Institute for Employment Research

The Research Institute of the Federal Employment Agency

IAB-Discussion Paper 3/2012

Articles on labour market issues

The time trend in the matching function

Friedrich Poeschel

The time trend in the matching function

Friedrich Poeschel (IAB)

Contents

Abstract

We revisit the puzzling finding that labour market performance appears to deteriorate, as suggested by negative time trends in empirical matching functions. We investigate whether these trends simply arise from omitted variable bias. Concretely, we consider the omission of job seekers beyond the unemployed, the omission of inflows as opposed to stocks, and the failure to account for vacancy dynamics. We first build a model of all labour market flows and use it to construct series for these flows from aggregate data on the U.S. labour market. Using these series, we obtain a measure for employed and non-participating job seekers. When we thus include all job seekers, the estimated time trend remains unchanged. We similarly obtain measures for inflows into unemployment and vacancies. When these are included, the magnitude of the time trend is halved but remains significant. When we account for basic vacancy dynamics, the estimated time trend can be fully explained by omitted variable bias. As suggested by this result, we present evidence that empirical matching functions can be interpreted as versions of the law of motion for vacancies: the coefficients in matching functions coincide with the coefficients in the law of motion after correcting for omitted variable bias.

Zusammenfassung

Wir betrachten die scheinbar abnehmende Leistungsfähigkeit des Arbeitsmarktes, angedeutet durch negative Zeittendenzen in empirischen Matchingfunktionen. Wir untersuchen, ob diese Tendenzen schlicht aus Verzerrungen aufgrund ausgelassener Variablen entstehen. Im Einzelnen berücksichtigen wir ausgelassene andere Arbeitssuchende neben den Arbeitslosen, ausgelassene Zugänge neben den Bestandszahlen und ignorierte Dynamik der Vakanzen. Wir erstellen zunächst ein Modell aller Übergänge auf dem Arbeitsmarkt und konstruieren damit Zeitreihen anhand von US-Arbeitsmarktdaten. Mithilfe dieser Reihen ermitteln wir Angaben zur Menge der Arbeit suchenden Beschäftigten und Nichterwerbspersonen. Die geschätzte Zeittendenz einer Matchingfunktion, die diese Arbeitssuchenden miteinbezieht, bleibt unverändert. Wir ermitteln weiter Angaben zu den Zugängen in Arbeitslosigkeit und Vakanzen. Werden diese miteinbezogen, so halbiert sich die Größe der Zeittendenz; sie bleibt aber signifikant. Sobald wir jedoch die Dynamik der Vakanzen berücksichtigen, können wir die geschätzte Zeittendenz vollauf mit Verzerrungen aufgrund ausgelassener Variablen erklären. Wir präsentieren erste Hinweise darauf, dass empirische Matchingfunktionen folglich als Versionen des Bewegungsgesetzes für Vakanzen interpretiert werden können: Die Koeffizienten in den Matchingfunktionen stimmen mit denen im Bewegungsgesetz überein, wenn die Verzerrung durch ausgelassene Variablen korrigiert wird.

JEL classification: J63, J64

Keywords: matching function, time trend, labour market performance, omitted variable

Acknowledgements: I am grateful to Melvyn Coles, Hermann Gartner, Thomas Lubik, Christian Merkl, Fabien Postel-Vinay, Robert Solow, Gesine Stephan, Margaret Stevens, Michael Stops, and participants at the Le Mans SaM workshop (Université du Maine) and at the IAB for helpful and encouraging comments. Klaus Oeckler helped with additional data. Financial support from the Centre Cournot pour la Recherche en Economie is gratefully acknowledged. The paper was partly written at the Ecole Polytechnique, and I thank that institution for its hospitality. All errors are my own.

1 Introduction

The performance of labour markets has typically been described by matching functions as known from the work of Pissarides (see Pissarides (2000), for example). The standard matching function relates the number of matches H in a labour market, i.e. hirings, to the stocks of vacancies V and unemployed job seekers U: $H = m(V,U)$. It is almost always estimated using a Cobb-Douglas specification. In many cases, a time trend is also included to examine how labour market performance has changed over time, so that $H = m(V,U,t)$. Petrongolo/Pissarides (2001) report that a time trend was included in 7 of the 16 studies of the aggregate matching function that they survey. Overall, the empirical results in these studies clearly suggest that there is a highly significant negative time trend, implying that labour market performance appears to deteriorate over time. At the disaggregate level of occupations, Fahr/Sunde (2004) likewise find significant negative time trends in most cases, and not a single instance of a positive time trend; their findings are confirmed by Stops/Mazzoni (2010). Indeed, deteriorating labour market performance might have assumed the status of a stylised fact in labour economics, driven by the nearly ubiquitous finding of negative time trends and closely related findings on shifting Beveridge curves (see e.g. Jackman/Layard/Pissarides (1989)).

The literature has not reached a consensus on why labour market performance appears to be steadily deteriorating. Probably, institutional changes play an important role (see Nickell/Nunziata/Ochel (2005) for an overview). For example, if the unemployment benefit becomes more generous, then unemployed job seekers can afford to reject job offers with comparatively low wages that they would otherwise have accepted. This way, a rise in the unemployment benefit leads to fewer matches than before at any given level of vacancies and unemployed job seekers. Empirically, this will register as a decrease in labour market performance. However, estimation procedures can account for institutional changes, and they seem to explain but a part of the negative time trends (see Petrongolo/Pissarides (2001)). At the same time, the forces that improve labour market performance may be harder to account for, although such forces most likely exist. For example, over the last two decades, the increasing use of the internet in job search and recruitment might have accelerated matching at any given level of vacancies and unemployed (see Kuhn/Mansour (2011)). Therefore, the common finding of negative time trends has remained puzzling.

In this paper, we examine whether the negative time trend arises as a merely statistical product: the estimate of the time trend may be biased downwards, so that a seemingly negative time trend appears when the actual time trend is zero or even positive. Such a negative time trend then would not correspond to any actual changes in labour market performance. We hypothesise specifically that estimates for the time trend in matching functions suffer from omitted variable bias. To our knowledge, this potential explanation for the negative time trend has not been advanced before. We consider two variables that are omitted by the standard matching function $H = m(V,U,t)$ although they might be relevant in the true matching process between unemployed and vacancies: other job seekers (employed or non-participating) and current inflows into vacancies and unemployment. We thirdly investigate the consequences when basic vacancy dynamics are ignored.

