The Job Openings and Labor Turnover Survey (JOLTS) has quickly captured the attention of macroeconomists studying labor markets after the survey’s launch in December 2000. The enthusiasm of macro-labor economists about JOLTS is easy to understand: job openings (more commonly referred to as vacancies) play a crucial role in equilibrium models of unemployment that have been developed in the 1980s and 1990s. These models (following the pioneering work of Diamond [1981, 1982a, 1982b], Mortensen [1982a, 1982b], and Pissarides [1984, 1985]) have proved to be very fruitful in analyzing a wide range of aggregate labor-market issues: the existence of unemployment as an equilibrium phenomenon, the ongoing high rate of worker reallocation observed in labor markets, or the effect of policies that influence the operation of labor markets. Data on vacancies comparable to the series available in JOLTS had never been collected previously in the United States. Moreover, not only did JOLTS provide a much-needed superior measure of vacancies, it did so at a time when research on models emphasizing the role of vacancies has been very active. In addition, the JOLTS series had the unintended fortunate timing of beginning just as the long expansion of the 1990s was coming to an end. Capturing the state of the labor market just before the start of the 2001 recession thus allowed JOLTS to be informative about cyclical variation with a relatively short time span. In light of these facts, it is hard to overstate the enthusiasm of the macro-labor research community in response to the availability of the JOLTS. Faberman’s chapter in this volume (chapter 2) is an excellent overview of this new data source and should be on the reading list of anyone wishing to work with the JOLTS data.

Éva Nagypál is an assistant professor of economics at Northwestern University.
Despite the enthusiasm that the launch of JOLTS has created, this new data source has not yet been closely scrutinized to determine how it can be used to validate prevailing theories of recruitment. My chapter’s intent is to push the discourse on JOLTS in this direction. I start by reviewing some methodological and conceptual issues that arise when using JOLTS data. In particular, I first discuss the issue of labor turnover measurement and the problem of missing separations in the JOLTS data. I then discuss how the JOLTS definition of vacancies relates to the definition of vacancies used in theoretical models, and highlight how possible discrepancies between the definitions need to be taken into account when doing empirical work using JOLTS data. In the second part of the chapter, I use the publicly available JOLTS data to study empirically the widely used theoretical construct of the matching function. This allows me to demonstrate one of the many ways that the JOLTS data can serve to test existing theories of labor-market dynamics and provide new evidence to inform the development of these models in new directions. Throughout, the concepts and definitions that I use are equivalent to those used in Faberman’s chapter, though I limit myself to using the publicly available aggregate and industry data.

3.1 Consistency of JOLTS Turnover Data

A distinct advantage of JOLTS is that it directly measures gross worker flows from the employer perspective (i.e., hires and separations) as opposed to simply measuring net employment change at establishments. Thus, the JOLTS gives a richer picture than available from other data sources about the margins that firms use to adjust their level of employment. There is, of course, a tight relationship between hires, separations, and net employment change at the level of an establishment, since, by definition,

\[ \Delta e_J = e_{J+1} - e_J = h_J - s_J \]

where \( e_J \) is the level of employment at establishment \( J \) at the beginning of period \( t \), and \( h_J \) and \( s_J \) are the number of hires and separations at establishment \( J \) during period \( t \). Summing over all establishments in some set \( J \) (for example, the set of all nonfarm establishments, or the set of establishments in a particular industry) gives two alternative measures of employment growth over period \( t \):

1. \[ \Delta e^1_J = \sum_{j \in J} \Delta e_J = \sum_{j \in J} e_{J+1} - e_J = e_{J+1} - e_J \]
2. \[ \Delta e^2_J = \sum_{j \in J} h_J - s_J = h_J - s_J \]

where \( e_J \) is the level of employment across all establishments in \( J \) at the beginning of period \( t \), and \( h_J \) and \( s_J \) are the total number of hires and sepa-
rations at all establishments in $J$ during period $t$. Equation (1) gives a way to measure aggregate employment growth using employment data, the best measure of which, for the same universe of establishments as the one covered by JOLTS, is given by the Current Employment Survey (CES). This measure of employment growth can then be compared with the aggregate employment growth calculated using equation (2) based on labor turnover data in JOLTS, giving a way to assess the consistency of the JOLTS turnover data.

