TESTING THE CONNECTION BETWEEN REPLACEMENT AND JOB FLOWS

CARLOS HENRIQUE CORSEUIL - IPEA

Abstract

The present paper investigates the relation between replacement and job creation as proposed by a particular theoretical model of employment dynamics. As opposed to the standard procedure in the literature, we check whether replacement explains job creation. Another contribution of the paper is to propose alternative measures for employment dynamics and it's components. Our estimates points to a negative relation as predicted by the theoretical model. Finally the magnitudes of the estimated coefficients also confirm the predictions and these findings are robust to alternative procedures dealing with measurement error.

Nesse artigo investigamos a relação entre criação de emprego e substituição de trabalhadores tal como proposto em um modelo teórico sobre dinâmica do emprego. Investigamos se o processo de substituição de trabalhadores é um determinante da criação de emprego. Outra contribuição desse artigo é de propor medidas alternativas para ambos os processos mencionados. Nossas estimativas apontam para uma relação negativa entre substituição de trabalhadores e criação de emprego ao nivel do estabelecimento, tal como predito pelo modelo teorico. Além disso as magnitudes estimadas também estão de acordo com as predições. Finalmente nossos resultados se mostram robustos a procedimentos alternativos que lidam com a questão de erros de medida.

KEY WORDS: job creation, replacement, employment dynamics
PALAVRAS CHAVES: criação de emprego, substituição de trabalhadores, dinâmica do emprego
JEL CODES: J63, J23
ÁREA DA ANPEC: 12 - ECONOMIA DO TRABALHO

1 Introduction

Recent empirical papers have made important contributions in understanding employment dynamics by computing separate numbers for both components, one related with gross job flows and another related purely with worker flows. This later component is referred in the literature either as replacement or churning.

Besides the information on how much each of these components contributes to employment dynamics, some of the papers check if they are connected to each other. Usually they use empirical models where job flows is assumed as an explanatory variable for either the overall employment dynamics or for the replacement component. Albaek and Sorensen (1998) and Burgess et al. (2000) are representative papers of this literature. The later considers alternative forms of relations between these variables, including how lags of replacement affects job flows. Nevertheless none of the relations investigated in these papers is grounded in a theoretical framework [1].

The present paper investigates the relation between replacement and job creation as proposed by a particular theoretical model of employment dynamics[2]. As opposed to the standard procedure in the literature, we check whether replacement explains job flows, actually job creation to be more precise.

Another contribution of the paper is to propose alternative measures for employment dynamics and it's components. We take advantage of a matched employed-employee database with detailed information about job position filled by each worker to make use of some information not available in the data sets used in previous papers.

Our findings confirms that replacement is in fact a relevant explanatory variable for job creation. Moreover our estimates points to a negative relation as predicted by the theoretical model.

The remainder of the paper contains a section to discuss the concepts and measures, followed by a section which introduces the data and comments some descriptive results. The econometric analysis is the subject of two other sections, which discuss the empirical model specification and it's results, respectively. A robustness analysis is the object of the final section before our conclusion. Comparisons with related papers will be scattered throughout the sections.

[1] When Burgess et al. (2000) tries to justify the relation mentioned above they write: "In the absence of a formal model, we simply highlight two issues...".(Burgess et al., 2000)[p. 480]

[2] The model is developed in a companion paper, Corseuil (2007).

2 Concepts and Measures

Employment dynamics can be broadly defined as the collection of worker x job matches formed and dissolved within a time period. Some of these flows correspond to shifts in labour demand since they are originated by firm's decision to change their employment structure. These are the case when a worker occupies a newly created job, or when the job where a separation occurred closes down. The literature refers to this component of employment dynamics as gross job flows, or to be more precise job creation in the case of match formation and job destruction in case match dissolution[3]. Some other flows correspond to changes in firm's labour force given a fixed employment structure. This is the case when a worker moves in a job position previously occupied by another worker. This component of employment dynamics will be referred in this paper as replacement[4].