For all three cases, we explore theoretically and empirically how the omission affects the estimated time trend. To this end, we build a comprehensive model of labour market flows, which allows us to construct data series that are otherwise unavailable.

Our findings are directly related to the studies that have established the result of a negative time trend: the omissions we consider are all pervasive among the aggregate matching functions listed in Petrongolo/Pissarides (2001). Omissions from the matching function as such have received some attention. Broersma/van Ours (1999) highlight the error that results when measures of matches and job seekers do not correspond, e.g. precisely because employed job seekers are omitted. Mumford/Smith (1999) argue that matches involving unemployed job seekers should not be analysed in isolation from matches involving employed job seekers. Sunde (2007) analyses the problems that arise when only registered vacancies are observed alongside unemployed job seekers. However, none of these papers explores the consequences for the time trend. Rather, where they at all propose any solutions to the problem of omitted variables, they rely on unique data sets and can therefore hardly be replicated.[1]

Closest to this paper is work by Gregg/Petrongolo (2005). Their analysis includes measures of inflows into vacancies and unemployment (while also making the other two omissions). Estimating unemployment and vacancy outflow equations, they do not find a negative time trend in the vacancy outflow equation. They then speculate that an increase over time of the vacancy stock relative to the vacancy inflow might be linked to the negative time trend in a standard matching function. Even more interestingly for us, the magnitude of the significant negative time trend they report for the unemployment outflow equation roughly halves as they move from the standard model of random matching to a model with inflows. Where we analyse inflows as a potential source of omitted variable bias in a matching function, we present a similar finding. We can also conclude that the omission of other job seekers does not seem to generate any part of the time trend, as it remains unchanged when they are included. By contrast, when we investigate how the law of motion for vacancies affects the estimated time trend, we are able to explain its entire magnitude as a consequence of omitted variable bias. In fact, our results raise the possibility that empirical matching functions reflect merely the law of motion for vacancies, a link that appears to be unknown in the literature.

Below we proceed as follows. In section 2, we describe our data, find a negative time trend using the common approach, and offer a critique of previously advanced explanations for this finding. In section 3, we propose an empirical model of labour market flows. We then use series constructed from the model to examine whether the time trend arises from the omission of job seekers beyond the unemployed. Section 4 similarly considers the omission of inflows into unemployment and vacancies as a potential source of omitted variable bias. Finally, section 5 explores the consequences for the estimated time trend and matching functions more generally when the law of motion for vacancies is ignored. Section 6 collects our conclusions.

[1] For example, to implement Sunde's (2007) approach, one needs to observe whether the job seeker in a match was previously unemployed or employed and whether the vacancy was registered or not.

2 The puzzle: negative empirical time trends

2.1 Data

We use two data sets. The first is the *Job Openings and Labor Turnover Survey* (JOLTS) from the U.S. Bureau of Labor Statistics.[2] Following the Bureau, JOLTS is a sample of around 16,000 observations collected on a monthly basis from establishments across the United States. These establishments may be firms in the private non-agricultural sectors or government bodies at the local, State, and Federal levels. Data collection started in December 2000 and continues to date. We will use the observations from January 2001 to February 2011, giving us $T = 122$ monthly observations. Here lie the first advantages of this data set for our purposes: we are not aware of any major institutional changes over this period that would likely reduce the efficiency of matching, thereby leading to a negative time trend. Rather, we can imagine that an ongoing shift towards internet-based recruitment and job search over this period increased labour market efficiency. Another advantage is the monthly frequency of the data, which should allow us to largely avoid the issue of aggregation bias and any consequences for the time trend.

Establishments participating in JOLTS report information on their hirings, vacancies, and separations. Following the definitions of the Bureau of Labor Statistics, hirings in period t (where the period is a month) are defined as the number of workers added to the firm's payroll in period t. This includes seasonal workers and rehired staff after a layoff at least a week earlier, but does not include staff from temporary help agencies or similar contractors. A vacancy in period t is defined as an unfilled position, part-time or full-time, on the last business day of period t. An unfilled position exists if a specific position is currently not held by anyone but could be taken up within 30 days, and if the firm engages in recruitment efforts to fill the position. This does not include positions that must be filled internally, nor positions for which somebody has been hired but work has not yet begun. Here lies a further advantage of the JOLTS data: these definitions, the use of payroll changes where possible, and the census-like way of data collection ensure that measurements are good, even exceptionally good in the case of vacancies. Therefore, in contrast to many other data sets, measurement error is unlikely to be a major problem in JOLTS.

The second data set we use is the U.S. Current Population Survey (CPS), conducted by the Census Bureau and beginning in 1940. The information in the CPS comes from a monthly representative survey of around 60,000 U.S. households (thus around 110,000 individuals). Based on responses about activities of the household members during the reference week, which is normally the week including the 12th of the month, the individuals' labour market status according to CPS definitions is inferred. A person from age 16 who holds a job is classified as employed, be it full-time, part-time, or temporary work. A person from age 16 is classified as unemployed if the person does not currently hold a job but would be available for work and, unless on temporary lay-off, has actively sought work for at least four weeks. This definition matches the economic definition of unemployment rather well, so that there should not be major measurement error from this source. Finally, a person

[2] The data used for this version of the paper were obtained after the revisions to JOLTS in early 2011.

Table 1: Descriptive statistics

Variable	mean	st. dev.	min.	max.
H_t	4829	870	2705	6243
V_t	3679	809	2038	5691
TS_t	4855	909	3059	8125
E_t	140669	3402	135701	146584
U_t	9262	2894	6023	15628
U_t^{QS}	5998	2228	3628	10866
U_t^{LR}	3265	692	2273	4966
$U_t^{<5}$	2787	231	2297	3522
N_t	77414	3932	70083	86216
H_t/U_t	0.579	0.209	0.178	0.972
V_t/U_t	0.449	0.187	0.134	0.945

Figures are monthly levels given in thousands, except H_t/U_t and V_t/U_t. Sources: CPS and JOLTS, 01/2001 to 02/2011.

from age 16 neither classified as employed nor classified as unemployed is deemed not to be participating in the labour force. The sum total of individuals thus classified corresponds to the non-institutional civilian population from the age of 16, as people in institutions such as prisons and in the army are disregarded.

