Estimating Vacancies from Firms' Hiring behavior: The Case of a Developing Economy*

Leonardo Fabio Morales†
lmoralzu@banrep.gov.co

José Lobo‡
jjloboc@unal.edu.co

Abstract

We propose a procedure that allows recovering an estimate of vacancies from firms' information on hires and separations. Using estimated vacancies, we analyze the aggregated behavior of vacancies for the Colombian labor market. In addition, we estimate matching functions to conclude that the matching formation process for the Colombian labor market is random; this finding support the idea that frictions are mainly due to informational restrictions, and not explained by a structural mismatch. Our method might be useful in developing economies, where there are no good official sources of information on the matter.

JEL Classification Codes: J60, J63, J23

Keywords: Turnover, Vacancies, Layoff, Labor Demand, Labor Force and Employment.

*We want to thank participants in the research workshop at the Colombian central bank, and participants of the 21st conference of the International Network for Economic Research for their valuable feedback. We also thank Carolina Velez and Estefany Hernandez for their research assistance. We are fully responsible for the opinions expressed here, as well as for possible mistakes. The views and conclusions expressed in this paper are those of the authors, and they do not imply any commitment on the part of the Banco de la República or its Board of Governors. The algorithm proposed and implemented in this paper was executed using STATA version 16 MP on an Intel(R) Xenon(R) CPU workstation running at 2.10GHz with 128 GB of memory. Stata algorithms to implement the estimation method were written by the authors, and they are available at the corresponding author's website: https://sites.google.com/site/leonardomoraleszurita/home/researchlm/formal-labor-market-vacancies-in-colombia.

† Researcher at the Colombian Central Bank.
‡ Research Assistant at the Colombian Central Bank.

1. Introduction

Vacancies are an element in the process that determines worker and job flows. In the theoretical literature, vacancies play a crucial role in Equilibrium Unemployment Theory models; together with unemployment, vacancies are the attributes of matching functions, the most salient theoretical innovation of search models (Petrongolo & Pissarides, 2001). Despite the importance of vacancies in theoretical models and its critical influence in the determination of a firm's labor stock, relatively few empirical studies analyze vacancies, and its relation with hires and separations at the establishment level (Davis, Faberman, & Haltiwanger, 2013). Most of the studies on vacancies use data from a few developed countries, where information on firms' open positions is available.

In the case of developing countries, the literature on vacancies is even scarcer, however, some studies attempt the construction of vacancy indexes; they use for this purpose the number of Help Wanted Advertisement posts in newspapers (Álvarez & Hofstetter, 2014; Arango, 2013). For some countries, this is the only way to measure vacancies, given the absence of official sources of information on the matter (Álvarez & Hofstetter, 2014). Some studies suggest that job vacancies based on Help Wanted ads may capture the movements and dynamics of vacancies (Abraham & Wachter, 1987; Amoah, 2000); nevertheless, it would be impossible for that methodology to capture the level of this variable. An additional issue with the Help Wanted ads is that in recent times, with the popularization of the internet, posting ads in written newspapers is a less frequent practice. There are corrections suggested in the literature to deal with this problem.

The empirical literature on job markets flows is wide, especially in developed countries, where there are sources of information that allow the measurement of worker and job flows with precision (Burgess, Lane, & Stevens, 2000; Davis, Haltiwanger, & Schuh, 1996; Morales, Hermida, & Davalos, 2019). Nevertheless, in most of the studies, vacancies are implicit in the labor market flows analyzed, and they do not play a direct role in the intensity of job and worker movements. Even though vacancies are the center of any job creation process, they are relegated to an almost unnoticeable place in the analysis of labor worker and job dynamics. This is the case, possibly, because of the difficulty of linking data sources that allow for the analysis of job and worker flows and vacancies at the same time.

In the traditional microeconomic view of firms' behavior, they must solve the problem of reaching their optimal employment level; they have mainly two tools to do this: layoffs and vacancies. In regards to vacancies, firms will open vacancies for two reasons. On the one hand, if they want to create new positions; therefore, vacancies play an active role in determining job creation. We will refer to this type of vacancies as expansion vacancies. On the other hand, firms might open vacancies if there is a fraction of separations that they need to replace; in that case, this share of separations translates into vacancies. By this way, there is a direct relationship between vacancies and labor churning. We will refer to this type of vacancies as replacement vacancies.

In this paper we adopt a view of the concept of vacancies, in which they play a more direct and meaningful role in the process that governs workers and job movements. We do this by modeling a hiring function that depends on separations and expansion vacancies. We justify the importance of a hiring function in a theoretical framework. For our empirical work, we use an employer-employee linked panel from Colombia, which is based on administrative records from the social security system, usually refered as PILA. Based on the estimation of the hiring function, we propose a methodology that allows the researcher to recover an estimate of vacancies. For this purpose, standard information on hires, separations, and employment at the establishment level is required. With the modernization of the social security systems, this type of data is becoming available for many developing countries; in addition to the PILA for Colombia, there are similar data sets for Brazil (RAIS), and for Mexico (IMSS), just to mention few examples.

We estimate expansion and replacement vacancies (flows and stocks); we show the consistency of our methodology using Monte Carlo experiments. Our estimation of aggregated vacancies behave realistically in the context of the formal Colombian labor market during the study period. Furthermore, we document some important relationships of vacancies with other aggregated variables as unemployment and hires. The relationships of our estimated vacancies with other labor market variables preserves expected properties suggested by theory. As previous studies remark, vacancies collected directly from firms can be the subject of measurement and aggregation problems, which complicates the inference of economic relationships between vacancies and other variables using raw data from these sources (Davis et al., 2013). The ratio of hires to vacancies that we obtain from our estimation, seems to behave more according to theory predictions, than previous computation of this ratio using raw data on observed vacancies for other countries.

Traditional measures of job flows are proxies for real processes of job creation and destruction. In some situation this proxies might be imprecise; for instance, a firm might have destroyed job positions in months of positive net growth. Similarly, the magnitude of job creation is not necessarily described precisely by firms' employment growth; for instance, for a given period, hires may be the result of worker replacement from previous separations. Hires may also be the result of the creation of new vacancies in the past. These limitations of job flows have been remarked previously in the literature using the following paradox first proposed by (Davis et al., 1996).

> "Some newly created and newly destroyed jobs may not show up as plant-level employment changes. For example, a plant may destroy ten assembler jobs and create ten robotics technician jobs, so that total employment does not change."

As a by-product of our estimated vacancies, we propose alternative measures of creation and destruction of job positions. Using the flow of expansion vacancies, we generate a rate of creation of new job positions. Using a similar logic, we propose a measure of destruction of job positions.

Finally, our methodology is very appealing for the estimation of matching functions; given that we are able to recover vacancies flows and stocks, we can test if the matching formation process for the Colombian labor market is random or stock-flow type. We find support for the random matching formation structure over the stock-flow structure. This results means that the most relevant search frictions in the Colombian labor market are due to informational lacks, mainly on job locations. Results does not support the existence of a structural mismatch between workers and available jobs; seekers unmatched in previous rounds will match, if sufficient time have passed to overcome the informational restrictions.

In the second section of this paper, we describe a theoretical framework that justifies the use of a hiring function. In the third section, we propose a simple model of firms' hiring behavior, and we describe a methodology that allows predicting vacancies from it. In the fourth section, using Monte Carlo experiments, we evaluate the performance of our methodology. In the fifth section, we apply the methodology using Colombian data and get an estimation of the formal labor market vacancy-stock. In the sixth section, we proposed vacancies' based measuments of job positions' creation and destruction. In section sixth, we analyze the aggregated behavior of vacancies, hiring, and unemployment. In section sixth, we show the results of matching function estimations, and finally offer conclusions and some implications for policy in ninth section.

2. Theoretical Framework

From the firm's point of view, flows of separations, hires and expansion vacancies are derived from the solution to the maximization-of-profits optimization problem. Standard approaches to the firm's problem consider labor as a homogeneous input and the stock of workers as the control variable the firm chooses for the maximization of profits. As previous literature has pointed out, in a model with homogeneous labor the following will not occur at the same time for a single firm: vacancies, hires, and temporary layoffs (Holt & David, 1966); this is because no hires or vacancies are necessary if it is optimal for the firm to reduce its size. Data from real firms shows that shrinking firms have a good deal of hiring (Davis & Haltiwanger, 2014); this evidence is observed in developing economies as well (Flórez, Morales, Medina, & Lobo, 2020; Morales & Medina, 2016). Therefore, in this framework, we consider several types of labor, each one with different wage rates. In addition, because of congestion in the market, firms may not be able to reach their desired size immediately; therefore, we take into account frictions in the market by introducing a hiring function, which determines the firms' hires as a function of the stock of vacancies.

In this framework, a firm decides the number of separations and expansion vacancies subject to a technological restriction and subject to a firm hiring function $h_t^j(\cdot)$, which maps the stock of vacancies to hires at a particular period. By deciding separations and vacancies, firms control their desired employment size each period. At the beginning of a given period *t* firms know the employment stock at the end of the previous period e_{it-1}^j, in addition, firms realize that some of the employees that belong to the last period employment stock leave voluntarily, this amount is represented by $\hat{s}_{it}^j$. Each period, a firm has to make decisions on separations

($\check{s}_{it}^{j}$) and expansion (new) vacancies ϑ_{it}^{j}, for each type of labor input *j*. The profit function[4] of a firm in a competitive market, in a given period *t*, can be represented as:

$$\pi_{i,t} = p_t \cdot f\left(e_{it}^{j}\right) - \sum_{j=1}^{J} w_t^{j} \cdot e_{it}^{j} \qquad (1.1)$$

$$\text{where, } e_{it}^{j} = e_{it-1}^{j} - \hat{s}_{it}^{j} - \check{s}_{it}^{j} + h_t^{j}\left(\hat{s}_{it}^{j}, \vartheta_{it}^{j}, V_{it-1}^{j}\right) \quad (2.1)$$

In the previous equations, e_{it}^{j} stands for the amount of labor available for production at the end of period *t*; equation (2.1) describes the dynamics of the stock of each labor type in a firm. Firms must optimally decide $\check{s}_{it}^{j}$ and ϑ_{it}^{j} taking into account that they start the period with $\hat{s}_{it}^{j}$ fewer employees. We assume that since $\hat{s}_{it}^{j}$ were quits, firms want to fill those positions again. Firms will increase $\check{s}_{it}^{j}$ if they find reducing their size to be optimal; firms may increase expansion vacancies ϑ_{it}^{j} if they find increasing the size of the firm to be optimal. The expression $h_t^{j}\left(\hat{s}_{it}^{j}, \vartheta_{it}^{j}, V_{it-1}^{j}\right)$ represents the firm's hiring function, which determines the firm's hiring. Since there are frictions in the labor market, firms may not be able to fill all vacancies in a given period, therefore, in general $h_t^{j}(\cdot) \leq \hat{s}_{it}^{j} + \vartheta_{it}^{j} + V_{it-1}^{j}$; for this reason, the hiring function depends on the stock of vacancies that, at the end of the previous period, had not been filled yet V_{it-1}^{j}.