Previous papers have proposed specific measures of employment dynamics that could be decomposed in gross job flows and replacement. The standard procedure is to measure employment dynamics at plant level as the total number of accessions (H) and separations (S) within a time period (usually year or quarter and denoted by t). Concerning the job creation (JC) and the job destruction (JD) component, one of them corresponds to the net employment growth (ΔN), depending on the sign, and the other is null. Albaek and Sorensen (1998) and Burgess et al. (2000) are some of the papers using this methodology. Their measure for replacement (REP) is defined as the difference between either total hirings and job creation or total separations and job destruction.

These standard measures for employment dynamics and the respective components will be denoted with a "s" superscript and can be described as below when defined at plant level (p)[5]:

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

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

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

[3]This component of employment dynamic represents transitions for both workers and jobs. In the case of jobs the transitions happens from inactive to active or vice-versa.

[4]Note that it represents transitions only for the workers, since jobs remain active by construction.

[5]Notation was chosen arbitrary since there is no consensus in the literature.

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

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

Note that the gross job flow measures, and consequently the replacement one, implicitly require the assumption that occupations are homogeneous within the establishments. Another drawback of this procedure is the temporal inconsistence between the gross job flow measures and the others. For instance one establishment might have increased the number of jobs in three units at the beginning of period t, and then at the end of the period it decided to cut two of the three new jobs. The application of the standard methodology would result in two replacements, whereas in fact there was none.

Some papers deal with these problems described above amending the standard procedure in one of the dimensions pointed above. The first modification makes replacement temporally consistent with job flow measures. The other relax the assumption on homogeneity of jobs within an establishment.

Concerning the first amendment, Davis and Haltiwanger (1998) suggest the following concept for employment dynamics: 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 aggregate level due to data constrain[6].

Abowd et al. (1999) build up in this insight using some extra information that they have available to propose analogous measures at the establishment level. They implicitly define employment dynamics as *the number of workers x establishments matches that either exists in t but not in t-1 or the other way around*. The first case is denoted as "entry (rate) excluding within year entry" while the second case has analogous terminology using exit instead of entry[7]. They are able to compute such measures at establishment level. As job flows can also be computed at this level of aggregation, they have all components of employment dynamics computed at establishment level[8]. Their employment dynamics measures can be describe as:

[6]Their establishment level data allow them to compute only job flow rates. Hence they are forced to use household survey for worker flows. The two measures could then be matched only at aggregate level.

[7]The terminology literally translates how they measure these quantities.

[8]This follows from considering replacement as the residual between employment dynamics and job flows.

$$ED'^{+}_{p,t} = H_{p,t} - H^{*}_{p,t}$$

$$ED'^{-}_{p,t} = S_{p,t} - S^{*}_{p,t}$$

where H^* (S^*) represents matches formed (extinct) at t which did not last until the end of period t (which also started at t).

$$REP'_{p,t} = ED'^{+}_{p,t} - JC^{s}_{p,t} = ED'^{-}_{p,t} - JD^{s}_{p,t}$$

The definition for total employment dynamics in Abowd et al. (1999) differs slightly from the one provided by Davis and Haltiwanger (1998). The later proposes to measure the number of workers involved in the afore mentioned transitions, while the former proposes to measure the number of matches involved in conceptually identical transitions, as in our broad definition at the beginning of this section.

The empirical counterpart of the expressions place of work or job, which is part of the definition of employment dynamics, originates another important departure from the standard practice. As opposed to the establishment, which is the default option, one can use occupational categories within each establishment, when this is available.

This alternative procedure was implemented by Hamermesh et al. (1996) and Lagarde et al. (1996) when measuring job creation and job destruction. However they stick to overall accessions and separations as measures of employment dynamics[9]. Using j to denote occupational categories, the components are then defined as[10]:

$$JC_{p,t} = \sum_{j} (\Delta n_{p,j,t}) \cdot I(\Delta n_{p,j,t} > 0)$$

$$JD_{p,t} = \sum_{j} [-\Delta n_{p,j,t}] \cdot I(\Delta n_{p,j,t} \leq 0)$$

$$REP''_{p,t} = H_{p,t} - JC_{p,t} = S_{p,t} - JD_{p,t}$$

We propose an alternative procedure to measure employment dynamics and it's components integrating these two amendments discussed above.