To the extent that the JOLTS and the CES cover the same universe of establishments and the JOLTS weighting scheme is explicitly adjusted to match the CES level of employment, the correspondence between the two measures of employment growth should be very close. Beyond sampling error, there is only one reason that the correspondence between the two measures of employment growth cannot be expected to hold month by month—the difference in reference periods. The JOLTS turnover data refer to the period between the first day of the month and the last day of the month, while employment in the CES measures employment during the pay period that includes the twelfth of the month. Calculating employment growth over horizons longer than a month, however, should diminish both the effect of any sampling error and the effect of the difference in the reference period.

Figure 3.1 plots aggregate employment growth from December 2000 onwards calculated from the CES data and from the JOLTS data using equations (1) and (2). According to the CES data, total employment declined by 59,000 workers in the United States between December 2000 and December 2004, which is in line with the poor employment performance of the U.S. economy during and following the 2001 recession. At the same time, according to the JOLTS data, the number of employed grew by 4.64 million during the same period, representing over 3.5 percent of total employment. This is a large discrepancy. To the extent that (a) the CES is a much larger survey that is designed explicitly to determine the level of employment in the United States and (b) the stock of employment is easier to measure than the flow into and out of employment, one can attribute all the discrepancy between the two measures of employment growth to measurement problems in the JOLTS turnover data. This discrepancy has been identified earlier by Wohlford et al. (2003). In fact, as a result of internal studies by BLS staff that uncovered the same discrepancy, there have been some changes in 2002 in the way the JOLTS data were collected, with the survey instrument redesigned for schools and temporary help agencies. These changes have reduced the size of the above discrepancy, but have not eliminated it. To show this, figure 3.2 plots aggregate employment growth for four year-long periods based on the CES and the JOLTS data. The overstatement of employment growth by JOLTS was the largest early in the survey, between December 2000 and December 2001 (2.29 million), but it
Fig. 3.1 Aggregate employment growth in the JOLTS and in the CES data since December 2000

Fig. 3.2 Aggregate employment growth in the JOLTS and in the CES data since the beginning of the year for each year between 2001 and 2004
remained positive in all subsequent years; it was 0.57 million between December 2001 and December 2002, 1.00 million between December 2002 and December 2003, and 0.84 million between December 2003 and December 2004.

Moreover, the aggregate annual employment growth discrepancy of 0.7 percent for 2003–2004 masks substantial industry variation in annual employment growth discrepancy (measured as $\frac{1}{2} \times \sum_{i=Dec2004}^{Jan2003} (\Delta e_{it}^2 - \Delta e_{it}^1)$ for industry $i$), which is plotted on the vertical axis of figure 3.3. As can be seen for 2003–2004, the annual overstatement of employment growth by JOLTS varies from a high of 2.58 percent in the Federal Government to a low of –3.13 percent in construction. This large industry variation implies that the mismeasurement of labor turnover in the JOLTS is a larger problem than seems at first from the aggregate data.

There is reason to believe that the discrepancy in the JOLTS arises in large part due to the mismeasurement of the separation rate. To show this, I calculated for each two-digit North American Industry Classification System (NAICS) industry the average JOLTS separation rate for the period January 2003–December 2004 and the average separation rate from the Current Population Survey (CPS) for the same period. On average, the separation rate calculated from the CPS is 1.9 times as large as the separation rate calculated from JOLTS. This is due both to the understatement of separations in JOLTS and to the overstatement of separations in the CPS due to the well-known classification problem (Nagypál 2006). There is large cross-industry variation in the ratio of the JOLTS to the CPS separation rate, however, ranging from the JOLTS separation rate being a third of the CPS separation rate in education to two-thirds in mining. Moreover, it is exactly the industries that have a very low measured JOLTS separation rate compared to the CPS separation rate that have the largest overstatement of their employment growth in the JOLTS hires and separation data. This can be seen from figure 3.3, where I plot the average annual employment growth discrepancy for the period January 2003–December 2004 between the JOLTS and the CES against the ratio of the JOLTS separation rate to the CPS separation rate for each industry. This evidence is suggestive that the underestimation of the separation rate is a key reason that the JOLTS data overstate employment growth in the U.S. economy.