From the solution of its dynamic optimization problem, a firm obtains policy functions $\check{s}_{it}^{j\,*}, \vartheta_{it}^{j\,*}$, which are sequences of optimal choices for the control variables. The optimal hiring behavior is given by $h_t^{j}\left(\hat{s}_{it}^{j}, \vartheta_{it}^{j\,*}, V_{it-1}^{j*}\right)$, where V_{it}^{j*} is the stock of vacancies derived from the optimal controls for separations and expansion variables ($\check{s}_{it}^{j\,*}, \vartheta_{it}^{j\,*}$). In this model, the traditional measures of worker and job flow in the empirical literature: hires, separations, job creation, and job destruction, will be represented by: $\sum_j h_t^{j}$, $\sum_j\left(\check{s}_{it}^{j} + \hat{s}_{it}^{j}\right)$, $\sum_j 1_{\left\{\Delta e_{it}^{j}>0\right\}}\Delta e_{it}^{j}$, $\sum_j 1_{\left\{\Delta e_{it}^{j}<0\right\}}\Delta e_{it}^{j}$, respectively.

[4] For easiness in the notation, we ignore capital inputs; the production technology of the firm is a function of all types of labor.

There are three important takeaways from this section. First, vacancies play a fundamental role in the dynamics of worker and job movements. Second, hiring behavior is the link between plant size and vacancies, and it is affected by rigidities of the labor market; therefore, in general, firms cannot reach their desired size instantaneously. Analyzing firm's hiring behavior, represented by the function $h_t^j\left(\hat{s}_{it}^j, \vartheta_{it}^{j^*}, V_{it-1}^{j^*}\right)$, is crucial for the understanding of fluidity of labor inputs at the level of the firm. Third, in the analysis of firm hiring behavior, there should be a clear distinction between replacement vacancies and expansion vacancies; even though both determine hires, they can be governed by different dynamics. In the next section, we propose a simple empirical strategy to estimate the h_t^j function; this will allow us to get an estimation of the flows and stocks of vacancies for each firm in the economy.

3. A Simple Model of Firm's Hiring Behavior.

3.1 A Model of the firm's hiring behavior.

The total amount of vacancies of a firm is the result of two elements. On the one hand, the share of all firm's separations that is replaced eventually; in other words, there are some vacancies that firms open with the purpose of replacing workers that are gone. On the other hand, there are vacancies that firms open for expansion purposes. The first type of vacancies does not imply job creation, they are an important component of the churning rate, and we will refer to them as replacement vacancies. The second type of vacancies is the primary source of job creation, and we will refer to them as expansion vacancies. In this model, no vacancy is discarded without being filled, and all hiring has the purpose of filling a previously open vacancy.

Let us denote by s_t the total number of separations that take place in a firm during period t. On average, firms substitute a fraction of all separations, we assume this portion is constant, and we denote it as π; therefore, πs_t represents the total number of separations generated in period t that will eventually be replaced. Therefore, firms will hire πs_t workers from replacement vacancies generated at period t, but due to congestion in the labor market, it will take several periods for the firm to fill in these positions. The dynamics of this hiring behavior is modeled using a lags polynomial, the period t hires that correspond to the filling of replacement vacancies are represented as:

$$h_t^r = \sum_{\tau=0}^{L} \tilde{\theta}_\tau(\pi s_{t-\tau}) = \sum_{\tau=0}^{L} \theta_\tau s_{t-\tau} \qquad (3.1)$$

where, $\sum_\tau^l \tilde{\theta}_\tau = 1, \ \ \sum_\tau^l \theta_\tau = \pi$

The term θ_τ, which is equivalent to $\pi\tilde{\theta}_\tau$, accounts for the proportion of all replacement vacancies generated in a previous (or current) period τ that are filled in the current period, and therefore, they become hires. In regards to the expansion vacancies, they represent the creation of jobs at the level of the firm; due to congestion in the labor market, it will take several periods for the firm to fill these new positions. The period t hires that correspond to the filling of expansion vacancies are represented as follows:

$$h_t^e = \sum_{\tau=0}^{R} \phi_\tau \vartheta_{t-\tau} \qquad (3.2)$$

where $\sum_{\tau=0}^{R} \phi_\tau = 1$. The term ϕ_τ accounts for the proportion of all expansion vacancies generated in previous (or current) period $t-\tau$, which are filled in the current period, and therefore, they become hires. For instance, the term $\phi_1 \vartheta_{t-1}$ represents the number of expansion vacancies generated at the previous period that are filled in the current period t. In both types of vacancies, replacement or expansion, the length of the polynomial (L and R, respectively) determines the total amount of periods that it takes for vacancies generated in a given period to be filled. The vacancies that are filled at period t are represented by the following hiring function:

$$h_t = \sum_{\tau=0}^{L} \theta_\tau s_{t-\tau} + \sum_{\tau=0}^{R} \phi_\tau \vartheta_{t-\tau} + \epsilon_t \qquad (3.3)$$

where, $\epsilon_t \sim N(0, \sigma_e)$ is an iid measurement error. The hiring function in equation 3.3 represents the accounting identity that explains hiring as a function of exclusively two components: replacement vacancies and expansion vacancies. This simple function determines the hiring of the firm because hires are originated by the existence of vacancies. We will refer as V_t^e and V_t^r as the stock of expansion and replacement vacancies; they are computed using equations 3.4 to 3.6. These formulations express stocks V_t^e and V_t^r, as the sum of all vacancies that were generated in previous periods, but they have not been filled completely (In appendix A, we offer more details).

$$V_t^e = (1 - \phi_0 - \phi_1 - \cdots - \phi_{\mathrm{R}-1})\, \vartheta_{t-R-1} + (1 - \phi_1 - \cdots - \phi_{\mathrm{r}-2})\, \vartheta_{t-R-2} + \cdots + (1 - \phi_0)\, \vartheta_t \qquad (3.4)$$

$$V_t^r = (\pi - \theta_0 - \theta_1 - \cdots - \theta_{L-1}) s_{t-L-1} + (\pi - \theta_0 - \theta_1 - \cdots - \theta_{L-2}) s_{t-L-2} + \cdots + (\pi - \theta_0) s_t \qquad (3.5)$$

$$V_t = V_t^e + V_t^r \quad (3.6)$$

Employment of the firm, in a given period, is given by the level of employment in the previous period, plus contemporaneous hires, minus contemporaneous separations; this is $e_t = e_{t-1} + h_t - s_t$. Substituting equation 3.3 into this former expression, we obtain the following expression for gross employment changes:

$$\Delta e_t = (\theta_0 - 1) s_t + \sum_{\tau=1}^{L} \theta_\tau s_{t-\tau} + \sum_{\tau=0}^{R} \phi_\tau \vartheta_{t-\tau} + \epsilon_t \quad (3.7)$$

Equations 3.3 and 3.7 represent the hiring behavior of the representative firm and the movement equation of the firm's level of employment; the variables $e_{t,} s_t, h_t$ in those equations are observable factors in this paper, nevertheless ϑ_t, the flow of expansion vacancies is not observed.

The empirical literature on job and worker flows use plant-level data, which are usually samples of firms from specific economic sectors (Davis et al., 1996). In some cases, studies use information for all formal firms in a specific US state (Burgess et al., 2000). In other cases, studies used linked employer-employee panel datasets for the universe of formal firms for a specific country (Morales & Medina, 2019). In all those cases, employment, separations and hires ($e_{t,} s_t, h_t$, respectively) are observable variables, in most cases vacancies are not. Datasets containing information on vacancies are very uncommon; one example for the US is the Job Openings and Labor Turnover Survey (JOLTS).

3.2 Predicting vacancies from firms' hiring behavior

In this subsection, we propose a methodology that allows predicting the firm's flows and stocks of replacement and expansion vacancies, using for this purpose firms' information on separations, hires, and employment. Information at the firm level on hires and separations is much more common that information on vacancies. There are many countries with statistical or administrative information on these variables; for instance, in addition to the PILA for Colombia, there are similar data sets for Brazil (RAIS), and for Mexico (IMSS).