[9] As a matter of fact Hamermesh et al. (1996) implement this procedure in a different way than the one being described here. They rely on direct information about whether a new hiring was motivated by job creation or replacement, and analogous information for separation.

[10] Replacement is not explicitly defined in those papers but both of them implicit refer to this component drawing comparisons between employment dynamics and gross job flows.

That is employment dynamic will be measured as *the number of workers x jobs matches that either existed in t but not in t-1 or the other way around.* This can be represented as:

$$ED^{+}_{p,t} = \sum_{j}[NM_{p,j,t} - NM^{*}_{p,j,t}]$$

$$ED^{-}_{p,t} = \sum_{j}[BM_{p,j,t} - BM^{*}_{p,j,t}]$$

where $NM_{p,j,t}$ ($BM_{p,j,t}$) denotes all matches formation (dissolution) along 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 t. Note that workers coming from other occupational categories within the same establishment are included in our definition of $ED^{+}_{p,t}$ but not in $ED^{'+}_{p,t}$. An analogous comment holds for $ED^{-}_{p,t}$.

Concerning job creation and job destruction, we use the last definitions mentioned above, denoted by $JC_{p,t}$ and $JD_{p,t}$. According to this procedure, our replacement measure will be defined as:

$$REP_{p,t} = ED^{+}_{p,t} - JC_{p,t} = ED^{-}_{p,t} - JD_{p,t}$$

3 Data and Aggregate Results

3.1 Data

Our data comes 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 payers establishments are supposed to send information about all employees which have worked anytime during the reference year [11].

We will use information from 1994 to 2001. Although information is available since 1986, we have some reasons not to use the whole period. First, there was an upward trend in coverage in the late 1980's. Moreover the recent availability of some variables and changes in definition of others in 1994 provide other reasons [12]. Sample size in the original dataset is about 2 to 2.5 million registered establishments per year. We leave some industries out of

[11] The absence of non-tax payers prevent us to claim that the data refers the whole universe of Brazilian establishments.

[12] It is also possible to claim that some structural transformations in Brazilian economy in the early 1990s provide yet another reason.

the sample. Mining, Utilities, Health, Education, Public Sector and Social Services were excluded due to a massive concentration of state operated companies, while Agriculture establishments were left out of the sample due to coverage problems, and Construction were excluded due to it's idiosyncratic employment dynamics and labour relations. We are left with manufacturing, services and trade[13].

This is a matched employer-employee longitudinal data, similar to the ones available for developed countries[14]. The important novelty of this data is to combine the matched employer-employee structure with detailed information available for workers occupation.

This information allow us to list workers id, in each one of the establishments' occupation categories, for consecutive years. We are then able to identify matches formation and dissolution at this level of observations. This is actually the procedure employed in our codes, which is equivalent to, but not the same as, the one described in the expressions for $ED^{+}_{p,t}$ and $ED^{-}_{p,t}$. Therefore it is clear that we take full advantages on the extra information available in RAIS and usually not available in other data sets used to measure employment dynamics[15].

We use 3 digits occupational categories within each establishments as the empirical counterpart of jobs[16]. We share the view expressed in Moscarini and Thomsson (2006) that "... measuring occupations at this level corresponds most closely to the notion of labor technology, with labor input being differentiated by the tasks involved...".

The use of such disaggregate categorization may rise some concerns about measurement errors. For instance, it may happens that a worker continuously employed in the same establishment, performing the same task, is classified in different but close related job categories in consecutive years. We will provide some robustness analysis for our main results where we consider two alternative procedures. The first one excludes such movements within establishments. The other is to consider the 2 digits occupational categories as the job classification. Moreover it worths saying that MTE does care very

[13]Some further screening procedures applied in the original data set are described in appendix A.

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

[15]Some data have the matched employer-employee structure but do not have information for occupation, while others do have information for occupation but lack either the worker or the establishment id.