The CPS underwent a major re-design in 1994, but all the series we will use begin later and have been collected under the same standards. It is worth noting that the definitions used for these classifications appear to accord with the definitions used for the JOLTS data.[3] For example, workers on temporary lay-off are counted as unemployed in the CPS, and the recall of a temporarily laid-off worker is counted as a hiring in JOLTS. For both the JOLTS and the CPS data, table 1 offers descriptive statistics for the variables and derived variables that we will employ, over the period from January 2001 to February 2011. From JOLTS, we use the series on hires (H_t), vacancies (V_t), and total separations (TS_t). E_t, U_t, and N_t denote CPS series on the stocks of workers who are respectively employed, unemployed, and not participating in month t. We obtain the series U_t^{QS} by adding CPS stocks of currently unemployed job leavers and job losers, and U_t^{LR} by adding currently unemployed labour force entrants and re-entrants. While we use several of the CPS series on the unemployment stock by unemployment duration, we report here only the most relevant, the stock $U_t^{<5}$ of currently unemployed individuals who have been in this state for less than 5 weeks.

2.2 Some observations from a standard analysis

The most widely used matching function relates the flow of hirings H_t to stocks of vacancies V_t and unemployed job seekers U_t:

$$H_t = m(V_t, U_t)$$

[3] For further details and comparisons, data definitions for both JOLTS and CPS data can be obtained from http://www.bls.gov/bls/glossary.htm.

Table 2: Matching function regressions using OLS ($T = 122$)

dependent variable	v_t	u_t	θ_t	t	const.	seas. dum.	adj. R^2
(1) h_t	0.845*	0.302*	–	-0.0018*	-1.087	no	0.64
s.e.	0.106	0.106	–	0.0004	1.764		
(2) h_t	0.298*	-0.172*	–	-0.0008*	7.384*	yes	0.95
s.e.	0.059	0.055	–	0.0002	0.949		
(3) f_t	–	–	0.772*	-0.0016*	0.172*	no	0.93
s.e.	–	–	0.026	0.0004	0.023		
(4) f_t	–	–	0.750*	-0.0017*	-0.039	yes	0.98
s.e.	–	–	0.013	0.0002	0.021		

* Significant at 1% level.

where $m(\cdot)$ is the matching function. We focus on the canonical Cobb-Douglas specification of $m(\cdot)$. Then a matching function with a time trend may be written as

$$H_t = V_t^{\beta} U_t^{\alpha} e^{K+\gamma t+\epsilon_t} \tag{1}$$

where β is the elasticity of matches with respect to the stock of vacancies, α is the elasticity of matches with respect to job seekers, K is a constant, γ is the coefficient of time t, and ϵ_t is random error. When the constant is interpreted as the speed of matching (see e.g. Lindeboom/van Ours/Renes (1994)), the time trend γt will indicate how the performance of the labour market has changed over time. In particular, a negative value for γ implies that over time fewer hirings result from given stocks of vacancies and unemployed, so that labour market performance deteriorates.

In logarithms, equation (1) returns a linear model equation:

$$h_t = K + \beta v_t + \alpha u_t + \gamma t + \epsilon_t \tag{2}$$

where the lower-case letters h_t, v_t, and u_t denote logarithms. Given time series on H_t, V_t, and U_t, equation (2) can be estimated. As this equation is linear, OLS will give the best results provided the assumptions of the Gauss-Markov theorem hold. Line (1) in table 2 reports results for such a regression. Because a strong element of seasonality may be expected in the data, we can also extend the model in equation (2) by a full set of monthly dummies to account for seasonality, with December as reference category:

$$h_t = K + \beta v_t + \alpha u_t + \gamma t + \sum_{i=month} \delta^i D^i + \epsilon_t \tag{3}$$

The results are reported in line (2) of table 2. While the results in line (1) suggest plausible values for the elasticities β and α, the negative result for α in line (2) does not make much sense: it would mean that a rise in job seekers leads to a fall in hirings.

The results in line (1) support the hypothesis of constant returns to scale (CRS) in matching: doubling the inputs V_t and U_t in equation (1) will then ceteris paribus also double H_t if $\beta + \alpha = 1$. The F-test statistic for this hypothesis is 0.51 from $F_{(1,118)}$, while the critical value from $F_{1,100}$ is 3.94 at the 5% significance level, so that we fail to reject this hypothesis. Under the restriction of CRS, we can rewrite equation (1) as

$$\frac{H_t}{U_t} = \left(\frac{V_t}{U_t}\right)^{\beta} e^{K+\gamma t+\epsilon_t} \tag{4}$$

and then linearise this to obtain

$$f_t = K + \beta\theta_t + \gamma t + \epsilon_t \tag{5}$$

where $f_t = \ln(H_t/U_t)$ is the logarithm of the *job-finding rate* of the unemployed and $\theta_t = \ln(V_t/U_t)$ is the logarithm of *market tightness* V_t/U_t: the more vacancies there are relative to job seekers, the tighter is the labour market. Line (3) in table 2 gives the results for a regression based on this equation. They plausibly suggest an elasticity $\beta = 0.772$ for vacancies and $1 - \beta = 0.228$ for unemployed job seekers. This time, extending the model in equation (5) by monthly dummies does not lead to implausible results, as line (4) shows. Finally, and importantly for this paper, all regressions that produced plausible estimates also indicate a highly significant negative time trend of about equal magnitude.[4]

It is well known that the assumptions of the Gauss-Markov theorem might not apply to the model in equation (1), so that OLS should not be used. A lot of attention has been devoted to the issue that hirings are measured as the total flow in a period but then regressed on stocks that change during this period, not least as an immediate consequence of hirings (see Petrongolo/Pissarides (2001) for an overview). Another issue is endogeneity bias that might arise because firms choose to open more vacancies when labour market performance is particularly high, in the sense that ϵ_t is particularly large. Such endogenous firm behaviour would generate a correlation between ϵ_t and the explanatory variable V_t (see Borowczyk-Martins/Jolivet/Postel-Vinay (2011)).