Let us assume that firm's job creation follows a Poisson counting process; therefore, the flow of expansion vacancies for all firms, at a given period t, is identically and independently distributed Poisson with media and variance ϑ_{jt} (i.e., $\vartheta_{\mathrm{jt}} \sim \mathrm{poisson}(\boldsymbol{\vartheta}_{\mathbf{jt}})$). The flow of hires and separations are observed in our data for each firm; nevertheless, vacancies are not; with the available data, an estimable version of equation 3 is:

$$h_{jt} = \sum_{\tau=0}^{L} \theta_{\tau} s_{jt-\tau} + \alpha_{\mathrm{jt}} + u_{\mathrm{jt}} \quad (3.8)$$

where $\widehat{\alpha}_{\mathrm{jt}}$ is a firm and time fixed effect (i.e. $\widehat{\alpha}_{\mathrm{jt}} = \widehat{\alpha}_{\mathrm{j}} + \widehat{\alpha}_{\mathrm{t}}$). The set of different intercepts α_{jt} are the expected value of hires when $s_{\mathrm{jt}-\tau} = 0$ for all τ, therefore:

$$E\left[h_{jt} \middle| s_{jt-\tau} = 0, \forall\, \tau\right] = \sum_{\tau=0}^{R} \phi_{\tau} \vartheta_{jt-\tau} \quad (3.9)$$

An estimation of this previous expectation in equation 3.8 is the individual and time varying intercepts:

$$E\left[h_{jt} \middle| s_{jt-\tau} = 0, \forall\, \tau\right] = \alpha_{jt} \quad (3.10)$$

Assuming that in equation 3.3 $E\left[\epsilon_{jt} \middle| s_{jt-\tau}, \vartheta_{jt-\tau}\right] = 0$, it can be seen from equations 3.9 and 3.10 that intercepts α_{jt} in equation 3.8 would capture the effect of the polynomial $\sum_{\tau=0}^{R} \phi_{\tau} \vartheta_{jt-\tau}$, which is unobserved in this equation; therefore, $\sum_{\tau=0}^{R} \phi_{\tau} \vartheta_{jt-\tau} = \alpha_{jt}$. We assume that the flow of the expansion vacancies is such that $\vartheta_{jt} \sim poisson(\boldsymbol{\vartheta}_{jt})$, therefore, the mean of these vacancies may change by firm and time. A reasonable assumption is that the mean of this process is stable in time, in the sense that for a small set of periods it may remain fixed. If expectation of $\boldsymbol{\vartheta}_{jt}$ does not change during the R+1 periods of the expansion vacancies polynomial, then we can express the mean of the expansion vacancies as follows:

$$\alpha_{jt} = E\left[\sum_{\tau=0}^{R} \phi_\tau \vartheta_{jt-\tau}\right] = \sum_{\tau=0}^{R} \phi_\tau E[\vartheta_{jt-\tau}] = \boldsymbol{\vartheta_{jt}} \sum_{\tau=0}^{R} \phi_\tau = \boldsymbol{\vartheta_{jt}} \quad (3.11)$$

The last equality comes from the assumption that $\sum_{\tau=0}^{R} \phi_\tau = 1$ in equation 3.2. An interpretation of equation 3.11 is the following: if the mean of the process that generates expansion vacancies is stable, at least during R+1 periods, the time-varying intercepts in equation 3.8 are equal to the mean of the expansion vacancies in each period. Therefore, by estimating $\widehat{\alpha_{jt}}$ we can get an estimation of the expected value of expansion vacancies, for a given period, ϑ_{jt}. In the first stage of the procedure, we propose in this paper, we estimate equation 3.8 and collect all $\widehat{\alpha_{jt}}$ coefficients.

In the second stage of the procedure, we use all estimated parameters $\widehat{\alpha_{jt}}$ to simulate Poisson distributed variables $\tilde{\vartheta}_{jt} \sim poisson(\widehat{\alpha_{jt}})$. These variables are realizations of the process that generates the expansion variables for each firm j. We test the sensitivity of the vacancies computation to using different distributional assumptions; result turns out to be very similar, regardless of the distribution used (see Appendix C). Once we have simulated realizations of the expansion vacancies, we estimate the following version of equation[5] 3.7:

$$\Delta e_{jt} = (\theta_0 - 1)s_t + \sum_{\tau=1}^{L} \theta_\tau s_{jt-\tau} + \sum_{\tau=0}^{R} \phi_\tau \tilde{\vartheta}_{jt-\tau} + \epsilon_t \quad (3.12)$$

Using bootstrap techniques, we repeat the procedure described before for an amount I of iterations; then, we compute estimators for θ_τ and ϕ_τ as sampling averages from the sample of bootstrap iterations. Analogously, we obtain standard errors for the estimators. With our estimators $\hat{\theta}_\tau$ and $\hat{\phi}_\tau$, in each iteration, we use equations 3.4 and 3.5 to compute stock of replacement vacancies $\hat{V}_t^r$, the stock of expansion vacancies $\hat{V}_t^e$, and the stock of total vacancies $\hat{V}_t = \hat{V}_t^e + \hat{V}_t^r$. We use Monte Carlo simulations to assess the ability of the procedure described before to get unbiased and precise predictions of the different vacancies stock. We explain these Monte Carlo experiments in the next section.

[5] Since we assume in the theoretical model of hiring behavior that $\sum_{\tau=0}^{R} \phi_\tau = 1$ (i.e. all expansion variables are filled eventually), we impose such restriction in the second stage of our applied work.

3.3 Possible Extensions of the Methodology

In the previous section, the set of coefficients estimators $\hat{\theta}_\tau$ and $\hat{\phi}_\tau$, the primary input for the estimation of vacancies stock, describe the hiring behavior of a representative firm in a particular time framework. Nevertheless, there might be heterogeneity across different firms; for instance, these parameters might vary across industries or firm sizes because technologies might be different across these categories. The methodology we propose can adequately accommodate these and other sources of the observed variation. The algorithm explained in section 3.2 can be performed for a subsample of firms according to observable characteristics; therefore, a different family of estimators $\widehat{\theta^k}_\tau$ and $\widehat{\phi^k}_\tau$ can be estimated for category *k*. Finally, using these coefficients and equations 3.4 and 3.5, the stock of vacancies can be computed. Likewise, estimators $\hat{\theta}_\tau$ and $\hat{\phi}_\tau$ can change in long periods because of changes in demographics, preferences, or education. As before, the methodology can accommodate this type of variation, allowing these coefficients to vary in time. Different regimes across subperiods can be modeled to allow such variation.

4. Monte Carlo Simulations.

We test the procedure presented in the previous section using Monte Carlo experiments. We simulate 1000 firms, in 100 time periods; all firms start in the first period with 1000 workers. In an iterative process, we generate 1000 samples of a data generator process[6] characterized for polynomials of length L=3 and R=3. Another characteristic of the data generation process is that separations and expansion vacancies follow a Poisson distribution with mean $\bar{s}_t$ and $\bar{\vartheta}_t$, respectively. To introduce some stability in the data generator process, we assume that $\bar{s}_t$ and $\bar{\vartheta}_t$ vary by regimes within the whole period that the firm is observed, but within each regime they remain unaltered[7].

[6] The real value of the parameters in equation (3) are the following: $\phi_0 = 0.2, \phi_1 = 0.15, \phi_2 = 0.1, \phi_3 = 0.05$; $\theta_0 = 0.4, \theta_1 = 0.3, \theta_2 = 0.2, \theta_1 = 0.10$

[7] By a regime, we mean a collection of periods, each the same length, in which $\bar{s}_t$ and $\bar{\vartheta}_t$ are constant, which we represent by T. At period one we fix $\bar{s}_0$ and $\bar{\vartheta}_0$, for subsequent periods these means are determined as follows: $\bar{s}_T = \left(\frac{1}{T}\sum_{T=0}^{T} \bar{s}_T\right) + \kappa_t$; $\bar{\vartheta}_t = \left(\frac{1}{T}\sum_{T=1}^{T} \bar{\vartheta}_t\right) + \nu_t$ (3.13)

where, $\binom{\kappa}{\nu} \sim \mathbb{N}\left(\begin{matrix} 0 \\ 0 \end{matrix}, \Sigma\right)$ and $\Sigma = \begin{pmatrix} \sigma^2_{\bar{s}_T} & -1/4 \\ -1/4 & \sigma^2_{\bar{\vartheta}_t} \end{pmatrix}$.

Once the series of $\bar{s}_t$ and $\bar{\vartheta}_t$ are computed, separations and expansion vacancies are generated as follows: $s_{jt} \sim \text{Poisson}(\bar{s}_t)$ and $\vartheta_{jt} \sim \text{Poisson}(\bar{\vartheta}_t)$. In each iteration of the experiment we generate s_{jt} and ϑ_{jt} for 1000 firms in 100 time periods. Then, using equation 3.3, we generate hires[8] (h_{jt}). From the identity $e_t = e_{t-1} + h_t - s_t$, we generate firms' total employment in each sample[9]. In each iteration of the procedure, we generate a random sample of the following set of variables s_{jt}, h_{jt}, e_{jt} for each firm *j* in all *t* periods; we use this information to develop the procedure described in section 2. We compute final estimators for θ_τ and ϕ_τ in equation 3.12, as sampling averages of estimators from single iterations; in addition, from the procedure we estimate $\widehat{\bar{\vartheta}_{jt}}$, which is our estimate for expansion vacancies. With estimators $\hat{\theta}_\tau$ and $\hat{\phi}_\tau$, we use equations 3.4 to 3.6 to compute aggregated stocks and flows of expansion, replacement, and total vacancies $\hat{V}_t, \hat{V}_t^e, \hat{V}_t^r, \hat{v}_t, \hat{\vartheta}_{jt-\tau}$, respectively. The flow of replacement vacancies is the share of separations, which are replaced $\widehat{\pi s}_t$ (see equation 3.1).

We compare these predictions with the real values of the Monte Carlo experiments for assessing the performance of the procedure in the prediction of vacancies stocks and flows. Graph number 4.2, shows the predictions of vacancy flows and stocks for three different specification models, but the same data generator process. In all cases, for each iteration of the Monte Carlo experiments, data are generated using a polynomial of length L=3 and R=3 in the hiring function (equation 3.3). The first column in the panel of graphs represents an estimation with L=2 and R=2, the second column presents the estimation with L=3 and R=3, and finally, the last column presents the estimation with L=4 and R=4. The first row of the graph's panel presents the predictions for the stock of total vacancies; the second and third row shows the prediction for the flow of replacement and expansion variables, respectively. In appendix B, we present estimation results for each specification.

Equation 3.13 determines the evolution of the means of the separation and expansion vacancies process; these means are moving averages of up to period $t \epsilon T$ plus a random shock centered in zero and negatively correlated between separations and expansion vacancies.

[8] Measurement error is assumed to be $\epsilon_i \sim \mathbb{N}(0, \sigma_\epsilon^2)$, where σ_ϵ is a fraction of the sampling standard deviation of hires.

[9] In equation 7 we introduce additional measurement error which is assumed to be $\epsilon_i \sim \mathbb{N}(0, \sigma_\epsilon^2)$, where σ_ϵ is a fraction of the sampling standard deviation of hires.

As it can be seen from Graph 4.2, as long as the model is well specified (i.e., the correct amount of lags are included) the method we proposed in Section 3 does a good job in predicting the flows and stocks of replacement, expansion and total vacancies. In such a case, the real value of vacancy stock is always contained within the 95% confidence interval of the prediction. When the model is over-specified (L=4, R=4) the estimation still shows a good performance in predicting the flow of replacement and expansion vacancies, but this performance is poorer in the prediction of the stock of total vacancies. The same is the case when the model is under-specified (L=2, R=2), the prediction of the total vacancy stock is biased, and the same happens to the prediction of replacement vacancy flow. In estimations with different specification, which are not shown for the sake of ease in the presentation, we find the same conclusion; inconsistent estimation of the stock of total vacancies is obtained from over- and under- specified models.