[16]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 about the categorization used in this paper (CBO-94) and the ISCO-88.

much about the quality of this information because some labour and pension regulations in Brazil depends on the employee occupation[17].

3.2 Aggregate Results

Table 1 below brings the numbers for each of the measures defined above as a share of the average employment level. We can see that new matches corresponds to 45% of average employment level, among which 17% corresponds to replacement and 28% represents newly created jobs. Results are similar for broken matches which amounts to 43%, with a job destruction rate of 26%. From this table we can calculate that 38% (40%) of new (broken) matches corresponds to replacement. These numbers are considerably lower than the analogous ones from the literature mentioned above.

Table 1: Employment Dynamics 1994-2001

	positive	*negative*
employment dynamics	**45.3**	**43.1**
replacement	17.0	17.0
job flow	28.2	26.1

Note that we have important methodological divergences that may compromise comparisons with previous papers. In fact both amendments that we implement tends to diminish the replacement, relative to the standard procedure. First the procedure to make the measures temporally consistent tends to decrease the overall employment dynamics numbers, hence the numbers for replacements as well. Second, the use of occupational categories for jobs as opposed to establishments tend to increase the numbers for gross job flows, since it uncovers some smoothness introduced by the aggregation at the establishment level (i.e. some categories may create jobs while others may destruct in the same establishment). As replacement corresponds to the residual part of employment dynamics after taking out job flows, then lower levels are expected when this last procedure is implemented.

Indeed the share of replacement in overall employment dynamics is considerably lower in our results than in papers using other measure. For instance, with respect to the standard methodology, Albaek and Sorensen (1998) find that the share of match formation related to replacement amounts to almost

[17]In 1994 when introducing the CBO-94 in RAIS, MTE made a informational campaign among employers about this classification system.

60%, and the same holds for dissolutions. Burgess et al. (2000) uses the same methodology for quarterly data and find 70% for replacement's[18] share in the sum of hirings and separations[19].

Perhaps the most illustrative comparison for this porpouse is provided by Orellano and Pazello (2006) which applies the standard methodology to the same data source than we use. Their numbers allow us to calculate that replacement corresponds approximately to 62% of employment dynamics[20].

These different magnitudes show that improving the measures of employment dynamics and their components is not a minor point. The test to be performed later could be severed affected by measurement errors of such magnitude.

4 Empirical Model

4.1 Background

In a companion paper it is developed a labor market model where firms decide simultaneously about job creation and workers replacement[21]. The main feature of the model is imperfect information about workers productivity, which is revealed only after production takes place.

This framework predicts that, conditioned on firm size, the higher the number of replacements the smaller the job creation.

The mechanism is the following, due to constant returns to scale, firms will always try to hire workers. The decision on how many workers the firm will look for, and how she will distribute them among replacement and new job positions will depend on the revealed quality of it's matches with their employees. Those firms with a bad composition will tend to have a relatively high number of new matches which will mostly be used to replace workers. Those with a good composition will tend to have a relatively low number of new matches which will mostly be used to fill newly created jobs.

Note that the predicted relation between replacement and job creation is not a causal one. The causality in the theoretical framework goes from matches quality to both replacement and job creation.

[18]Replacement is referred as churning in their paper

[19]The effect of each of these amendments on the relevance of replacement can be accessed in different papers. Abowd et al. (1999) provide the numbers to evaluate the effect of the first amendment while Hamermesh et al. (1996) allows an evaluation of the second amendment.

[20]Their sample differs from our in terms of spatial and temporal coverage.

[21]In the theoretical model the distinction between firm and establishment (or plant) is not relevant.

Among empirical papers dealing with employment dynamics, there is only one (as far as we know) which investigates something connected to the prediction mentioned above. That is the paper of Burgess et al. (2000) which brings ambiguous evidence using a different approach than ours. Some of their exercises suggests a positive relation between replacement and gross job flows, while others suggests a negative relation[22]. It should be stressed that they use the standard measures for job flows and replacement, which differs from our measures as shown before.