In order to avoid such problems, one might use an instrumental variable (IV) approach. To this end, we employ a generalised method of moments estimator with a robust weight matrix. The regression in line (5) of table 3 estimates a model like in equation (2) but instruments v_t and u_t respectively by v_{t-1} and u_{t-3}; the correlation coefficients are respectively $r_{v_t,v_{t-1}} = 0.85$ and $r_{u_t,u_{t-3}} = 0.98$. If we used u_{t-1} as instrument instead, the correlation with the instrumented u_t would be almost perfect. The results are poor: the estimated elasticities in line (5) appear far too high. When monthly dummies are included in line (6), the estimated coefficient of u_t is close to 0 and insignificant. The regression in line (7) instruments only u_t by u_{t-3} and produces results close to those in line (1) of table 2. When only v_t is instrumented, the results are equally implausible as in line (5). This could mean that u_{t-3} is a good instrument while v_{t-1} is not; but since the results in line (7) are close to

[4] We have also run the regression in line (3) as a rolling-window regression with 60 months covered by the window. Monthly dummies are disregarded here due to the reduced number of observations for each regression. The point estimate for the time trend is negative in almost all instances. If taken at face value, this would indicate that labour market performance is almost always declining, not just its long-run average.

Table 3: Matching function regressions using IV ($T = 121$)

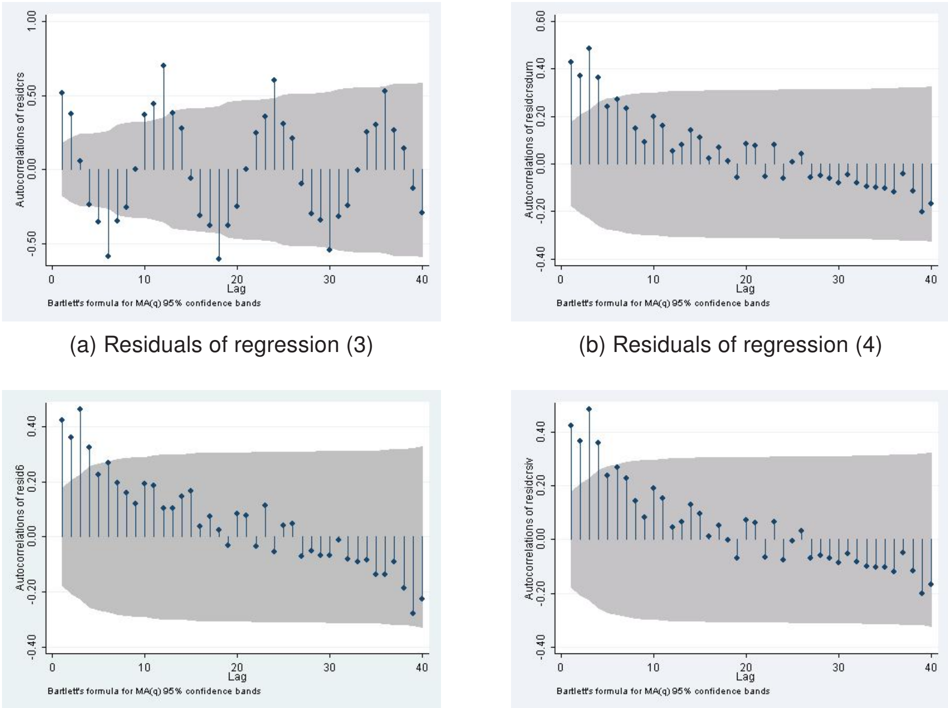
dependent variable	v_t	u_t	θ_t	t	const.	seas. dum.	R^2
(5) h_t	1.704*	1.034**	–	-0.0035*	-14.667	no	0.50
s.e.	0.574	0.506	–	0.0011	9.211		
(6) h_t	0.494*	0.017	–	-0.0013*	4.115	yes	0.94
s.e.	0.136	0.122	–	0.0003	2.183		
(7) h_t	0.907*	0.353*	–	-0.0021*	-2.033	no	0.67
s.e.	0.121	0.126	–	0.0004	2.072		
(8) h_t	0.481*	0.006	–	-0.0012*	4.315*	yes	0.94
s.e.	0.069	0.067	–	0.0002	1.145		
(9) f_t	–	–	0.777*	-0.0016*	0.179*	no	0.93
s.e.	–	–	0.028	0.0003	0.025		
(10) f_t	–	–	0.754*	-0.0017*	-0.034	yes	0.99
s.e.	–	–	0.010	0.0002	0.025		

GMM with a robust weight matrix was used in all cases for estimation.
* Significant at 1% level.
** Significant at 5% level.

those in line (1), it also appears that a good instrument only for u_t either does not help, or that there was no problem in regression (1) to begin with. In any case, also the results in line (7) give way to an insignificant estimated coefficient of u_t when monthly dummies are included in line (8).

Line (9) in table 3 estimates the specification with CRS in equation (5) but instruments θ_t by θ_{t-1} ($r_{\theta_t,\theta_{t-1}} = 0.96$), and line (10) estimates this specification with monthly dummies. In both cases, the results are remarkably close to the results in lines (3) and (4) of table 2 based on OLS. The results also seem to change little with the inclusion of monthly dummies. The time trends are without exception negative and significant at the 1% significance level, with almost identical magnitudes. It thus appears that the CRS specification delivers plausible and particularly robust results across the estimations considered here. This paper will therefore often focus on this specification.

In turn, the results for the matching function specifications that do not assume CRS are not satisfactory: while line (7) gives the results closest to the estimates that one would expect, this regression neither includes monthly dummies nor instruments for v_t. Hence it is likely to suffer from the endogeneity bias mentioned above because the vacancy data in JOLTS counts job openings on the *last* business day of the month, so that v_t is not pre-determined. This failure of the standard matching function to materialise in the JOLTS data should be in and of itself suspicious. The results for the CRS specification do not counterbalance this failure, as they might reflect as little as a positive correlation between the job finding rate and market tightness, rather than a structural relationship between hirings, jobseekers, and



(a) Residuals of regression (3) (b) Residuals of regression (4)

(c) Residuals of regression (6) (d) Residuals of regression (10)

Figure 1: Autocorrelation of the regression residuals

vacancies. Whatever problem affects the standard matching function here, it might also be the source of the negative time trend.