Further examination of the JOLTS employment growth discrepancy across industries also reveals that a relevant characteristic of industries that is correlated with the size of this discrepancy is the average level of

1. The CPS started using the NAICS industry classification of the JOLTS after January 2003. The separation rate in the CPS can be derived by matching the Basic Monthly Survey across two consecutive months and calculating the ratio of the sum of employer-to-employer and employment-to-nonemployment transitions between the two months to the number of employed workers during the first month.
education in the industry. Figure 3.4 plots the average years of education in each two-digit NAICS industry, calculated using CPS data from 2003–2004 against the average annual employment growth discrepancy for the period January 2003 to December 2004 between the JOLTS and the CES. Clearly, this figure implies that the overstatement of employment growth is a larger problem for more educated workers, a pattern that is worthy of further investigation and that could inform future revisions of JOLTS data collection.

To assess the impact of the employment growth discrepancy between the JOLTS and the CES on the measurement of labor turnover, I use a simple procedure to adjust hires and separations for this discrepancy by industry according to

\[
\tilde{h}_{it} = h_{it} + \max(0, \Delta e_{it} - \Delta e_{it})
\]

\[
\tilde{s}_{it} = s_{it} + \max(0, \Delta e_{it}^2 - \Delta e_{it}^1)
\]

where \(h_{it}\) (\(s_{it}\)) and \(\tilde{h}_{it}\) (\(\tilde{s}_{it}\)) is the measured and adjusted number of hires (separations) in industry \(i\) in month \(t\), respectively. To do this adjustment, I estimate the employment growth for month \(t\) for industry \(i\) in the CES by extrapolating the employment numbers for the pay period containing the twelfth of the month. To the extent that this adjustment merely requires that employment growth numbers match up industry-by-industry at the two-digit level as opposed to establishment-by-establishment, this procedure underadjusts the hires and separations numbers, thus giving a lower
bound on the true hiring and separation rate. This procedure results in an adjusted aggregate hiring rate of 3.62 percent as opposed to the measured hiring rate of 3.31 percent, and in an adjusted aggregate separation rate of 3.62 percent as opposed to the measured separation rate of 3.23 percent, a significant change.

### 3.2 What do the JOLTS Job Openings Measure?

Beyond giving a more detailed view of labor turnover, a distinct advantage of JOLTS is that it provides information on the number of job openings for a representative sample of U.S. establishments, thereby giving a much more direct measure of vacancy creation in the U.S. economy than was previously available (primarily through the use of the Help Wanted Advertising Index). Of course, to develop a measure of job openings, the BLS had to construct an appropriate empirical definition. Faberman reviews this definition in chapter 2. Here, I would like to discuss the impact of two choices in the construction of this empirical definition: the choice to measure the stock of vacancies at a point in time as opposed to their flow during a period, and the choice to include only vacancies for positions that can start within thirty days.