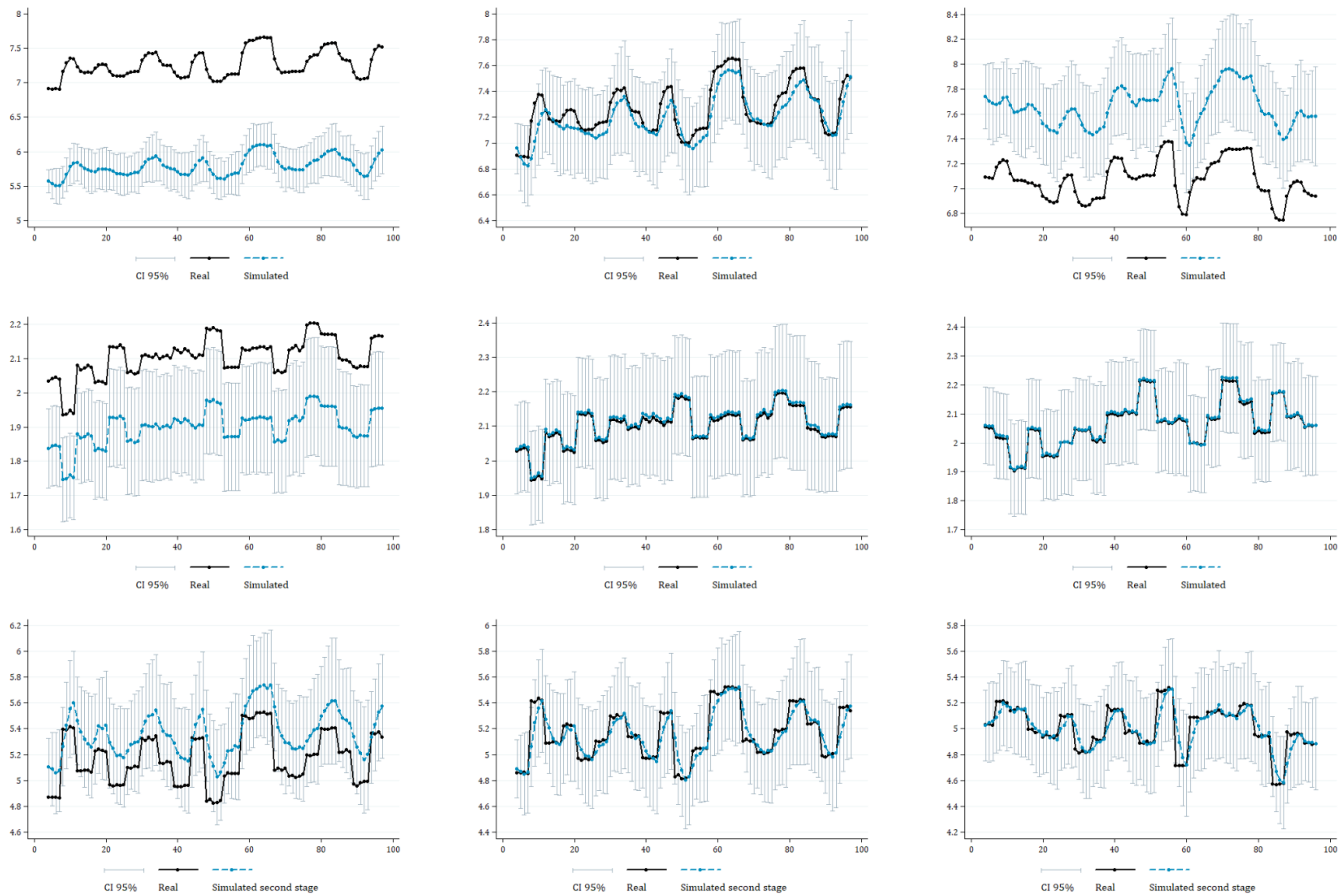
Graph 4.2: Vacancies Predictions

Specification | **L=2, R=2** | **L=3, R=3** | **L=4, R=4**

1. Total Stock of vacancies

2. Flow of replacement vacancies

3. Flow of expansion vacancies



Notes: Data are generated using a polynomial of length L=3 and R=3 in the hiring function. Since we assume in the theoretical model of hiring behavior that $\sum_{\tau=0}^{R} \phi_\tau = 1$ (i.e. all expansion variables are filled eventually), we impose such restriction in the second stage regressions.

5. Vacancies Stock Estimation for Colombian Labor Market

5.1 Data

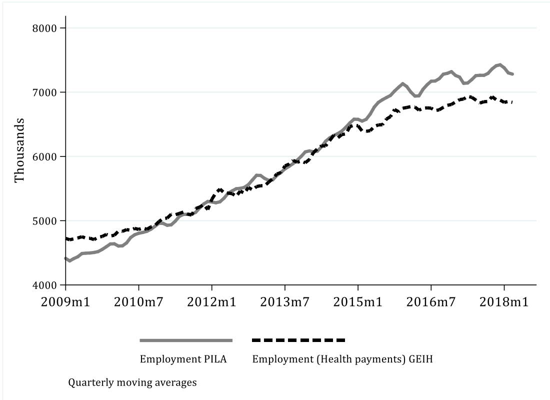
In this section we apply the procedure proposed in section number 3, using for this purpose a Colombian employer-employee linked panel. This panel is generated from administrative records from the “Integrated Record of Contributions to Social Security,” PILA, by its acronym in Spanish[10], which is provided by the Ministry of Social Security in Colombia. PILA is a unique source of longitudinal information containing wages, employment, economic activity, and other characteristics, for the employer and the employee (Morales & Medina, 2019). Using PILA, we analyze the evolution of each firm’s payroll and construct hires, separation and employment size (on a monthly basis). These previous variables, together with job creation and job destruction, are the most standard measures used in the literature on worker and job flows.

As previous studies using PILA have pointed out, one of the advantages of this data is that all labor flows can be computed for all economic sectors (Mejía, Morales, & Medina, 2017; Morales & Medina, 2019). As is shown in graph 5.1, the PILA captures the size of the formal labor market in Colombia well; the graph compares the total formal workers from PILA and the official household survey, GEIH[11]. The solid line represents the employment from firms registered in PILA with more than a single employee; the dotted line represents, at the national level, the formal employees from GEIH, which are formal in the sense that they and their employers pay social security contributions. Graph 5.2 presents the hires, separation flows, and traditional measures of job creation and job destruction. As Graph 5.2 shows, the formal Colombian labor market is relatively dynamic, in 2016 there were more than 7,700,000 formal workers on a monthly basis; during the same year, on average, more than 806,974 and 803,797 monthly hires and separations, respectively, were generated (Flórez et al., 2017).

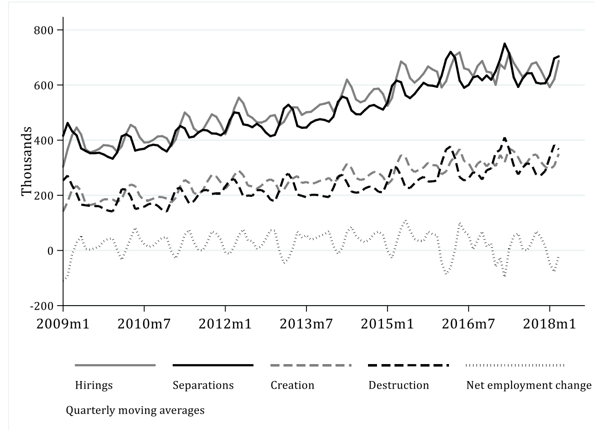
---

[10] As explained in Morales and Medina (2016), the Ministry of Social Security in Colombia collects information from all social security-related payments from employers. This information included firms’ characteristics as wages, employment size, and some information from employees as general socio-demographic characteristics.
[11]by its acronym in Spanish is the official source of all of the labor market indexes in Colombia. The survey is representative of the main 23 Colombian metropolitan Areas.

Graph 5.1: **Employment 2009-2016**[12]



Graph 5.2: **Fluidity Measures**



Source: PILA; Include establishments with at least two employees; authors' calculations

5.2 Vacancies Stock for Colombian Formal Labor Market

We perform the technique proposed in section 3 for two samples; the first one is a sample containing all formal firms with more than 50 employees (average during the whole period), and the second one containing all firms with less than 50 employees and at least five. Therefore, to obtain an estimation of the vacancy stock, we estimate versions of equation 3.8 (first stage) and equation 3.12 (second stage) for samples of big and small Colombian formal firms. To control for any kind of correlation between permanent unobserved heterogeneity and independent variables in equations 3.8 and 3.12 our results are based on fixed effect (FE) panel estimation. For the sake of comparison, we also present some of the results obtained from an Ordinary Least Squares (OLS); nevertheless, fixed effect panel regression will be our preferred specification.

An empirical matter that must be established for the estimation of equation 3.8 and equation 3.12 is the length of the polynomial. We start the estimation of equation 3.8 with parsimonious specifications with just one lag, and then we add more lags, one by one, until the last lag added is significant. The chosen model is the one that fulfills the criterion that all lags are significant and has the minimum root mean square error. In addition, we assume that the time in which firms fill out their replacement vacancies should be similar to the time in which they fill out their expansion vacancies. Therefore, once the optimal number lags are

[12] The employment time series from GEIH excludes self-employed and those whom health insurance payment is not in charge of the employer, employee or both.

determined in equation 3.8, this same number of lags is used in equation 3.12 for the θ and ϕ polynomials.

In table number 5.1 we present the estimation results of equation (3.12); this table is arranged in two different panels, one presenting the results of estimation for big firms (OLS and FE), and one presenting the results for small firms with more than 5 and less than 50 employees (OLS and FE). The best model we identified applying the criteria described before is a model with three lags. We determined the best specification using the FE model, and then we estimated the same specification by OLS for a matter of comparison. As explained in section 3.2, our algorithm is based on fixed effect (by firm and time) estimation. Estimations from our FE specification indicate that the filling process of expansion vacancies intensifies in the first period the position is created. Therefore, the contemporaneous ϕ coefficient is the largest of all. Regarding replacement variables, estimations show that the process of filling replacement vacancies is heavily concentrated in the first month after the replacement vacancy is created, that is why coefficient θ_1 is the largest.

