the product price (p). We denote the product of price and the firm's productivity component by $A'_{i,t}$. The expected profit at the beginning of the second period can easily be derived substituting $A'_{i,t}$ for p in the corresponding expression developed in Corseuil (2009) for the wage bargaining version of the theoretical framework. The adapted version becomes:

$$E_{\xi_2} E_{\mu_2}[\pi_2(r, jc \mid \xi_1, \ell_1)] = (1-\beta) \cdot \{A'_{i,t} \cdot [s + \theta \cdot (1-s)] - b\} \cdot m(\gamma_2) \cdot (jc + r) -$$
$$c^{jc}(jc + r) + (1-\beta) \cdot \ell_1 \cdot [(A'_{i,t} \cdot \xi_1 - b) + A'_{i,t} \cdot \theta \cdot (1-\xi_1) - b]$$
$$+ [c - (1-\beta) \cdot (A'_{i,t} \cdot \theta - b)] \cdot r.$$

The solution, which can also be derived easily employing the same strategy, becomes:

$$\underline{\ell''-} = \frac{z_2''-}{2 \cdot a}$$

and

$$\overline{\ell''-} = \frac{z_2''- + k''-}{2 \cdot a}$$

where $z_2''-$ and $k''-$ correspond to:

$$z_2''- = (1-\beta) \cdot \{A'_{i,t} \cdot [s + \theta \cdot (1-s)] - b\} \cdot m(\gamma_2) - c$$

and

$$k''- = [c - (1-\beta) \cdot (A'_{i,t} \cdot \theta - b)].$$

It is easy to see that conditioned on a given value of the firm productivity component ($A'_{i,t}$), the negative relationship between job creation and replacement holds in the same way as considered in previous sections. Although we can not directly observe variations in firms' productivity components, such variations can be inferred through variations in wages offered for new employees. The definition for these wages was given by the following expression in Corseuil (2009):

$$w^e = \beta \cdot p \cdot [s + (1 - s) \cdot \theta] + (1 - \beta) \cdot b.$$

So, when allowing for idiosyncratic productivity shocks, it becomes:

$$w^e = \beta \cdot A'_{i,t} \cdot [s + (1 - s) \cdot \theta] + (1 - \beta) \cdot b.$$

Therefore variations on entrant wages can only be caused by variations on $A'_{i,t}$, since all other parameters are constant.

6.2 Estimations Controlling for Entrant Wages

In this section we re-estimate the empirical model adding the average wage for recently hired employees in order to control for productivity shocks. The idea is that firms will offer higher wages for entrants when their idiosyncratic productivity performance is relatively better. Since the quality of the matches is not revealed for entrant workers, the proxy will leave the quality component as part of the residual terms, as specified in Section 5.2.

The results to be discussed in Table 7 are derived from the following model specification:

$$JC_{p,t} = \delta \cdot REP_{p,t} + \theta \cdot w^{e}_{p,t} + \gamma' \cdot X_{p,t} + \xi_{p,t},$$

where $w^{e}_{p,t}$ represents the average wage for those workers hired throughout year t.[26] Now the identification hypothesis becomes: There is no other non-observable component, apart from profitability shocks and quality of the matches, which possibly induces a negative correlation between job creation and replacement.

One can see from Table 7 that the negative correlation between replacement and job creation persists even when we control for a proxy of profitability shocks (wage paid for recently hired workers). The magnitude of the estimated values for the replacement coefficient increases slightly in the two non-manufacturing broad industry groups considered in our analysis and decreases more significantly in the manufacturing group. The effects of size and age also change very little in the non-manufacturing groups and more significantly in the manufacturing group. Finally, the effect of the initial wage is statistically not different from null in all three cases, as can be seen from the columns showing the respective t-values.

It should be noted that the number of observations decreased significantly in all three cases since the latter model is estimated only with establishments that hired at least one new worker in more than one year.[27] In principle this reduction could raise concerns about sample selection problems. However, the lack of variation in the coefficients between alternative model specifi-

[26]The wage is measured either at December of year t or at the moment of the separation if it happens before December.

[27]This last restriction is due to the inclusion of fixed effects.

cations suggests that, at least for the non-manufacturing groups, sample selection should not be an issue.

Overall, the results in this section show that the negative correlation between replacement and job creation is robust to an alternative and less restrictive identification hypothesis. The small changes toward greater magnitude of this effect among non-manufacturing samples suggest that the profitability shocks induce positive correlation between replacement and job creation. Therefore this element could not be responsible for the negative correlations that we have reported in the benchmark specification (at least for the service and trade industries).