In particular, the residuals in all regressions considered above indicate that misspecification may be an issue. Panel (a) in figure 1 calls for the inclusion of monthly dummies to account for seasonality. Panel (b) shows that strong autocorrelation remains also with dummies included. While these two panels depict the residuals for the CRS specification, the residuals for the standard matching function behave very similarly. Such autocorrelation may reflect an omitted variable that affects hirings similarly over a number of periods. With IV regression, the residuals of the standard and CRS specifications are respectively depicted by panels (c) and (d). They suggest that the instrumentation did not fix the problem.

2.3 An assessment of previous hypotheses

It has often been suggested that a growing share of long-term unemployed is responsible for the negative time trend. Underlying this hypothesis is the idea that long-term unemployed may be less employable than short-term unemployed, due to the loss of skills over time or due to stigmatisation, or that long-term unemployed become discouraged and search less intensively than short-term unemployed. In any case, counting long-term unemployed towards job seekers with the same weight as short-term unemployed would overstate the measure of job seekers. If, in addition, the share of long-term unemployed grows over time, then the measure of job seekers is increasingly overstated over time. When there is no corresponding tendency in hirings, this generates a negative time trend in the standard matching function.

To account for the composition of unemployment, Blanchard/Diamond (1990) let only a proportion χ of long-term unemployed U_t^{LT} be perfect substitutes for short-term unemployed U_t^{ST}. The linearised standard matching function then becomes

$$h_t = K + \beta v_t + \alpha \ln(U_t^{ST} + \chi U_t^{LT}) + \gamma t + \epsilon_t$$

As it is not clear how χ should be determined, we take a radical approach here and set $\chi = 0$. We thus assume that an unemployed job seeker is a perfect substitute for a newly-unemployed up to a certain unemployment duration, and not a substitute at all from then on. We then consider three different cut-off durations: 5 weeks, 15 weeks, and 27 weeks. The relevant measures of unemployed job seekers are easily obtained from the CPS data on stocks of unemployed with duration of less than 5 weeks, 5 to 14 weeks, 15 weeks or longer, and 27 weeks or longer.[5]

Even this radical approach, however, does not produce conclusive evidence. When the cut-offs 5 weeks, 15 weeks, and 27 weeks are used in OLS regressions without monthly dummies, the time trends are respectively -0.0009, -0.0011, and -0.0013 (all significant at the 2% significance level). First, it is worth noting that there is a significant negative time trend even when only newly unemployed are considered. The magnitude of this time trend does exhibit a tendency to increase as more long-term unemployed are counted in. Yet this result is not robust: as soon as monthly dummies are included, the magnitude of the time trend exhibits the reverse tendency, while the estimates of α are negative and signifcant, much like in line (2) of table 2. Nor do we find a clear pattern when we instrument vacancies and the respective unemployment measure by lagged values (be it one or three lags in the case of unemployment). Hence it appears that a growing share of long-term unemployed might at best explain a part of the time trend. This conclusion is in line with Blanchard/Diamond (1990) who find no evidence that short-term and long-term unemployed should at all be treated differently in the matching function.

The efficiency of matching might also seemingly deteriorate when, for example, unemployment benefits increase and unemployed job seekers consequently raise their reservation wages, rejecting a greater share of the job offers they receive. A tendency for workers to become more picky over time would then translate into a negative time trend. To see this, consider the matching function

$$m(z, V, U) = p(z)C(V, U) \tag{6}$$

where z is a reservation productivity level, $p(z)$ is the probability that the match productivity exceeds z, and $C(\cdot)$ is the 'contact function' capturing a meeting technology. $C(\cdot)$ could be any matching function specification that does not distinguish between contacts and actual matches, including the specification in equation (1). Yet any of these contacts only leads to a match with probability $p(z)$. If z increases over time, reflecting that agents become pickier, then the same number of contacts in the labour market will lead to fewer matches. Where a matching function does not distinguish between contacts and matches,

[5] With superscripts indicating the unemployment duration in weeks, we have $U_t^{<15} = U_t^{<5} + U_t^{5-14}$ and $U_t^{<27} = U_t^{<15} + U_t^{\geq 15} - U_t^{\geq 27}$.

this growing divergence can be picked up by a negative time trend.

As data on $p(z)$ cannot be obtained, it is not easy to assess any empirical link between selectivity and the time trend. Attempts to account for the generosity of unemployment insurance, where quantifiable, have failed to explain the time trend (see Petrongolo/Pissarides (2001)). An evaluation of such exogenous changes to institutions is beyond the scope of this paper. However, changes in selectivity can also arise endogenously from changes in labour market conditions: when unemployed workers understand that they are more sought after, they might become more picky. It is straightforward to argue theoretically, as we do in appendix A using the model in Stevens (2007), that workers' selectivity z is likely to depend positively on market tightness. As a first empirical test for such endogenous selectivity, we can rank our $T = 122$ observations by the level of V_t/U_t (in descending order). With an intercept dummy D_t^L equal to 1 if V_t/U_t falls into the bottom half of the ranking and otherwise $D_t^L = 0$, we repeat the regression in line (4) of table 2. If endogenous selectivity is important, matching should be quicker when V_t/U_t is low, so that the coefficient of D_t^L is positive and significant. However, we find an insignificant coefficient estimate of almost exactly 0, and no change to the other estimates.[6]