2. At the same time, given that the employment growth number is estimated using extrapolation and could contain errors, this procedure could possibly overadjust the hires and separations numbers.
To focus the discussion, consider the following simple continuous-time model of vacancy creation, where time is measured in months. Assume that a firm wishes to hire someone to start working at some known future date $t_s$. Due to search frictions in the relevant labor market, appropriate candidates are not always immediately available for hire; rather, they arrive to the firm at random times if the firm has a vacancy open. In particular, assume that if the firm has a vacancy open, suitable candidates arrive at Poisson rate $\lambda$, which (approximately) means that during a short period of length $\Delta$, the probability that a suitable candidate shows up is $\lambda \Delta$. Assume that hiring a candidate at time $t_c \leq t_s$ has a cost of $c_e(t_s - t_c)$ to the firm. Such a cost could arise due to having to incur some expenses to keep the candidate available between the time he or she is offered the position at time $t_c$ and the time he or she starts working at time $t_s$. Assume that hiring a candidate at time $t_d > t_s$ has a cost of $c_d(t_d - t_s)$ to the firm. Such a cost could arise due to forgone profits from starting the position late. Finally, assume that the firm chooses the time to open a vacancy to minimize the expected cost of hiring too early or too late compared to time $t_s$. Under these assumptions, one can show that the firm will optimally open the vacancy at time $t_v = t_s - l$, where $l$, the lead time to open a vacancy, is given by

$$l(\lambda, r_d) = \frac{\log (1 + r_d)}{\lambda}$$

where $r_d = c_d/c_e$ is the relative cost of delay. This simple model has the intuitive implication that the harder it is to find a suitable candidate (i.e., the lower is $\lambda$) and the higher is the cost of delay relative to the cost of early hiring, the earlier will the firm decide to open a vacancy relative to the time of the intended start of the job.

To see why this simple model is useful to think about the measurement of job openings in the JOLTS, assume that at each point in time firms wish to hire a fixed measure of workers. Then one can calculate the probability that a vacancy that is open at some point during the month $[t_o - 1, t_o]$ is observed at time $t_o$ (without any restrictions on when the position starts) to be

$$P^u(\lambda) = \frac{1}{1 + \lambda}.$$

The solid line in figure 3.5 plots this probability of observing a vacancy as a function of $\lambda$. The interpretation of this probability is simple: jobs with a higher arrival rate of suitable applicants have a vacancy open for a shorter period of time, hence these vacancies have a lower probability of being observed given a fixed frequency of observation. This is a well-known issue in duration models—whenever duration events are sampled using stock sampling (as in the JOLTS), events of short duration are less likely to be

3. Derivations of all the results shown are available upon request.
sampled. One can use statistical methods developed in duration analysis to address this issue (see Lancaster 1990) and reconstruct the flow of vacancies from the stock data.

The dependence of the probability of observation on the rate of arrival of suitable applicants has at least two important implications. First, different probability of observing vacancies due to different arrival rate of suitable applicants could be one explanation for why the vacancy-to-hires ratio varies substantially across industries, from a ratio of 0.30 in construction to a ratio of 1.48 in health. It is possible that the number of new vacancies opened per new hire is the same in these industries, and what is different is how long the average vacancy in the industry is open due to the relative ease with which a vacancy in construction can be filled and the relative difficulty with which a vacancy in health can be filled. This interpretation of the data is supported by the strong positive correlation between the vacancy-to-hires ratio and the average education of workers across industries, shown in figure 3.6. Second, to the extent that there is systematic variation in $\lambda$ over the business cycle, with recessions being times when openings are easier to fill and hence $\lambda$ is higher, the probability of observ-

![Graph showing the probability of observing a vacancy as a function of the arrival rate of candidates and the relative cost of delay in the simple model of vacancy creation.](image-url)

**Fig. 3.5** Probability of observing a vacancy as a function of the arrival rate of candidates and of the relative cost of delay in the simple model of vacancy creation

*Note: $P^u$ is the unrestricted probability while $P^r$ is the probability restricted to include only vacancies where the position is available within one month.*
ing a vacancy is procyclical in the above simple model, implying that the cyclical variation in the stock of vacancies overstates the variation in the flow of vacancies.