Table 5.1: Estimation Results of Equation 3.12

Parameter	Less than 50 employees						More than 50 employees					
	FE			OLS			FE			OLS		
	Coeff	Std. Err.	t	Coeff	Std. Err.	t	Coeff	Std. Err.	t	Coeff	Std. Err.	t
θ_0	-0,93	0,002	-427,38	-0,86	0,002	-379.26	-1,01	0,02	-56,97	-0,94	0,02	-52.13
θ_1	0,22	0,003	69,14	0,29	0,003	103.34	0,40	0,05	7,77	0,47	0,05	9.20
θ_2	0,05	0,002	25,44	0,10	0,002	46.03	0,14	0,01	9,33	0,19	0,02	10.44
θ_3	0,02	0,002	8,72	0,08	0,002	40.90	0,07	0,02	3,15	0,12	0,02	7.54
φ_0	0,38	0,002	221,03	0,29	0,001	207.49	0,58	0,02	25,04	0,61	0,02	30.12
φ_1	0,26	0,002	163,62	0,25	0,001	197.15	0,18	0,01	12,29	0,16	0,02	10.54
φ_2	0,19	0,002	125,61	0,23	0,001	178.51	0,08	0,01	6,11	0,10	0,02	5.47
φ_3	0,17	0,002	108,60	0,24	0,001	176.65	0,16	0,03	5,43	0,13	0,02	6.36
π		0,36			0,62			0,60			0,84	

Notes: Robust standard errors are computed and clustered by firm. The coefficient that multiplies, contemporaneous separations is $(\theta_0 - 1)$ as denoted in equation 3.12. The π parameter is obtained as $\sum_{\tau}^{l} \theta_{\tau}$. As expressed in equation 3.12, the regression does not include an intercept; nevertheless, results do not change importantly in regressions including an intercept. In these regressions we impose the restriction $\sum_{\tau=0}^{R} \phi_{\tau} = 1$, as it is assumed in the theoretical model.

When firms are born during the period of study, there is a flow and a stock of vacancies, before the very existence of the firm. In this case, we need a correction to recover the vacancies that generated the hires from the first period that newborn firms are observed.

When a firm is born, there is an amount of vacancies equivalent to $\mathrm{h}_{\mathrm{j},0} - \phi_0 \tilde{\vartheta}_{\mathrm{j},0}$, which are the hires in the first period that filled out previously created positions. These additional vacancies correspond to flows and stocks from a firm's pre-existing period. Since our models reveal that firms take up to three months to fill out our their vacancies, we assume that the flow of creation vacancies during the previous three months the firm is born is just $\frac{1}{3}\left(\mathrm{h}_{\mathrm{j},0} - \phi_0 \tilde{\vartheta}_{\mathrm{j},0}\right)$. For newborn firms, the stock of expansion vacancies during this pre-existing period is just the summation of these pre-existing flows. We assume all the vacancies in the pre-existing period are filled out in the period the firm is born[13].

Before presenting the results of our vacancies estimation, we comment on robustness of the methodology to distributional assumptions. In our method we assume that the expansion vacancies are distributed Poisson (i.e., $\tilde{\vartheta}_{jt} \sim poisson(\widehat{\alpha_{Jt}})$); where $\widehat{\alpha_{Jt}}$ are the intercepts estimated in the first stage of the procedure. In Appendix C, we present the stock of total vacancies, computed with the entire sample, using additional discrete and continuous distributions as an alternative to the Poisson. In all cases, these alternative distributions have positive real number or integers as support, which is the space in which simulated vacancies would make sense[14]. The estimation of the total vacancy stock is very similar, regardless of which distribution is used. Furthermore, even though the levels of the predictions may present with minimal differences, the behavior of the series is virtually the same. In light of this evidence, we assume that the arbitrary use of Poisson distribution does not have important implications in the computation of vacancies.

In graph 5.3, we present the total stock of formal vacancies in Colombian labor market, which we obtain from the FE coefficients. We compute the stock of expansion, replacement, and total vacancies using the formulas 3.4, 3.5, and 3.6, respectively. In this graph, we aggregate stock vacancies for small and big firms. The average vacancy stock estimated from FE is

[13] There are different alternatives to develop a correction for newborn firms; one can assume that the behavior of expansion vacancies is equivalent before and after the origin of the firm. We could use the equation for the first period hires of the firm (i.e. $h_0 = \sum_{\tau=0}^{R} \phi_\tau \vartheta_{t-\tau} + \epsilon_t$) and from there solve close forms for the pre-existence expansion vacancies.

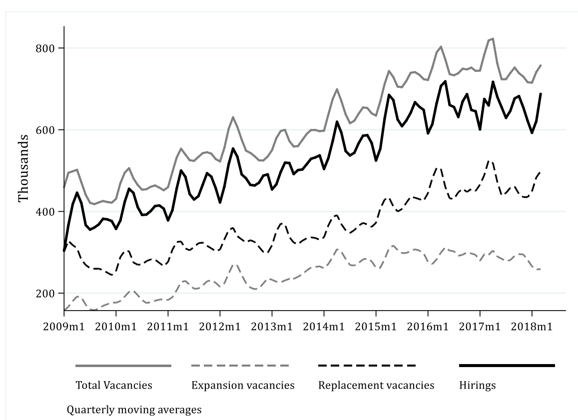
[14] We compute total stock of vacancies with negative binomial distribution, chi-squared distribution, and inverse Gaussian distribution. We explore additional distributions, with very similar results.

665k on average for the study period[15]. From our preferred specification, the FE model, we estimate that 53% of all vacancies correspond to expansion and the remaining 47% correspond to replacement vacancies.

Graph 5.2 shows hires, separations, and employment change from PILA. The average monthly hiring for the study period, (see Graph 5.1), was nearly 574k hires per month, this amount of hires must be generated by a considerable stock of vacancies. The stock of vacancies must be greater than the flow of hires in a given month; this is because there are frictions in labor market and the hiring that requires filling all expansion and replacement vacancies does not occur contemporaneously. Our total stock of vacancies, 665k on average for the study period, is consistent with the sizeable amount of hires observed in Colombian economy for the same period. The magnitude of workers movement is important, therefore, an aggressive and dynamic stock of vacancies must support the magnitudes of these flows.

The official record[16] of vacancy stock, reported by the government agency in charge of gathering information on all vacancies, says that the stock of vacancies in Colombia was 309k, on average for the year 2016; this amount of vacancies is substantially small, considering the magnitude of the worker flows observed in PILA.

Graph 5.3: Vacancy Stocks



[15] This similar to the one would have been obtained from for the OLS estimation (683k).Nevertheless, the composition of this total stock is different; predictions from OLS models overestimate the stock of replacement vacancies and underestimate the stock of expansion vacancies

[16] These records were constructed from quarterly or monthly bulletins available in the *Servicio Público de Empleo* website. “Servicio Público de Empleo" is a relatively new government agency in charge of collecting information on all vacancies in Colombia.

Notes: The stock for expansion, replacement, and total vacancies is constructed using the formulas 3.4, 3.5, and 3.6, respectively, and the coefficients reported in table 5.1. The vacancies and hires presented in this graph correspond to firms with at list 5 employees; also, outliers in the top 99.5% of the distribution of average size are excluded from the analysis.

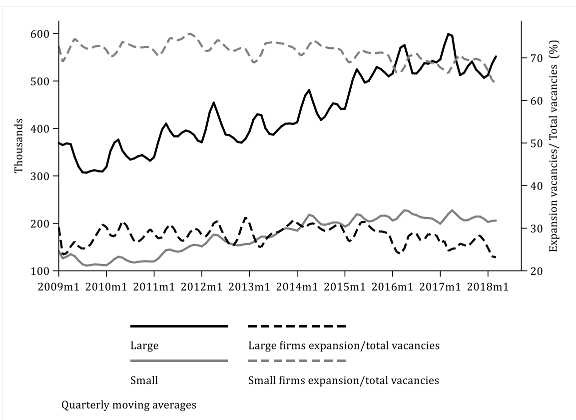
Graph 5.5 presents the vacancies stock for small and big firms; as can be seen from this graph, most of the vacancy stock (nearly 74%) comes from firms with more than 50 employees. Nevertheless, the proportion of expansion vacancies over total vacancies is greater for small firms than for big firms; the percentage of expansion vacancies is 79% in the case of the former and 44% in the case of the latter. Regarding the whole sample, the proportion of expansion to total vacancies is 53%. In the literature on labor market flows, it has been reported that relative to its size, small-medium firms tend to create more jobs than their larger counterparts. In a recent paper, Flórez et al. (2020) conclude that average net employment growth rate is substantially higher for formal firms with less than 50 employees in Colombia. There is plenty of international literature reporting similar patterns for other countries (Anyadike, Bonner, & Hart, 2011; Baldwin & Picot, 1995; Birch, 1981; Broersma & Gautier, 1997).

Graph 5.5 presents the flow of replacement and expansion vacancies estimated from the procedure proposed in section 3.2. We compute the flow of replacement vacancies as $\widehat{\pi}s_t$ (separations in a particular period that will be replaced). The flow of expansion vacancies is computed as the sum of the simulated variable $\widehat{\bar{\vartheta}}_t$ in both type of firms (section 3.2). Graph 5.6 also presents the flow of SPE vacancies, which comes from the governmental agency in charge of collecting information on vacancies in Colombia (Servicio Público de Empleo). This governmental agency has an online platform where all firms are required to post any vacancy they have open.

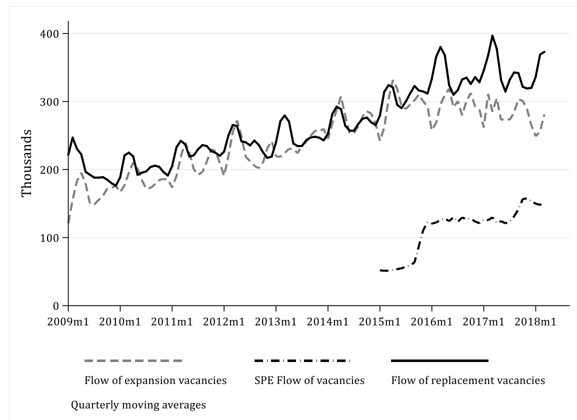
Our estimations reveal that on a monthly basis the formal labor market in Colombia, on average for the study period, produces a continuous flow of 531K vacancies. This flow of 531K monthly vacancies creates 570k hires per month, on average, registered for the Colombian formal labor market during 2009 and 2016 (see Graph 5.2). The number of vacancies registered by the SPE in Colombia is extremely low (an average of 84k for 2014-2016), and it does not reflect the magnitude of monthly hires computed from PILA.