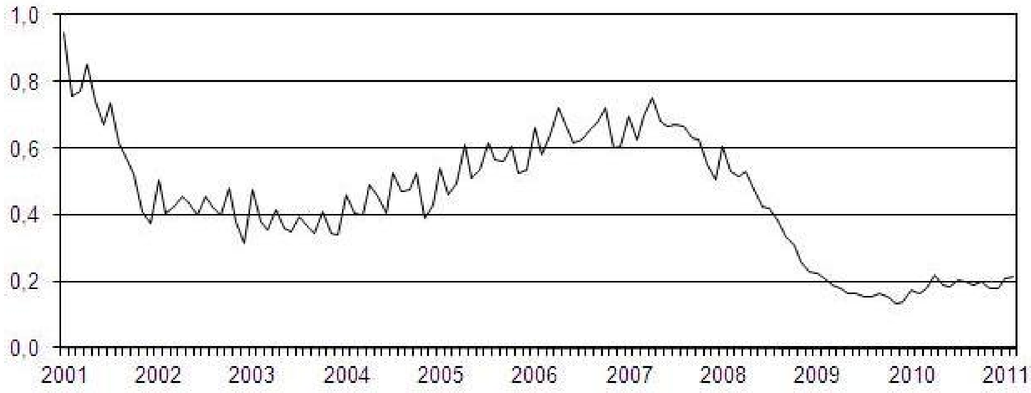
Worse, even if there was a positive link between selectivity and market tightness, this would imply a *positive* time trend for the time period under consideration: as panel (a) in figure 2 depicts, market tightness has tended to fall between January 2001 and February 2011. The correlation between θ_t and time is accordingly -0.60. We therefore conclude that the negative time trend is likely not generated by an increase of workers' pickiness over time.

In an oft-cited contribution, Bleakley/Fuhrer (1997) investigate the inward shift of the Beveridge curve for the U.S. in the late 1980s. Such shifts may capture the same changes as time trends in matching functions do. Indeed, among three potential explanations for the observed shift, Bleakley/Fuhrer (1997) first advance changes in matching efficiency. In this rare instance, their estimation of a standard Cobb-Douglas matching function produces a positive time trend, corresponding to the observed inward shift. While they conclude that shifts in the parameters K and α can equally well account for the time trend, they only speculate about the economic reasons behind this trend.

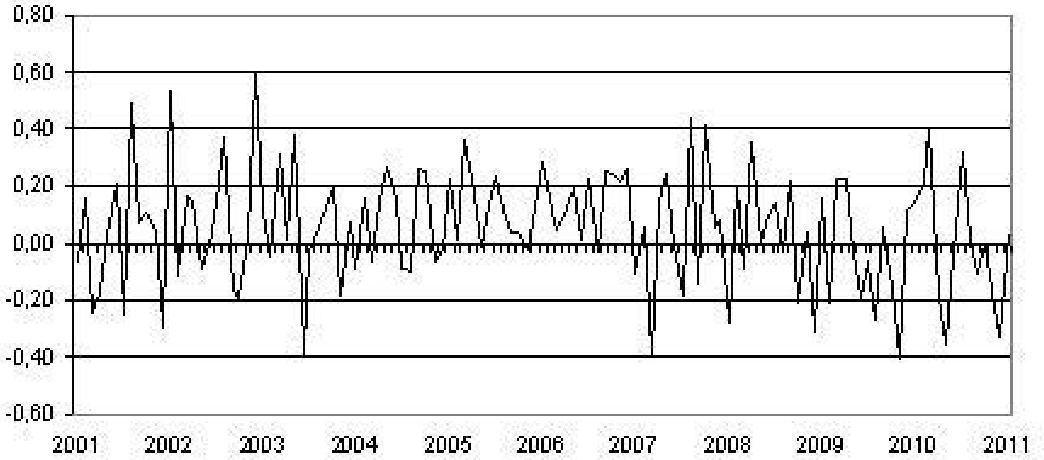
The second explanation they consider concerns changes in labour force growth: a permanently lower inflow of labour market entrants into unemployment would reduce unemployment levels at given vacancy levels and thus shift the Beveridge curve inwards. As Bleakley/Fuhrer (1997) offer a primarily graphical argument, we evaluate it using exactly corresponding graphs. Panel (b) in figure 2 suggests that this argument cannot explain the negative time trend we found above for the period January 2001 to February 2011. The slight tendency for labour force growth to fall over this period would, following this logic, again lead to a positive time trend.

Finally, Bleakley/Fuhrer (1997) point to a decrease in labour market "churning", measured

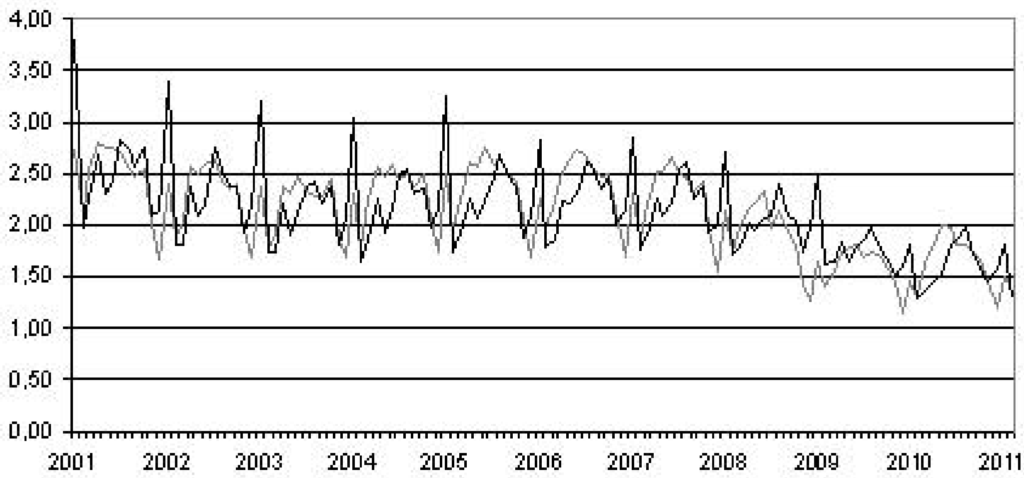
[6] Sedláček (2010) jointly estimates vacancies and the speed of matching. He finds a procyclical speed of matching on the U.S. labour market (albeit for an earlier time period than considered here). To confirm this finding, our crude empirical test would have to produce a negative and significant coefficient estimate for D_t^L because market tightness is procyclical.