7 SUMMARY AND CONCLUSIONS

The present paper investigated whether and how replacement and job creation are correlated. It does so through an empirical model, based on a theoretical framework developed in a companion paper and recommendations from the empirical literature on determinants of gross job flows. One prediction that comes from the theory is a negative correlation between replacement and job creation. The estimation results confirmed the prediction that establishments with higher replacement levels tend to have lower levels of job creation, conditioned on the relevant set of controls.

We generalize the theoretical framework encompassing idiosyncratic productivity shocks. Another version of the empirical model is then developed where we used a proxy variable to idiosyncratic productivity shocks to relax an identification assumption of the previous version. The results were confirmed in this extended version of the empirical model as well as in some robustness checks done in our measurement procedure.

A rich database with a matched employer-employee structure allowed another important contribution of the paper, which is to propose an alternative procedure to the measurement of the components of worker flows. The measurement procedure explored the availability of detailed information on the occupational category of each worker to track the occupational structures of the establishments. We argued in favor of our proposed measurement procedure by showing that the use of the standard measurement procedure may produce biased results in the estimation of our empirical model, even when this model is correctly specified. Finally, such database also allows the introduction of relevant control variables in the empirical model, such as establishment fixed effects and time varying establishment characteristics.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA

Table 1: The evolution of employment structure in a hypothetical establishment

$DECEMBER_{t-1}$		$JUNE_t$		$DECEMBER_t$	
y	z	y	z	y	z
A	D	A	D	A	D
B		B		B	C
C		C			
		E			
		F			

Note:
y and z denote occupational categories, while A to F identify workers.

Table 2: Worker Flows 1994-2001 (%)

	adjusted measure		**standard measure**	
	positive	*negative*	*positive*	*negative*
worker flows	**45.3**	**43.1**	**56.8**	**52.8**
replacement	17.0	17.0	39.3	39.3
job flows	28.2	26.1	17.5	13.5

Table 3: Regression results for the benchmark specification using the proposed measurement procedure for worker flows components

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.28	-149.5	-0.41	-125.3	-0.60	-353.7
size	0.23	324.7	0.20	191.5	0.34	517.7
dummy age > 3	-0.34	-5.9	-0.72	-6.4	-0.24	-20.1
n.obs	2,267,550		1,095,565		3,131,188	
R^2	0.52		0.47		0.57	

Notes:
(1) Dependent variable : job creation
(2) Model specification includes establishment fixed effects and interactions of years and industry categories.

Table 4: Regression results for the benchmark specification using standard measurement procedure for worker flows components

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	0.04	111.6	0.02	24.3	0.08	145.9
size	0.07	150.0	0.07	114.4	0.05	120.4
dummy age > 3	-0.50	-12.1	-1.11	-14.8	-0.25	-29.3
n.obs	2,267,550		1,095,565		3,131,188	
R^2	0.52		0.47		0.57	

Notes:

(1) Dependent variable : job creation

(2) Model specification includes establishment fixed effects and interactions of years and industry categories.

Table 5: Regression results for the benchmark specification - Small plants

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.35	-200.1	-0.33	-110.8	-0.24	-187.1
size	0.31	349.3	0.28	217.4	0.25	369.2
dummy age > 3	-0.29	-40.8	-0.47	-31.4	-0.19	-42.7
n.obs	2,178,742		1,001,288		3,084,876	
R^2	0.62		0.60		0.60	

Notes:
(1) Dependent variable : job creation
(2) Model specification includes establishment fixed effects and interactions of years and industry categories
(3) Small Plants are those employing less than 50 workers.

Table 6: Regression results for the benchmark specification - Young plants

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-1.13	-175.1	-1.82	-153.8	-0.52	-106.4
size	0.47	131.3	0.63	102.3	-0.09	-32.7
n.obs	576123		276837		920784	
R^2	0.78		0.79		0.78	