Finally, we comment on the similarity of our vacancies prediction with the results from a different methodology. In Appendix D, we present a comparison between the vacancy index computed from our methodology and an adjusted "help wanted" index, which is calculated using the method proposed in Arango (2013). Surprisingly enough, both series behave similarly, and they share a very similar trend; the correlation between both series is 0,84.

Graph 5.4:



Graph 5.5



Notes: The vacancies and hires presented in this graph correspond to firms with at least 5 employees; also outliers in the top 99.5% of the distribution of average size are excluded from the analysis. The flow of expansion variables correspond to $\tilde{\vartheta}_{jt}$ and the flow of replacement vacancies correspond to πs_t, as is explained in section 3.1 and 3.2. The flow of vacancies from "Servicio Público de Empleo" (SPE) comes from the governmental agency in charge of collecting information on vacancies in Colombia. Reports from SPE are available at www.serviciodeempleo.gov.co.

6. Using Vacancies to Measure Job Creation, Job Destruction, and Worker Replacement.

In the literature on worker and job flows, a very standard measure of Job Creation (Job Destruction) is the employment net growth rate, when this growth is positive (negative). Literature usually presents job creation and destruction as proportions of the second order moving average of employment. These traditional job flow measures could be imprecise. In the case of job Creation, for instance, firms might have destroyed jobs in periods of positive net growth; and it is the same in the case of the job destruction, a firm might have destroyed job positions in periods of positive net growth.

Estimating vacancy flows and stocks allow us to propose alternative measures to job creation and destruction of job positions. With our estimates of expansion and replacement vacancies stocks and flows, we can distinguish creation (or destruction) of new positions and employment growth as two different phenomena.

We propose as an alternative measure of job creation, the ratio of expansion vacancies' flow to total employment size[17], which we refer as Job Positions Creation (JPC). We argue that our measure is more precise than the traditional measure because only expansion vacancies should be associated with the creation of new job positions. As an alternative measure of job destruction, we propose the ratio: total separations minus the flow of replacement vacancies (this are separations that are not replaced) over employment size, we denote this ratio as Job Position Destruction (JPD). Since the separations that are never replaced are equivalent to destroyed job positions, we can identify destruction of job positions independently from employment reductions.

In graph 6.1 we show the creation of new job positions (JPC), expressed as the flow of expansion vacancies as a proportion of firm size, which is 4.7% on average for the study period. The magnitude of the traditional job creation rate is 4.9% on average for this period. Job Creation rate is higher than our measure of Job Positions Creation, but they are not remarkably different[18]. The Job Positions Destruction (see Graph 6.2), which is total separations minus the flow of replacement vacancies, expressed as a percentage of employment size, is 4.5% on average for the study period. The magnitude of the job positions destruction process is just slightly higher than the magnitude of job destruction suggested by the job destruction rate, which is 4.1% on average for the same period[19]. The similarity between our measures and traditional ones reflects that positive (negative) changes in employment are good proxies of creation (destruction) of new positions. Nevertheless, this is an empirical matter and there might be situations in which is not the case.

---

[17] Employment size is computed as a second order moving average of the total number of workers, which is a standard practice in the literature (Davis et al., 1996).

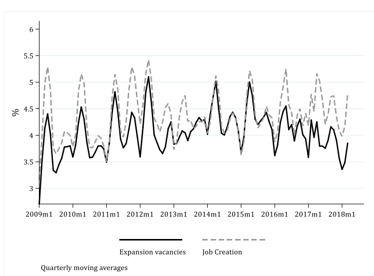
[18] From equation (3.3) it can be seen that the traditional job creation rate will overestimate the creation of new job positions if the flow of previous expansion vacancies plus the number of previous and current replacement vacancies, filled out in the current period are smaller than contemporaneous total separations. (i.e. $if\ h_t - s_t > 0 \rightarrow h_t - s_t > \vartheta_t \leftrightarrow \sum_{\tau=1}^{R} \phi_\tau \vartheta_{t-\tau} + \sum_{\tau=0}^{L} \theta_\tau s_{t-\tau} > s_t$).

[19] From equation (3.3) it can be seen that the traditional job destruction rate will sub-estimate the destruction of previously existing job positions if the flow of current separation that will be created, from the current period on, is smaller than contemporaneous hires (i.e. $if\ h_t - s_t < 0 \rightarrow |h_t - s_t| < 0 \leftrightarrow h_t > \pi s_t$).

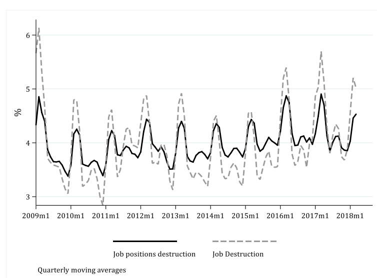
Another standard flow in the literature of labor dynamics is the churning rate, which accounts for the excess of worker reallocation to accommodate job reallocation; churning is computed as (hires - job creation + separations - job destruction)/employment. The idea behind the churning flows is that there is a constant flow of workers quitting; simultaneously, there is a flow of hiring and firing by employers to improve the quality of their workforce (Burgess et al., 2000). Therefore, churning flows implicitly capture worker replacements for fixed positions. Nevertheless, churning rate is a static measure, which does not take into account that hires, at the establishment level, are the ocuppation of vacancies that could have been generated several months before.

In order to capture this idea, we construct the ratio of the flow of replacement vacancies to separations; this is the proportion of all separations that were replaced, we refer to this rate as Worker Replacement Rate (WPR). In Graph 6.3, we show this ratio together with the ratio of churning to worker reallocation rate. As Graph 6.3 shows, the churning as a proportion of all worker reallocation is, on average, 50%; this is a bigger percentage in comparison with our estimations of worker replacement rate, which was 47% on average for the study period.

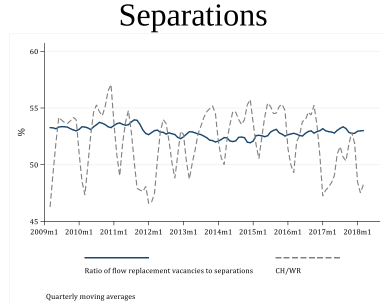
Graph 6.1: Expansion Vacancies vs Employment



Graph 6.2: No Replaced Sep. vs Employment



Graph 6.3: Eventually Replaced Separations vs Separations



Notes:

Graphs 6.1 and 6.2: expansion vacancies, stands for the aggregated flow of expansion vacancies as a proportion of employment size, $\tilde{\vartheta}_{jt}$, in equation 3.12. Job position destruction represents the difference between total separations in the flow of replacement vacancies $(1-\pi)s_t$ in equation 3.1. Job creation and destruction corresponds to traditional measures in the literature $1_{\{\Delta e_{it}>0\}}\Delta e_{it}$ and $1_{\{\Delta e_{it}<0\}}\Delta e_{it}$, respectively, where e_{it} stands for employment levels (see sections 3.1 and 3.2). All measures are expressed as a proportion of the employment order 2 moving average, $\frac{1}{2}(e_{it} - e_{i,t-1})$.

Graphs 6.3: The first series represent the ratio flow of replacement vacancies to total vacancies, which is $\pi s_t/s_t$ in equation 3.1. The second series is the ratio Churning Rate to Worker Reallocation Rate.

6. The Aggregated Behavior of Vacancies, Hiring, and Unemployment.

Using our estimation of vacancies stocks and flows, we compute vacancy rates and vacancy yields for the urban formal labor market, and we estimate a Beveridge curve. All evidence we present in this section show that our vacancies behave accordingly with theory and the economic conditions in Colombia at each moment of our study period. The Vacancy Rate (VR), at period t, is defined as the ratio $\sum_j V_{jt} / \sum_j \left(V_{jt} + O_{jt} \right)$, which represents the vacancies (V_t) as a percentage of total jobs. The total amount of jobs is the summation of vacancies and the total employed population, for the main 23 Colombian Metropolitan Areas. The vacancy yield (VY), at period *t*, is defined as the ratio $\sum_j h_{jt} / \sum_j V_{jt-1}$, which represents hires as a percentage of the stock of vacancies at the end of the previous month.

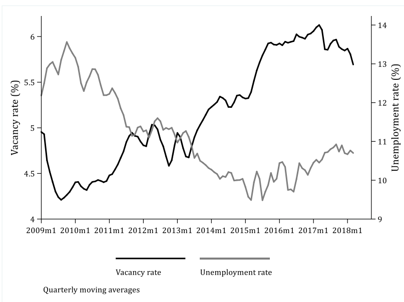
As presented in Graphs 6.4, during this paper's study period, formal vacancy rate went up from 4.5% in the first quarter of 2009 to 6% the last quarter of 2015; this period coincides with an intense reduction on unemployment (Graph 5.1). By the end of 2016, vacancy rate shows a change in trend, and in 2017 decreases importantly; all these coincides with a period of increase in unemployment rate. The vacancy yield rate we obtain has a mean of 0.8 during the study period, and is always less than one (see Graph 6.5). Traditional models of search, predict hires will be a proportion of vacancies (Pissarides, 2000); therefore, one would expect this measure to be smaller than one. The ratio of hires to vacancies that we obtain from our estimation, seems to behave more according to theory predictions, than previous computation of this ratio using raw data on observed vacancies for other countries [20].