(a) Ratio of vacancies (from JOLTS) to unemployed (from the CPS)



(b) Percentage change of the labour force, based on CPS data on employment and unemployment stocks



(c) Hirings (grey line) and separations (black line) in percent of the working-age population, based on flow data from JOLTS and a stock from the CPS

Figure 2: Graphical evaluation of previous explanations for negative time trends

as the magnitude of flows into and out of employment. If the magnitude of both flows decreases, there will be simultaneously lower levels of vacancies and unemployment, amounting to an inward shift of the Beveridge curve. Panel (c) in figure 2 shows that "churning" also decreased over the period considered here. For this to explain the negative time trend, the matching function would have to exhibit increasing returns to scale, so that a simultaneous fall in U_t and V_t disproportionately reduces H_t. The fragility of the standard matching function in our regressions above makes it hard to say much about the returns to scale. In fact, section 3.4 below finds evidence that, after an appropriate modification, an IV regression suggests constant returns to scale.

In conclusion, previous attempts to explain the phenomenon of negative time trends have had limited success. In this paper, we theoretically and empirically examine three potential explanations that, to the best of our knowledge, have not been investigated before. We find that negative time trends can thus be fully explained.

3 Bias from the omission of other job seekers

3.1 Theory

While most studies take the stock of unemployed workers to be the stock of job seekers, we also want to account for job seekers who currently hold a job or do not count towards the labour force. Because vacant jobs may be taken up just as well by such workers, the matching function will only be specified correctly if the stocks of employed and non-participating workers also enter. Several papers, for example Broersma/van Ours (1999) and Sunde (2007), have found that the estimated elasticities and returns to scale will be biased if only the stock of unemployed workers is used. However, the consequences for the estimated time trend have apparently not received any attention. To fill this gap, this section will investigate to what extent a negative time trend may result from omitted variable bias when employed and non-participating job seekers are omitted from the matching function.

We consider a more general matching function that relates the flow of hirings H_t to stocks of job seekers J_t and vacancies V_t:

$$H_t = m(J_t, V_t) = m(U_t + \phi_t(E_t + N_t), V_t)$$

The parameter ϕ_t denotes the average search intensity of employed and non-participating workers, relative to unemployed workers whose search intensity is normalised to 1. Assuming that $m(\cdot)$ again takes a Cobb-Douglas functional form, we have

$$H_t = V_t^{\beta} J_t^{\alpha} e^{K+\gamma t+\epsilon_t} = V_t^{\beta} \left[U_t + \phi_t(E_t + N_t)\right]^{\alpha} e^{K+\gamma t+\epsilon_t} \tag{7}$$

Since most studies in the literature either assume CRS or find evidence of CRS, let us also consider the special case $\beta + \alpha = 1$. Then dividing by U_t gives

$$\frac{H_t}{U_t} = \left[\frac{V_t}{U_t}\right]^{\beta} \left[\frac{U_t + \phi_t(E_t + N_t)}{U_t}\right]^{1-\beta} e^{K+\gamma t+\epsilon_t}$$

After linearising, this is

$$f_t = K + \gamma t + \beta\theta_t + (1-\beta)x_t + \epsilon_t \tag{8}$$

where

$$x_t = \ln \frac{U_t + \phi_t(E_t + N_t)}{U_t}$$

However, suppose a matching function is estimated that only includes unemployed workers and thereby omits x_t :

$$f_t = K' + \gamma' t + \beta'\theta_t + \epsilon_t \tag{9}$$

The textbook treatment of omitted variable bias in the context of multivariate regression leads to a simple expression for the expectation of the estimated coefficient $\hat{\gamma}'$:

$$\mathbb{E}(\hat{\gamma}') = \gamma + (1-\beta)\tau \tag{10}$$

where τ is the coefficient of t in an OLS regression of x_t on the other explanatory variables.

3.2 Measurement

In order to avoid this bias, stocks of employed and non-participating job seekers should also be included in the matching function. Since only some of these workers seek jobs, these stocks have to be weighted by ϕ_t, which requires information on flows between labour market states, in particular the various flows that generate matches. In principle, the CPS gross flow data provide such information for the U.S. labour market. These flows are obtained by comparing responses of individuals surveyed repeatedly in the CPS: if they report a different labour market status than the last time they were surveyed, this will count as a transition, and the sum of such transitions gives the respective flow. However, until recently the BLS did not even publish these flows because they were inconsistent with changes in the stocks.

As the gross flow data derive only from changes in labour market state, they cannot offer any series for the important flow of job-to-job transitions. In addition, two serious issues with the available flows have been discovered. Firstly, when individuals leave or enter the CPS sample, a comparison of their responses cannot be made, and these observations on transitions do not seem to be missing at random. Secondly, an individual's labour market status may be recorded incorrectly for some reason. While such classification errors might well cancel out in the data on stocks, they accumulate in the data on flows. For example, a constantly unemployed individual may be classified as employed once, generating a spurious transition, and is likely later classified correctly again, which generates a second spurious transition.[7]

Abowd/Zellner (1985) and Poterba/Summers (1986) find that the CPS gross flow data therefore very substantially overestimate the actual labour market flows, and they propose

[7] These problems apparently extend to flow data based on similar surveys. For example, the Australian data used by Mumford/Smith (1999) suffer from much the same problems.

procedures for adjusting the data but admit that these procedures may themselves suffer from various problems. In any case, matching function regressions that rely on unadjusted CPS gross flow data should be treated very cautiously, including Bleakley/Fuhrer (1997) and studies as recent as Jolivet (2009). Nagypál (2008) points out that, due to the re-design of the CPS in 1994, the adjustment procedures cannot be used for data after 1994. She even finds that the incidence of classification error has grown after the re-design.[8]

Since the re-design the CPS has also asked employed individuals whether they still work for the same employer as before. Fallick/Fleischman (2004) use this information to construct a series for job-to-job transitions. From comparisons with other data sets, Nagypál (2008) finds that this method might overstate job-to-job transitions by as much as 30%. Moreover, the issue remains that missing observations are not missing at random. Finally, even if these defaults were corrected, this will not allow us to construct series for the other labour market flows of interest. Unfortunately, weights for employed and non-participating job seekers in the matching function cannot be determined from job-to-job transitions alone.[9]

An alternative approach is to develop models that, given some more reliable data such as CPS stocks, allow to construct flow data. Shimer (2005) builds a job ladder model that allows the construction of some flows (within employment and between unemployment and employment) from CPS data on stocks, but without using the CPS gross flow data. While the model is set in continuous time and thereby avoids time aggregation issues, its main drawback is the limitation to employment and unemployment as the only labour market states, which precludes in particular that matches may arise from non-participating job seekers. Adding non-participation would greatly increase this model's complexity and probably render it intractable.[10]