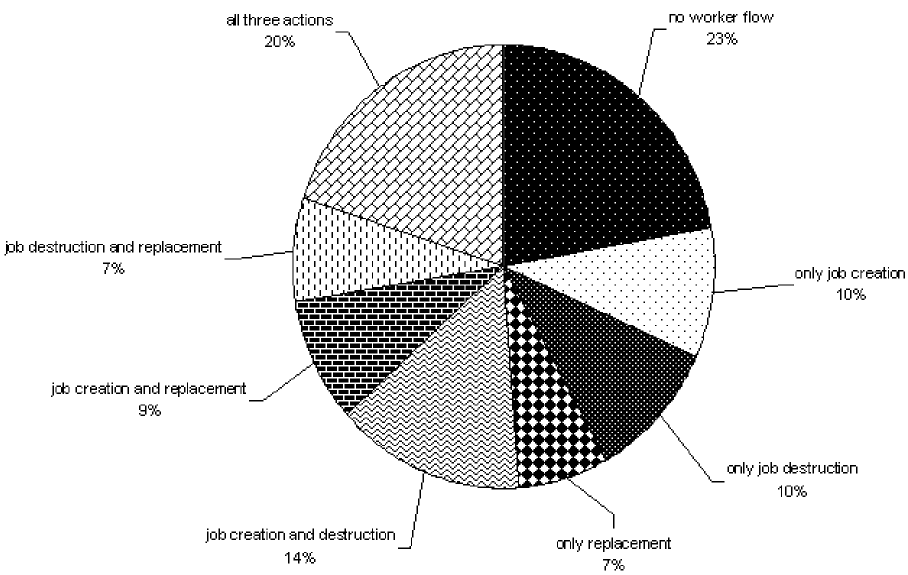
Notes:
(1) Dependent variable : job creation
(2) Model specification includes establishment fixed effects and interactions of years and industry categories
(3) Younger Plants are those operating for at most three years.

Table 7: Regression results for the generalized specification controlling for firms' initial wage

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.34	-92.2	-0.29	-59.0	-0.62	-200.7
entrant wage	-0.04	-0.3	-0.11	-0.5	-0.02	-0.6
size	0.23	159.0	0.15	85.1	0.34	243.7
dummy age > 3	-0.37	-2.6	-0.48	-2.1	-0.25	-7.5
n.obs	955285		556981		1331941	
R^2	0.53		0.52		0.57	

Notes:
(1) Dependent variable : job creation
(2) Model specification includes establishment fixed effects and interactions of years and industry categories
(3) The sample is restricted to establishments which have hired new workers during year t.

Figure 1: Distribution of possible combination of worker flows components across establishments



Appendices

A DATA CLEANING

The following screening procedures were applied to the original data:

- deletion of individuals with invalid identification codes (missing or zero).
- deletion of establishments with invalid identification codes (missing or zero).
- deletion of establishments with discontinuous data reporting from 1992 to 2002. This avoided overestimation of job creation and job destruction figures due to establishments that, although in operation, failed to have their information processed in a particular time period.
- Union of the following duplicated job codes: 073 and 193 - Social worker; 074 and 194 - Psychologist; 093 and 110 - Accountants; 162 and 454 - Decorator.

B THE RELATIONSHIP BETWEEN $COV(JC^S_{P,T}, REP^S_{P,T} \mid X_{P,T})$ AND $COV(JC_{P,T}, REP_{P,T} \mid X_{P,T})$

In this section we investigate how the signs of $cov(REP^s_{p,t}; JC^s_{p,t} \mid x_{p,t})$ and $cov(REP^a_{p,t}; JC^a_{p,t} \mid x_{p,t})$ are related to each other. We depart from expression (5) reproduced below, omitting the conditioning to simplify the notation.

$$cov(REP^s_{p,t}; JC^s_{p,t}) = cov(REP^a_{p,t}; JC^a_{p,t}) + cov(REP^a_{p,t}; -\varepsilon^{jc}_{p,t}) + cov(\varepsilon^r_{p,t}; JC^a_{p,t}) + cov(\varepsilon^r_{p,t}; -\varepsilon^{jc}_{p,t})$$

Our claim is that $cov(REP^s_{p,t}; JC^s_{p,t})$ and $cov(REP^a_{p,t}; JC^a_{p,t})$ may have opposite signs. This happens because of the divergent signs among the last three components in the expression above. We develop below each of these components in order to derive their respective signs.

$$cov(REP^a_{p,t}; -\varepsilon^{jc}_{p,t}) = cov[REP^a_{p,t}; (JC^s_{p,t} - JC^a_{p,t})] = 0;$$

The second equality follows because the replacement does not affect employment growth either at the establishment level (JC^s) or at the occupational category level (JC^a).

$$cov(\varepsilon^r_{p,t}; JC^a_{p,t}) = cov[(REP^s_{p,t} - REP^a_{p,t}); JC^a_{p,t}] > 0;$$

The inequality holds since REP^s counts part of JC^a as replacement.

$$cov(\varepsilon^r_{p,t}; -\varepsilon^{jc}_{p,t}) = cov[(REP^s_{p,t} - REP^a_{p,t}); (JC^s_{p,t} - JC^a_{p,t})] < 0;$$

The same relationship mentioned above for REP^s and JC^a drives this inequality. Therefore, the combination of a null, a positive and a negative term may flip the sign of $cov(REP^a_{p,t}; JC^a_{p,t})$.