Finally, in Graph 6.6 we show a Beveridge curve (BC), which is the relationship between vacancy and unemployment rates. This relationship has its roots in search equilibrium models that predict a negative relationship between vacancies and unemployment. Graph 6.6, presents a locally weighted regression of vacancy rate explained by the unemployment rate. The relationship depicted in Graph 6.6 is in line with the theoretical predictions. The estimated BC shows that from 2011 to 2014 tightness of the Colombian market increases substantially (movements to the left along the BC). After 2015, Colombian economy

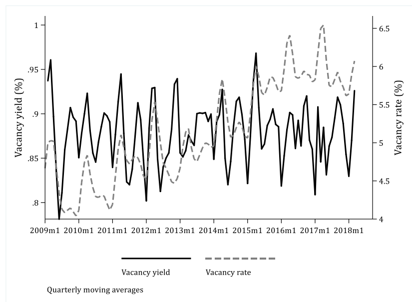
[20] Davis et al. (2013), use JOLTS, a survey of firms that report open positions. The vacancy yield reported in that study is substantially greater than one. In reference to this apparent inconsistency, authors expressed in the paper: (we) "cannot confidently infer the economic relationship between vacancies and hires from raw JOLTS data." In this paper authors' express their concern on the fact that vacancies collected from surveys might be the subject of measurement problems.

decelerated in comparison with previous years, and the unemployment raised. For this period, our Beveridge curve shows, as expected, a reduction on the tightness of the market (movements to the right along the BC)[21].

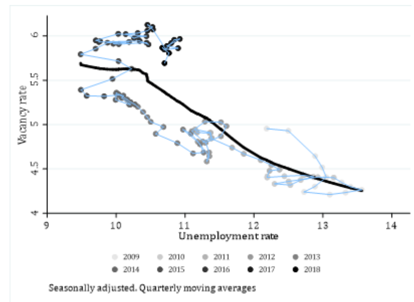
Graph 6.4



Graph 6.5



Graph 6.6



7. Stock and Flow Matching Function Estimation.

Our methodology is very appealing for the estimation of matching functions; given that we are able to recover vacancies flows and stocks, we can test if the matching formation process for the Colombian labor market is random or stock-flow type. There is an important debate in the literature on applied models of equilibrium-unemployment, on the nature of the matching process. On the one hand, Random Matching (RM) assumes that the search frictions are due to informational lacks, mainly on job locations. Therefore, employees and job positions than in a given period were not matched, will feed their respective stocks for the next period. On the other hand, Stock-Flow matching (SFM) assume that after a matching round, the unmatched employees and jobs are simple no match for each other. This could be explained by a mismatch given heterogeneity of agents and vacancies in terms of ability and preferences (Petrongolo & Pissarides, 2001; Sasaki, 2008)

Under the SFM structure, the flows of new vacancies are the most relevant element in the production of new matches; this is because stocks of seekers and vacancies, accumulated from previous rounds, will not match each other. In the case of the RM structure, the main element in the production of new matches is the stock of vacancies; seekers and vacancies

[21] There have been previous attempts to estimate a Beveridge curve for the Colombian urban labor market. Using information from Help Wanted advertisements collected in Arango (2013), Álvarez and Hofstetter (2014) estimate a Beveridge curve with properties similar to what is shown in Graph 7.2, but with substantially lower vacancy rates. For the period 1976-2012, Álvarez and Hofstetter (2014) present an average vacancy rate of nearly 2%. Since HelpWanted advertisements methodologies cannot recover the actual level of real vacancies, they cannot identify the level of vacancy rate either.

unmatched in previous rounds will match, if sufficient time have passed to overcome the lack of information on new job positions' location. Testing which structure of matching fits better with the average Colombian urban labor market would require estimating the following equation:

$$ln(H_{a,t}) = \alpha + \sum_{\tau=\underline{\tau}}^{\tau=\overline{\tau}} \beta_{V\tau,} ln(V_{a,t-\tau}) + \sum_{\tau=\underline{\tau}}^{\tau=\overline{\tau}} \beta_{U\tau,} ln(U_{a,t-\tau}) + \beta_v ln(v_{a,t}) + \beta_u ln(u_{a,t}) + \delta_a + \gamma_t + \epsilon_{a,t} \qquad (7.1)$$

where, $V_{a,t}$ and $U_{a,t}$ stand for stocks of vacancies and unemployed in MA a at time t; in addition $v_{a,t}$ and $u_{a,t}$ stands for last moth flows of vacancies and unemployed. Finally, δ_a, γ_t represents MA and time fixed effects, finally, $\epsilon_{a,t}$ is the error term. The economic and statistical significance of the parameters β_v and β_u provide a test for the fitness of SFM structure. We estimate equation (7.1) for a panel of the 23 main MAs of Colombia for the period September 2009 to march 2018.

Results of estimation are presented in table (7.1) for several specifications. The first column presents the estimation of a static model with no flows; the second column presents the estimation of static model with current period flows of vacancies and unemployed. In the third column, we present an estimation in which we use flow variables -vacancies and unemployed-, and first period lags of the stocks, instead of contemporaneous values. In the Finally, our most complete specification is presented in column (4), in this specification, we include polynomial lags (order 4) for the stock variables and we include vacancies and unemployed current period flows. In the dynamic specifications, (3) and (4), we do not include the contemporaneous stocks because, at least partially, the contemporaneous stock includes the contemporaneous flows, which are included in the regressions as a separate variable.

Our regressions support the random matching formation structure over the stock-flow structure; in all our specifications, contemporaneous flows (unemployed and vacancies) are not statistically significant. Furthermore, in all specifications, only the vacancy stocks are relevant in the generation of new matches. This evidence depicted a labor market, in which frictions caused by informational lacks are important in deterring new matches. Furthermore, because Colombian unemployment rate are high (Graph 6.4), in each period there is a sizeable amount of unemployed; therefore, variations of unemployed population, flow or stock, have a null effect in producing new matchings. In circumstances like this, the classical policy recommendation is enhancing the capability of intermediary institutions in the market; these institutions are key in implementing active labor market policies, which are any managerial or technological effort to help workers and employers to facilitating their match.

Table 7.1: Random vs Stock-Flow Matching Function

	Log (Hires)			
	(1)	(2)	(3)	(4)
	Fixed effects	Fixed effects	Fixed effects	Fixed effects
Log (Vacancy stock)	0.383***	0.498**		
	(0.120)	(0.185)		
Log (1st Vacancy stock lag)			0.747***	0.681***
			(0.076)	(0.060)
Log (2nd Vacancy stock lag)				0.105*
				(0.056)
Log (3rd Vacancy stock lag)				0.132**
				(0.057)
Log (4th Vacancy stock lag)				0.019
				(0.047)
Log (Unemployed)	-0.045	-0.002		
	(0.039)	(0.050)		
Log (1st unemployed lag)			-0.038	-0.052
			(0.028)	(0.034)
Log (2nd unemployed lag)				0.037*
				(0.020)
Log (3rd unemployed lag)				0.012
				(0.026)
Log (4th unemployed lag)				-0.048
				(0.043)
Log (Last month unemployed)		-0.042	-0.027	-0.018
		(0.029)	(0.021)	(0.019)
Log (Vacancy flow)		-0.134	-0.037	-0.099
		(0.166)	(0.070)	(0.074)
Constant	7.927***	8.180***	4.376***	2.027
	(1.447)	(1.455)	(1.228)	(1.428)
Time FE	YES	YES	YES	YES
MA FE	YES	YES	YES	YES
Observations	2691	2691	2668	2599
R-squared	0.977	0.977	0.980	0.981

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: This is a panel regression at the level of metropolitan area (MA), the panel covers from September 2008 to march 2018. Standard errors are clustered at the level of MA.

9. Conclusions and Discussion

In the literature on labor market flows and dynamics, the study of the link between vacancies and traditional worker and job flows measure is surprisingly scarce, only a few empirical studies analyze vacancies and its relation with hires and separations at the firm's level. One of the reasons for this gap in the literature is the difficulty of making a valid inference of the number of vacancies in the economy. There are very limited data sources containing worker/job flows and vacancies at the same time, there are surveys like JOLTS for the US, but they are available for only a few developed countries and contain information only for some limited economic sectors. In addition, when survey information on vacancies is available, there are measurement and aggregation issues that difficult the inference of economic relationships between vacancies and worker flows.

The main contribution of this paper is proposing a methodology that allows the researcher to recover an estimate of vacancies; for this purpose, standard information on hires, separations, and employment at the establishment level is required. With the modernization of the social security systems, this type of data is becoming available for many developing countries; in addition to the PILA for Colombia, there are similar data sets for Brazil (RAIS), and for Mexico (IMSS), just to mention few examples. The methodology we propose is based on the estimation of a hiring function, a function that is expressed in terms of replacement vacancies and expansion vacancies. The distinction between expansion and replacement vacancies has been mentioned previously in the literature (Lazear & Spletzer, 2012), but it has not been explored in depth. From the estimation of the hiring function, we get the parameters we use in the prediction of flows and stocks of replacement, expansion, and total vacancies. Monte Carlo experiments provide evidence that estimations of vacancies stock are consistent, as long the specification for the hiring function is dynamically complete. Finally, the methodology is flexible and can adequately accommodate longitudinal and cross-sectional variation sources by performing the algorithm for heterogeneous samples.

We find that expansion vacancies are nearly 53% of total vacancies; the most important component of the total vacancy stock comes from firms' expansions or creation, which is not surprising for a period of strong employment growth. Big firms, with more than 50 employees, create 74% of total vacancies. Nevertheless, the proportion of expansion vacancies over total vacancies is greater for small firms than for big firms; the percentage of expansion vacancies is 79% in the case of small firms, and 44% for large firms.

Our methodology allows computing alternative job flow measures, which we argue, can distinguish the creation (destruction) of new positions and positive (negative) employment growth as two different phenomena. We propose as well a measure of worker replacement which is similar in spirit to the ratio churning/worker-reallocation. There are differences between our proposed measures and traditional job flow and churning rates; nevertheless, they are not very sizeable. In a sense, this similarity provide validation to the interpretation of traditional job flows as creation and destruction of job positions.

Indirect tests for the validity of our estimated vacancies are encouraging. For instance, we obtain a vacancy yield that exhibits a predicted-by-theory behavior, which is in line with the economic intuition from search models. For the average firm, the ratio is always smaller than one. This is not a standard finding; previous estimations for other countries have estimated this ratio to be substantially higher than one. In addition, our estimated vacancies exhibit consistent behavior in relation with variables measuring the business cycle. Vacancy rates show a pro-cyclical behavior, and we find a downward-sloped estimated Beveridge curve.