The above simple model also helps us understand a second potential problem with the JOLTS measurement of vacancies. Recall that the JOLTS definition of a vacancy requires that work could start within one month of the day of measurement. This means that vacancies that are opened with long lead times (either because \( \lambda \) is low or because the relative cost of delay in hiring is high) are not counted in the JOLTS definition of job openings. In particular, one can show that the probability that a vacancy that is open at some point during the month \([t_o - 1, t_o]\) is observed at time \(t_s\) given that only vacancies where the position is available within a month (i.e., where \(t_s \leq t_o + 1\)) are counted is

\[
P^r(\lambda, r_d) = \begin{cases} 
\frac{1}{1 + \lambda} \frac{e^\lambda}{1 + r_d} & \text{if } \lambda < \log (1 + r_d) \\
\frac{1}{1 + \lambda} & \text{if } \lambda \geq \log (1 + r_d).
\end{cases}
\]

The two dashed lines in figure 3.5 plot this probability of observation as a function of \( \lambda \) for two different values of the relative cost of delay, a low value of \( r_d = 1 \) and a higher value of \( r_d = 3 \). Under the JOLTS definition, this simple model implies that jobs with a higher relative cost of delay and where suitable workers arrive less frequently are less likely to be observed and counted compared to the case where all vacancies are counted irrespective of the time the position is available. The reason for this is simple: for jobs with a higher relative cost of delay and where suitable workers arrive less frequently, it is optimal to open vacancies with a long lead time and, as a consequence, often workers are hired for these jobs long before they start working for the employer. Vacancies for such jobs (for example, those for academics) are systematically under measured using the JOLTS definition. This under measurement could go some way toward explaining why the education industry in figure 3.6 lies much below the regression line. The statistical tools to address this measurement issue are less readily available than the tools to use in case of stock sampling. In my opinion, the best way to address this measurement issue would be to acquire additional data on vacancies where work is expected to start further into the future than one month.

3.3 Using JOLTS to Study the Matching Function

The JOLTS data on vacancies allows for the empirical examination of the theoretical construct of a matching function using more direct measure of vacancies and new hires than previously available. In equilibrium models
of unemployment, the matching function is a theoretical construct that is used to describe how workers and firms meet in a frictional labor market. In particular, it posits that the flow of new matches between workers and firms is a function of the number of workers looking for employment and the number of vacancies that are opened by firms. Assuming that only unemployed workers look for employment (a commonly maintained assumption), the matching function posits that

$$m_t = m(v_{t-1}, u_{t-1})$$

where $m_t$ is the number of new matches created during period $t$, and $v_{t-1}$ and $u_{t-1}$ are the number of vacancies and unemployed workers looking to form employment relationships at the end of period $t - 1$. The number of new matches created can be measured using the hires data in JOLTS (i.e., $m_t = h_t$), so the JOLTS data provides two of the three data series necessary to estimate an aggregate matching function. Assuming a log-linear functional form for the matching function and an additive error term gives the empirical specification

$$\log h_t = \beta_v \log v_{t-1} + \beta_u \log u_{t-1} + \epsilon_t.$$ 

In terms of empirical implementation, there are several issues that need to be addressed. First, should one use seasonally adjusted or unadjusted data? Second, should the matching function be estimated using aggregate or industry-level data? Both of these questions turn out to be empirically relevant.
To show this, I first use the seasonally adjusted JOLTS data and the seasonally adjusted number of unemployed from the CPS and estimate equation (3) by ordinary least squares (OLS). Table 3.1 reports estimation results for this empirical specification using data from December 2000 to November 2004. Under this specification, the hypothesis that the matching function has constant returns to scale cannot be rejected, and the elasticity of the matching function with respect to vacancies is estimated to be 0.67. This estimate of the elasticity is substantially larger than the matching function elasticity of 0.3 to 0.5 derived by Petrongolo and Pissarides (2001), though given the small sample size, the standard errors on the estimates are rather large.