C ROBUSTNESS ANALYSIS CONCERNING ALTERNATIVE MEASURES

C.1 Restricting Within Plant Worker Flows

The implementation of the new measurement procedure requires a strategy to identify a job to job movement within the same establishment. We use the following identification assumptions:

i) Any movement across 3 digit occupational categories (within the establishment) corresponds to a job change, i.e. the task performed by the worker changes; and,

ii) Tasks are homogeneous within a 3 digit occupational category, which means that there can be no movement across jobs within the same occupation.

We proceed with relaxing the first assumption in two ways where we admit some movements captured in the data may be due to measurement error. First, we use the 2-digit codification for job categories. The aim of this procedure is to analyze how the results would vary if we assume that all within establishment movements across "similar" job categories are due to measurement error.

The regressions results analogous to Table 3 based on these measures are shown in Table 8. One can see that we get estimated coefficients extremely close to those shown in Table 3.

Since one may think that measurement errors are not restricted to "similar" job categories, we can restrict even further the within plant worker flows to check the robustness of our results. We then change the second assump-

Table 8: Regression results for the benchmark specification - 2-digit job categories

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.29	-159.6	-0.40	-125.9	-0.56	-338.8
size	0.22	315.0	0.18	185.5	0.33	494.6
age > 3	-0.34	-6.1	-0.80	-7.3	-0.25	-20.8
n.obs	2,267,550		1,095,565		3,131,188	
R^2	0.50		0.45		0.55	

Notes:
(1) Dependent variable : job creation
(2) Model specification includes establishment fixed effects and interactions of years and industry categories.

tion using alternative measures that remove any within plant movements. These measures can be defined as:

$$WF^{\diamond+}_{p,t} = \sum_j [H_{p,j,t} - H^*_{p,j,t} - H^{\diamond}_{p,j,t}]$$

and

$$WF^{\diamond-}_{p,t} = \sum_j [S_{p,j,t} - H^*_{p,j,t} - S^{\diamond}_{p,j,t}],$$

where $H^{\diamond}_{p,j,t}$ ($S^{\diamond}_{p,j,t}$) denotes all matches formed (dissolved) in which workers came from (went to) other positions in the same establishment.

Concerning job flows, we apply analogous procedures and define the following measures:

$$JC^{\diamond}_{p,t} = \sum_j \Delta n^{\diamond}_{p,j,t} \cdot I(\Delta n^{\diamond}_{p,j,t} > 0)$$

and

$$JD^{\diamond}_{p,t} = \sum_j \mid -\Delta n^{\diamond}_{p,j,t} \mid \cdot I(\Delta n^{\diamond}_{p,j,t} \leq 0),$$

where

$$\Delta n^{\diamond}_{p,j,t} = WF^{\diamond+}_{p,j,t} - WF^{\diamond-}_{p,j,t}.$$

Note that, although $\sum_j \Delta n^{\diamond}_{p,j,t} = \Delta n_{p,t}$, the equality does not necessarily hold for each component of this summation, i.e., in general we have $\Delta n^{\diamond}_{p,j,t} \neq \Delta n_{p,j,t}$. It follows that replacement is measured as:

$$REP^{\diamond}_{p,t} = WF^{\diamond+}_{p,t} - JC^{\diamond}_{p,t} = WF^{\diamond-}_{p,t} - JD^{\diamond}_{p,t}.$$

The regression results analogous to Table 3 based on these measures are shown in Table 9. One can see that such a procedure did not change our main results qualitatively. The replacement coefficients are now only slightly lower, in terms of absolute value, and remain smaller than one. The size and age coefficients also do not change substantially.

D RESULTS FOR LARGER AND OLDER ESTABLISHMENTS

Given that the sample is dominated by small establishments, one might question whether the mechanism described in the theoretical framework also holds for a sample excluding small establishments. We therefore repeat our basic analysis restricting the sample to establishments with more than 50 employees.