One important point raised by this 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.

Some other empirical models investigate the determinants of job flows[23]. These exercises are not connected with our model, since they leave out replacement level. However they point to other potentially relevant determinants for job flow.

The standard procedure is to include plant characteristics among explanatory variables, such as age, size and industry category. The last two variables are justified 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 relevance of age and size as the most relevant plant characteristics, whereas industry category and time effects also contributes to explain job flows.

4.2 Specification

We then propose the following model for establishment level observations.

$$JC_{p,t} = f(REP_{p,t}, x_{p,t}, AGE_{p,t}, \alpha_p, i * t, \epsilon_{p,t})$$

[22]They report three types of results. First a scatter-plot picture of contemporaneous employment dynamics (which they call worker flows) and job flow numbers at establishment level. This exercise suggests a negative relation between replacement and job flows. Later they regress job flow rates on establishment fixed effects and contemporaneous and lags churning (replacement) flow rates. The coefficient of the contemporaneous replacement appears as positive, but the one for the first lag shows negative sign. Moreover they show that churning rates has positive autocorrelation (also controlled for fixed effects).

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

Where α_p represents plant idiosyncratic effect, $x_{p,t}$ represents the establishment average size concerning t and $t-1$, $AGE_{p,t}$ is a binary variable for establishments older than 3 years and $i * t$ denotes interactions between dummies for year and industry (i) categories (two digits level)[24].

The empirical models specified in other papers, mentioned above, diverge from ours in some specification issues. First, their dependent variable are usually defined as a rate relative to the size. However the theoretical predictions to be tested with this model refers to the level of replacement and level of job creation.

The size measure itself is another delicate point and deserves a few extra words. Davis and Haltiwanger (1999) claim that the size effect on job flow is very sensible to whether it corresponds to the initial size or the average size between two consecutive years. They argue in favor of the later specification to avoid bias due to the "regression to the mean" effect. We follow their recommendation.

Since the quality of the match is not observed (for the econometrician), replacement becomes an endogenous variable in this regression framework (assuming the theoretical model is correct). Hence our concern with the specification of the empirical model is not to avoid endogeinity, but rather to make match quality the exclusive source for it. Hence we assume that, apart from the quality of the matches, there is no other non-observable time varying establishment effect that is simultaneously correlated with both replacement and job creation decisions. In appendix B we formalize this point and show that, under these conditions, a negative sign for the replacement coefficient should be expected in our regression.

For the sake of simplifying computational procedures, $f(.)$ will be parameterized as linear. Apart from excluding some establishments according to their industry category, as already mentioned before, we also exclude any births and deaths episodes. The reason is obvious, since these establishments could not choose to replace any worker.

[24] 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 employment dynamics is from 1994 to 1995 there is 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 Empirical Evidence

5.1 Job Flows and Replacement Frequencies

The model described above is asymmetric with respect to job creation and job destruction. It neglects the job destruction decision, making it exogenous, and discards the interaction between replacement and job destruction.

We may have an idea about how restrictive is this procedure by looking at the frequencies of each employment dynamics' components across establishments. Table 2 reports this information. The numbers refer to the cross establishments distribution of all combination of possible actions. For example the number in the first row and first column identifies the share of establishments that only create jobs (do not destroy any job nor replace any worker). Analogous explanations define the other numbers in the main diagonal. The number in the first row and second line shows the share of establishments that create and destroy jobs but do not replace any worker. Analogous explanations define the other numbers above the diagonal. Finally, the possibility of the three actions are taken simultaneously, is denoted as full action and the corresponding result is located close to the bottom/right corner.

Table 2: Distribution of possible actions across establishments (%)

	job creation	job destruction	replacement	
job creation	9.6	14.4	9.2	
job destruction		10.1	7.4	
replacement			7.2	
				all 20.0

Notes:
(1) 22.2% of the establishment opt for none of these actions.

The most relevant information that comes out of this table are twofold. The first one is the inexpressive share of establishments that combines replacement with job destruction. This is less frequent than combining replacement with job creation. In fact this is the less expressive combination among these involving any two actions.