The second column of table 3.1 reports estimation results when seasonally unadjusted data are used and \( \beta_{cm} \) is allowed to vary with the month \( m \). Now, the coefficients both on the vacancy rate and the unemployment rate are lower, though the hypothesis that the matching function has constant returns to scale still cannot be rejected given the small sample size. Even with the small sample size, however, one can reject the hypothesis that the scale parameter \( \beta_{cm} \) is the same for all months \( m \) even at the 99 percent level of confidence. Figure 3.7 plots the estimate of \( e^{\beta_{cm}} \), which can be thought of as an estimate of matching efficiency, for each month \( m \). There is a clear seasonal pattern in this matching efficiency, with the summer months representing a time when the same number of inputs into the matching function produce a significantly higher number of new hires. This strong seasonal pattern can also be clearly seen in the raw data for hires and vacancies plotted in figure 3.8, which shows that the number of hires is much more volatile over the year than the number of vacancies. There are two ways to interpret these findings. First, it is possible that there is seasonal variation in the process of matching. This could be due to the nature of employment relationships created over the seasonal cycle, with more temporary jobs filled by young workers being created over the summer, for example. Second, it is possible that there is consistent seasonal mismeasurement of va-

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>( \log h_t )</th>
<th>( \log h_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \log v_{t-1} )</td>
<td>0.668 (0.180)</td>
<td>0.531 (0.158)</td>
</tr>
<tr>
<td>( \log u_{t-1} )</td>
<td>0.378 (0.198)</td>
<td>0.185 (0.177)</td>
</tr>
<tr>
<td>Seasonally adjusted</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Month dummies</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of observations</td>
<td>47</td>
<td>47</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.579</td>
<td>0.958</td>
</tr>
</tbody>
</table>
cancies over the seasonal cycle, due to respondents’ interpretation of job openings referring to openings for permanent employment relationships.

Next, I estimate industry matching functions using the empirical specification

\[
\log h_{it} = \beta_i + \beta_m + \beta_v \log v_{it-1} + \beta_u \log u_{it-1} + \epsilon_{it}
\]

Fig. 3.7  Estimated matching efficiency for different months of the year

Fig. 3.8  Aggregate vacancies and hires between December 2000 and December 2004 (not seasonally adjusted)
where $\beta_i$ are industry and $\beta_m$ are month scale parameters. I measure $u_{t-1}$ by the number of unemployed workers whose last employment was in industry $i$. Even with the limited amount of data available, estimating this specification allows one to decisively reject the hypothesis that the elasticity of the matching function is the same across industries (i.e., $\beta_{v1} = \beta_{v2} = \ldots = \beta_{v18} and \beta_{u1} = \beta_{u2} = \ldots = \beta_{u18}$), and the hypothesis that the matching efficiency is the same across industries (i.e., $\beta_1 = \beta_2 = \ldots = \beta_{18}$), thereby rejecting the hypothesis that the matching function is stable across industries. Again, there are two ways to interpret these findings. First, it is possible that there is variation in the process of matching across industries due to the different characteristics of jobs and workers in these industries. Second, it is possible that the measurement issues that I discussed above systematically affect the measurement of vacancies and hires across industries. In any event, the lack of similarity in the matching function across industries raises the question whether there exists an aggregate matching function at all, as assumed in theoretical studies.

### 3.4 Conclusion

The Job Openings and Labor Turnover Survey (JOLTS) contains important new information that is useful to test existing theories of vacancy creation and to provide new insights into the process of matching in the labor market. In this volume, chapter 2 by Faberman is an excellent introduction to the data available in the JOLTS for anyone wishing to do research using these data. In this chapter, I have focused on several measurement issues that researchers using the JOLTS will have to confront and suggested ways that one might use the JOLTS data to further our understanding of labor market dynamics.

### References


---

4. Since the CPS started using the NAICS industry classification of the JOLTS only after January 2003, this equation is estimated using data from January 2003 to November 2004. Given the small number of observations, it is not possible to separately estimate the month scale parameter for each industry.

5. This implicitly abstracts from industry mobility, which is a strong assumption, but without it, it is not clear what the second input of an industry matching function should be.

6. Estimation results are available upon request.