Table 10 presents the results for this sub-sample in each of the three broad

Table 9: Regression results for the benchmark specification - Restricting any movement within plant

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.24	-143.5	-0.40	-125.3	-0.40	-246.6
size	0.23	296.6	0.15	175.4	0.21	377.2
age > 3	-0.30	-47.2	-0.90	-8.8	-0.24	-22.9
n.obs	2,178,742		1,095,565		3,131,188	
R^2	0.60		0.42		0.53	

Notes:
(1) Dependent variable : job creation
(2) Model specification includes establishment fixed effects and interactions of years and industry categories.

industry categories. The results reveal that replacement affects job creation in large establishments in the same way we described before. The coefficients are all negative, significant, and have absolute values lower than one, as was the case for the complete sample.

Table 10: Results from the empirical model for job creation - Large plants

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.29	-29.5	-0.43	-39.7	-0.68	-46.7
size	0.22	58.8	0.19	55.2	0.32	56.4
age > 3	-13.15	-7.1	-16.64	-9.6	-9.15	-9.4
n.obs	88,808		94,277		46,312	
R-Square	0.49		0.44		0.56	

Notes:.
(1) Model specification includes establishment fixed effects and interactions of years and industry categories.

Analogous estimations were carried out for a sub-sample of establishments in at least their fourth year of existence. Table 11 presents the results for this sub-sample in each of the three broad industry categories; the qualitative results are maintained.

Table 11: Results from the empirical model for job creation - Old plants

	service		**manufacturing**		**trade**	
	Coef.	*tValue*	*Coef.*	*tValue*	*Coef.*	*tValue*
replacement	-0.23	-109.5	-0.32	-88.3	-0.59	-306.4
size	0.21	268.3	0.18	161.9	0.35	475.7
n.obs	1691427		818728		2210404	
R-Square	0.52		0.45		0.58	

Notes:.
(1) Model specification includes establishment fixed effects and interactions of years and industry categories.

References

J. Abowd, P. Corbel, and F. Kramarz. "The entry and exit of workers and the growth of employment: an analysis of French establishments". *Review of Economics and Statistics*, 81:170–187, 1999.

J. Abowd and F. Kramarz. *The Analysis of Labor Market Using Matched Employed-Employee Data*, volume 3B of *Handbook of Labor Economics*. North Holland, 1999. editors Orley Ashenfelter and David Card.

K. Albaek and B. Sorensen. "Worker Flow and Job Flow in Danish Manufacturing, 1980-91". *The Economic Journal*, 108:1750–1771, 1998.

S. Burgess, J. Lane, and D. Stevens. "Job Flow, Worker Flow and Churning". *Journal of Labor Economics*, 18:473–502, 2000.

C. H. Corseuil. "Worker Flows and Firm Dynamics in a Labour Market Model", 2009. mimeo.

S. Davis and J. Haltiwanger. *Measuring gross worker and job flows*. Labor Statistics Measurement Issues. NBER, Chicago, 1998. editors Haltinwanger, Manser and Topel.

S. Davis and J. Haltiwanger. *Gross job flows*, volume 3B of *Handbook of Labor Economics*. North Holland, 1999. editors Orley Ashenfelter and David Card.

R. Ericson and A. Pakes. "Markov Perfect Industry Dynamics: A Framework for Empirical Work". *Review of Economic Studies*, 62:53–82, 1995.

D. Hamermesh, W. Hassink, and J. VanOurs. "New Facts About Factor-demand Dynamics: Employent, Jobs and Workers". *Annales d'Economie et de Statistique*, January:21–39, 1996.

T. Klette and S. Kortum. "Innovating firms and aggregate innovattion". *Journal of Political Economy*, 112:986–1018, 2004.

S. Lagarde, E. Maurin, and C. Torelli. "Flux d'emplois et flux de main-d'oeuvre en France: une etude de la periode 1987-1992". *Revue Economique*, 47:633–642, 1996.

G. Moscarini and K. Thomsson. "Occupational and Job Mobility in the US". *Scandinavian Journal of Economics*, 109:807–836, 2007.

M. Muendler, J. Poole, G. Ramey, and T. Wajnberg. "Job Concordances for Brazil: Mapping the Classificao Brasileira de Ocupacões (CBO) to the International Standard Classification for Occupations (ISCO-88)", 2004. mimeo.

V. Orellano and E. Pazello. “Evolução e Determinantes da Rotatividade da Mão-de-obra nas Firmas da Indústria Paulista na Década de 1990”. *Pesquisa e Planejamento Economico*, 36:179–207, 2006.

E. Yashiv. “The determinants of Equilibrium Unemployment”. *American Economic Review*, 90:1297–1322, 2000.