The other relevant information is the expressive share of establishments interacting two or more components of employment dynamics. 51% of the 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 employment dynamic models addressing only one action might be missing important features.

5.2 Estimations from the Empirical Model

As discussed above we would expect a negative and significant coefficient for $REP_{p,t}$. Moreover we had mention that those establishments with high number of replacement would also tend to have a high number of new matches. Therefore the expected result is actually a significant number between -1 and 0.

Table 3 presents the estimated coefficients by broad industry categories. It is interesting to note that results are in line with the model predictions in all three groups considered. The replacement coefficient is always negative with absolute value lower than one. An extra replacement tends to be compensated with 0.28 less jobs created in the service industry, 0.41 in manufacturing and 0.60 in trade. Therefore the higher the replacement level, the higher tends to be the total number of new matches[25].

Table 3: Results from the empirical model for job creation

	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
age > 3	-0.34	-5.9	-0.72	-6.4	-0.24	-20.1
n.obs	2267550		1095565		3131188	
R-Square	0.52		0.47		0.57	

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

Among the other coefficients it is interesting to analyze the ones related to size and age. We can see that older establishments tend to create less jobs, which is a standard result both in empirical labour economics literature and in empirical industrial organization literature.

[25] According to the theoretical model this is related to an assumption that replacement is cheaper than creating a new job.

The comparison involving the size effect is not so straightforward since the other papers apply alternative specifications, as mentioned before. 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.

6 Robustness Analysis

6.1 Restricting within plant worker flows

As already mentioned the use of 3 digit occupational categories may imply mismeasurement of the employment dynamic components. This problem tend to be more relevant in worker flows within establishments. A natural possibility to check the robustness of our results is to use the 2 digit codification for job categories. The aim of this procedure is to analyse how the results would vary if we consider that all within establishment movements across "similar" job categories are motivated by measurement error.

The regressions results analogous to table 3 based on these measures are shown in table 4. One can see that we get estimated coefficients extremely close to those shown in table 3.

Table 4: Results from the empirical model for job creation - 2 digits 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	2267550		1095565		3131188	
R-Square	0.50		0.45		0.55	

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

One may think that measurement errors are not restricted to "similar" job categories. So we can restrict even further the within plant worker flows to check how robust are our results. We then use in this section some alternative measures that abstract from any within plant movements. These measures can be defined as below:

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

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

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

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

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

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

where:

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

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, that is in general we have $\Delta n^{\diamond}_{p,j,t} \neq \Delta n_{p,j,t}$. It follows then that replacement is measured as:

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

The regressions results analogous to table 3 based on these measures are shown in table 5. One can see that such procedure did not qualitatively change our main results. The replacement coefficients are now only slightly lower, in terms of absolute value, which remains lower than one. The size and age coefficients also do not present any substantial change.

6.2 Alternative environment for the decision process

Given that the sample is dominated by small establishments, one might question whether the mechanism described in the theoretical framework also holds for big establishments. The question may be relevant if decisions on job creation and replacement are decentralized in big establishments. We then repeat our basic exercise restricting the sample for big establishments.

Table 6 presents the results for a sample restricted to establishments with more than 50 employees, in each one of the three broad industry categories.

Table 5: Results from the empirical model for job creation - 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	2178742		1095565		3131188	
R-Square	0.60		0.42		0.53	

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

The results reveal that replacement seems to affect job creation in big establishments in the same way as we were describing before. The coefficients are all negative, significant and with absolute value lower than one, as it was the case for the complete sample. This may indicate that even if decisions are decentralized in big establishments they are subject to a common budget constrain. Hence it seems that the replacement level in one unit of the establishment does affect the job creation level in another one.

Table 6: 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	88808		94277		46312	
R-Square	0.49		0.44		0.56	

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

Another possibility concerning the decision process regards to plants which are part of firms with other plants. It might then be the case that

decisions are centralized at firm level, which means that the estimated relation may not hold at the establishment level for a sample restricted to this group of establishments. Table 7 shows that once more our qualitative conclusions remain intact.