Finally using our predicted vacancies flows and stock, we estimate matching functions; we are able to test if the matching formation process for the Colombian labor market is random or stock-flow type. We find support for the random matching formation structure over the stock-flow structure; in all our specifications, contemporaneous flows are not statistically significant. This result implies that the most relevant search frictions in the Colombian labor market are due to informational lacks, mainly on job locations; therefore, it is not necessary the case that there is a structural mismatch given heterogeneity of agents and vacancies. The policy recommendation in this situation is enhancing the capability of intermediary institutions in the market; these institutions are key in implementing active labor market policies.

To the best of our knowledge, a method of estimating vacancies from hiring behavior has not been documented previously in the literature. The method we propose may be useful in developing economies, where there are no good official sources of information on open job positions, nevertheless there are public records of social security systems with information on firm's payroll. Applications or enhancements of the methodology proposed in this paper may contribute to the analysis of vacancies and its relationship with different labor market flow and indexes, especially in economies where data availability may be a limitation.

References

Abraham, K. G., & Wachter, M. (1987). Help-wanted advertising, job vacancies, and unemployment. *Brookings Papers on Economic Activity*, *1987*(1), 207–248.

Álvarez, A., & Hofstetter, M. (2014). Job vacancies in Colombia: 1976–2012. *IZA Journal of Labor & Development*, *3*(1), 1–11.

Amoah, B. (2000). Help-wanted index: The association between the HWI and employment and unemployment rates 1981-1999. *Perspectives on Labour and Income*, *12*(2).

Anyadike, D. M., Bonner, K., & Hart, M. (2011). Job Creation and Destruction in the UK. *Economics & Strategy Group, Aston University*.

Arango, L. E. (2013). Puestos de trabajo vacantes según anuncios de la prensa escrita de las siete principales ciudades de Colombia. *Borradores de Economía*, *793*.

Baldwin, J., & Picot, G. (1995). Employment generation by small producers in the Canadian manufacturing sector. *Small Business Economics*, *7*(4), 317–331.

Birch, D. L. (1981). Who creates jobs? *The Public Interest*, (65), 3.

Broersma, L., & Gautier, P. (1997). Job creation and job destruction by small firms: An empirical investigation for the Dutch manufacturing sector. *Small Business Economics*, *9*(3), 211–224.

Burgess, S., Lane, J., & Stevens, D. (2000). Job flows, worker flows, and churning. *Journal of Labor Economics*, *18*(3), 473–502.

Davis, S. J., Faberman, R. J., & Haltiwanger, J. C. (2013). The establishment-level behavior of vacancies and hiring. *The Quarterly Journal of Economics*, *128*(2), 581–622.

Davis, S. J., & Haltiwanger, J. (2014). *Labor market fluidity and economic performance*. National Bureau of Economic Research.

Davis, S. J., Haltiwanger, J. C., & Schuh, S. (1996). *Job creation and destruction*. MIT Press.

Flórez, L. A., Morales, L., Medina, D., & Lobo, J. (2017). *Labour flows across firms size, economic sectors and wages: evidence from employer-employee linked panel*. Banco de la Republica de Colombia.

Flórez, L. A., Morales, L. F., Medina, D., & Lobo, J. (2020). Labor flows across firm size, age, and economic sector in Colombia vs. the United States. Small Business Economics, 1-32.

Holt, C., & David, M. (1966). The concept of job vacancies in a dynamic theory of the labor market. In *The measurement and interpretation of job vacancies* (pp. 73–110). NBER.

Lazear, E. P., & Spletzer, J. R. (2012). Hiring, churn, and the business cycle. *The American Economic Review*, *102*(3), 575–579.

Mejía, J. F., Morales, L. F., & Medina, D. S. (2017). Trade liberalization and its effects on labor fluidity: Evidence from Colombia. *International Trade Journal*, *32*(1), 43–75. https://doi.org/10.1080/08853908.2017.1389324

Morales, L., & Medina, D. (2016). Labor Fluidity and Performance of Labor Outcomes in Colombia: Evidence from Employer-Employee Linked Panel. *Borradores de Economía*, (926).

Morales, L., & Medina, D. (2019). Fluidez del mercado laboral y resultados en materia de empleo en Colombia: evidencia derivada de datos enlazados de empleadores y empleados. *Revista CEPAL*, (127).

Morales, L. F., Hermida, D., & Dávalos, E. (2019). The Interaction between Formal and Informal labor dynamics: Revealing Job Flows from Household Surveys. Borradores de Economía, 1090, 34.

Petrongolo, B., & Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, *39*(2), 390–431.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Sasaki, M. (2008). Matching function for the japanese labour market: Random or stock–flow? *Bulletin of Economic Research*, *60*(2), 209–230.

Appendix

Appendix A, Computation of Vacancies Stocks

The stock of expansion vacancies consists of vacancies that were generated in previous periods, but that has not been filled completely. The following expression can represent this stock of expansion vacancies at the end of period t:

$$V_t^e = (1 - \phi_0 - \phi_1 - \cdots - \phi_{\mathrm{R}-1})\, \vartheta_{t-R-1} + (1 - \phi_1 - \cdots - \phi_{\mathrm{r}-2})\, \vartheta_{t-R-2} + \cdots + (1 - \phi_0)\, \vartheta_t \quad (3.4)$$

The first term of equation 3.4 expresses the fact that at period t most of the expansion vacancies generated at period $t - R - 1$ have already been filled, but there is a fraction $\phi_R = 1 - \phi_0 - \phi_1 - \cdots - \phi_{\mathrm{R}-1}$ that has not been filled yet and it will be filled the next period. This is because on average, the firms fill all expansion vacancies in R periods. Following the same logic, the last term of equation 3.4 represents the fact that at time t the only vacancies generated at that period that have been filled are $\phi_0\, \vartheta_t$. Analogously, the stock of replacement vacancies can be represented by the following equation:

$$V_t^r = (\pi - \theta_0 - \theta_1 - \cdots - \theta_{L-1}) s_{t-L-1} + (\pi - \theta_0 - \theta_1 - \cdots - \theta_{L-2}) s_{t-L-2} + \cdots + (\pi - \theta_0) s_t \qquad (3.5)$$

The first term of equation 3.5 expresses the fact that at period t there is a fraction $\theta_R = \pi - \theta_0 - \theta_1 - \cdots - \theta_{L-1}$ of replacement vacancies that have not been filled yet, and as in the case of expansion vacancies, it will be filled the next period. The last term of equation 3.5 represents the fact that at time t the only replacement vacancies generated from contemporaneous separations that have been filled are $\theta_0 s_t$.

The flow of expansion vacancies is the number of these vacancies generated at a given period, and it will be denoted as $v_t^e = \vartheta_t$. The flow of replacement vacancies at a given period is the number of the separations that will be replaced in that period or the future period, and it will be denoted as $v_t^r = \pi s_t$. The total number of vacancies, the vacancies stock at a given period t, is given by the addition of equation 3.4 and 3.5, and it can be represented as:

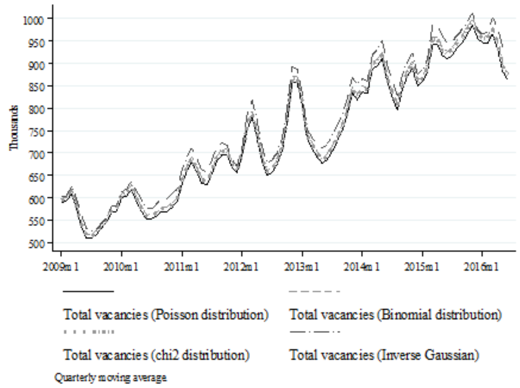
$$V_t = V_t^e + V_t^r \quad (3.6)$$

Appendix B: Monte Carlo Simulations.

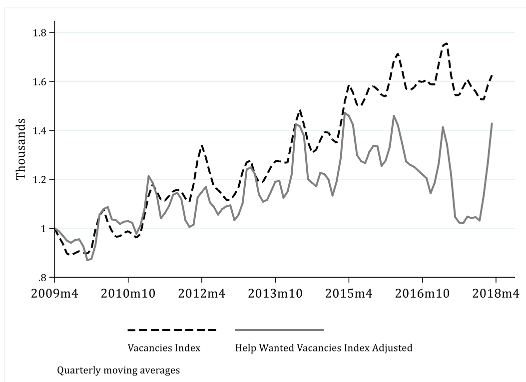
Series and coefficients		Number of lags 1	2	3	4
Mean Stock of vacancies	Real	6.99	7.25	7.25	7.07
	Estimation	3.19	5.77	7.20	7.66
	RMSE	4.27	1.94	1.35	1.57
Mean Flow of replacement vacancies	Real	2.02	2.10	2.10	2.06
	Estimation	1.41	1.89	2.11	2.07
	RMSE	0.72	0.25	0.10	0.11
Mean Flow of expansion vacancies	Real	4.98	5.16	5.16	5.01
	Estimation	5.58	5.36	5.15	5.00
	RMSE	1.32	1.18	1.16	1.12
Phi Estimation Set	**Phi0 = 0.40**	0.53	0.43	0.41	0.40
		0.00	0.00	0.00	0.00
	Phi1 = 0.30	0.47	0.32	0.29	0.28
		0.00	0.00	0.00	0.00
	Phi2 = 0.20		0.24	0.18	0.17
			0.00	0.00	0.00
	Phi3 = 0.10			0.12	0.09
				0.00	0.00
	Phi4 = 0.00				0.07
					0.00
Theta Estimation Set	**Theta0 = 0.20**	0.20	0.20	0.20	0.20
		0.00	0.00	0.00	0.00
	Theta1 = 0.15	0.15	0.15	0.15	0.15
		0.00	0.00	0.00	0.00
	Theta2 = 0.10		0.10	0.10	0.10
			0.00	0.00	0.00
	Theta3 = 0.05			0.05	0.05
				0.00	0.00
	Theta4 = 0.00				0.00
					0.00

Notes: The number of regimes is 22. The standard errors are in parenthesis. The same number of lags is applied to the phi and theta's polynomials.

Appendix C: Total Vacancies of different distributions for expansion vacancies



Appendix D: Vacancy Index from hiring behavior vs. Help Wanted Vacancy Index



Notes: Both indexes have January of 2009 as their base. The Help wanted vacancy index is an actualization of the results in Arango (2013) kindly provided by the author.