Table 7: Results from the empirical model for job creation - Multi-plant firms

	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
age > 3	-0.29	-40.8	-0.47	-31.4	-0.19	-42.7
n.obs	2178742		1001288		3084876	
R-Square	0.62		0.60		0.60	

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

7 Summary and Conclusion

The main goal of this paper is to test a theoretical prediction on how replacement and job creation are related to each other. The first contribution of the paper is to propose an alternative procedure to measure such components of employment dynamics. A rich dataset combining a matched employer-employee structure with detailed information on the occupational category of each worker allow us to combine, for the first time, the best available procedures to measure employment dynamics and gross job flows.

We then estimate an empirical model, based on some recommendations from the empirical literature on determinants of gross job flows and also on the theoretical framework which originated the prediction to be tested. The results confirm the prediction that establishments with higher replacement levels tend to have lower levels of job creation, conditioned on size. Moreover the magnitude of this effect is such that one extra replacement is associated with less than one extra job created, which is also in line with the theoretical framework being tested.

The results are also in line with those presented in other papers with respect to the effects of size and age on job creation. This is worth stressing

since we have used alternative measurement procedures for the job creation variable and also an alternative specification for the empirical model combining establishment fixed effects with these two controls and replacement. Finally it should be stressed that all these results are maintained in a robustness analysis.

The prediction tested in this paper is one of several others delivered by the same theoretical model. All the others are confirmed by stylised facts readily available from the literature. This fact gives support to the claim that the framework grounding our model is a relevant one to explain employment dynamics. The main message from the model is that firms are not passive. They try to revert their fortune when a poor performance in efficiency levels is revealed. One instrument used to achieve this goal is to try to improve the quality of their labour force, by replacing those workers in low productivity matches.

Appendices

A Data cleansing

The following screening procedure were employed to original data:

- deletion of individuals with invalid id code (missing or zero).
- deletion of establishments with invalid id code (missing or zero).
- deletion of establishments with more than one spell from 1992 to 2002. This is to avoid overestimation in job creation and job destruction figures due to establishment that although in operation did not have their information processed in a particular time period.
- Merge the following duplicated job codes pairwise: 073 and 193-Social worker; 074 and 194-Psychologist; 093 and 110-Accountants; 162 and 454-Decorator.

B Assumptions for the Empirical Model

Let $q_{p,t}$ represents an index of the quality of the matches in establishment p at time t. The theoretical model tells us that $REP_{p,t} = REP_{p,t}(q_{p,t})$ and $JC_{p,t} = JC_{p,t}(q_{p,t})$. Therefore it follows that:

$$E[\epsilon_{p,t} \mid REP_{p,t}(q_{p,t}); \Omega_{p,t}] \neq 0$$

where Ω denotes the set of all our explanatory variables in addition to replacement. This means that our estimation of replacement coefficient does not have a causal interpretation. However it does not mean that we can't extract any relevant information from this number. In what follows we show that under certain assumptions we can use this estimated coefficient to test the theoretical model relevant predictions related to how job creation and replacement interacts with each other.

The main assumptions are:

$$\epsilon_{p,t} = q_{p,t} + \mu_{p,t}$$

$$E[\mu_{p,t} \mid REP_{p,t}(q_{p,t}); \Omega_{p,t}] = 0$$

Then the following result can be easily derived:

$$\beta^{nq}_{rep} = \beta^{q}_{rep} + \frac{cov(q, REP)}{var(REP)}$$

Where β^{nq}_{rep} corresponds to the replacement coefficient in our regression (not controlled by q), while β^{q}_{rep} is the analogous coefficient in a regression controlled by q (and all the others explanatory variables that we use). According to the theoretical model this last coefficient is null. It means that our estimations of β^{nq}_{rep} captures the sign of $cov(q, REP)$, and hence should be negative if the model is correct.

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. "Imperfect Information in the Labour Market: Implications Towards Firm Dynamics", 2007. 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.

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.

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", 2006. mimeo.

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.