The approach we take in the next section is to build a simple but comprehensive accounting model of the labour market that treats all flows as unknowns. From the model's solutions, series for the unknown flows can be constructed using only the CPS data on stocks and some flow data from JOLTS. The JOLTS flow data are also collected by the CPS interviewers, but derive from changes in firms' payrolls and are thus likely free of the problems plaguing the CPS gross flow data: misclassification is much less likely to happen because firms have strong incentives to maintain correct payrolls. At the same time, the JOLTS data are equally regularly updated and equally easily available as CPS data. Therefore, such an accounting model has the potential to give both a more comprehensive and a more accurate view of the labour market than a standard job ladder model. Indeed, Nagypál's (2005) results strongly suggest that standard job ladder models cannot explain the observed job-to-job transitions. We thus hope that our alternative approach stimulates further effort in this direction.

[8] Fujita/Ramey (2006) propose a correction for the problem of missing observations, extending the method of Abowd/Zellner (1985). However, they apparently do not deal with the problem of classification error.

[9] Shimer (2005) discusses a method that uses information from the March supplement to the CPS on jobs held over the previous year. It appears that this method involves still more problems.

[10] In Shimer (2007), a model that includes all three labour market states is built and solved numerically, but the model requires data on flows to construct instantaneous transition rates. That is, this model does not produce any series for the flows, it rather takes the CPS gross flow data as an input without any adjustment, despite the serious issues in these data. Similarly, a very limited model in Nagypál (2008) takes adjusted CPS gross flows as inputs to obtain estimates for two transition rates.

3.3 An empirical model of labour market flows

To explicitly account for job-to-job transitions and other flows, we assemble a comprehensive model of the empirical flows on the labour market. The model exists of accounting equations that link CPS data on stocks to JOLTS data on flows. In principle, this suffers from a time aggregation problem - if an agent's spell between two movements on the labour market is short enough, it might not have been recorded in the monthly CPS data. However, this issue has typically been discussed in the context of quarterly data, which are much more affected by this problem. As we would view, for example, sufficiently short spells in unemployment between two jobs more as an actual job-to-job transition, not too much seems lost when many spells shorter than a month are not counted. In any case, Burdett/Coles/van Ours (1994) find that effects from time aggregation are the lower the higher the frequency of the data. Their findings suggest that such bias is unlikely to be a problem for matching function regressions based on monthly data. Another potential problem could be that the CPS data do not count very short spells while the JOLTS series on flows do; we return to this issue below.

Each worker in the model over the age of 16 is always in one of three labour-market states: employment (E), unemployment (U), and non-participation (N). Let EU_t denote the flow of workers in period t from employment to unemployment, and similarly for all other flows between states. Recall that E_t, U_t, and N_t denote the CPS stocks of workers in each state in period t, and that we can distinguish within unemployment the stock U_t^{QS} of workers who are unemployed following quits or forced separations and the stock U_t^{LR} of unemployed labour-force entrants or re-entrants. As before, the length of period t is a month, in accordance with our monthly data. The JOLTS data in turn provide us with series on hires (H_t) and on total separations (TS_t).

Those who are already in the labour force at age 16 or are still in the labour force beyond the age of 65 make up a small share of the total labour force at any point in time. The vast majority of workers in the labour force have come from state N after the age of 15 and return to that state before the age of 66. Hence, we can pretend that all flows into the working-age labour force originate in N, and all flows out of the working-age labour force lead to state N. Therefore, the growth of the working-age population directly only affects N and is then passed on to other states through the flows that originate in N. In principle, all unemployed must then have come from either employment or non-participation. However, U_t diverges slightly from the sum $U_t^{QS} + U_t^{LR}$ (by at most 2.2%). Therefore, we adjust U_t^{QS} and U_t^{LR} such that their sum equals U_t, while the ratio U_t^{QS}/U_t^{LR} is preserved by the adjusted series U_t^{QS+} and U_t^{LR+}:

$$U_t^{QS+} = U_t^{QS}\left[1+\frac{U_t - U_t^{QS} - U_t^{LR}}{U_t^{QS}+U_t^{LR}}\right], \quad U_t^{LR+} = U_t^{LR}\left[1+\frac{U_t - U_t^{QS} - U_t^{LR}}{U_t^{QS}+U_t^{LR}}\right]$$

We also make the simplifying assumption that unemployed workers in U_t^{QS+} and U_t^{LR+} move into employment during period t at the same average rate UE_t/U_{t-1} (the job-finding rate), and also move into non-participation at the same average rate UN_t/U_{t-1}. We can then write down the following stock-flow equations:

$$\Delta U_t^{QS+} = EU_t - \frac{UE_t + UN_t}{U_{t-1}} U_{t-1}^{QS+} \tag{11}$$

$$\Delta U_t^{LR+} = NU_t - \frac{UE_t + UN_t}{U_{t-1}} U_{t-1}^{LR+} \tag{12}$$

Equation (11) says that the change $U_t^{QS+} - U_{t-1}^{QS+}$ equals the inflow of workers from quits or lay-offs during period t, EU_t, minus the outflow from the existing stock U_{t-1}^{QS+} at rate $(UE_t + UN_t)/U_{t-1}$. Equation (12) says the same for unemployed workers who were previously not participating.

We next link the JOLTS series to the flows in our model. As hirings include workers from any of the three labour market states, we know

$$H_t = EE_t + UE_t + NE_t \tag{13}$$

Likewise, following a separation, workers can move to any labour market state:[11]

$$TS_t = EE_t + EU_t + EN_t \tag{14}$$

In our data, the difference $H_t - TS_t$ does not level with changes in E_t. This discrepancy cannot be due to time aggregation issues, as very short employment relations would contribute equally to H_t and TS_t and therefore leave $H_t - TS_t$ unaffected. Rather, H_t and TS_t reflect changes in firms' payrolls, while E_t is obtained from surveyed households and therefore also counts self-employed individuals.[12] Since self-employed do not match with firms, they should not feature in a matching model of the labour market. We therefore use $H_t - TS_t$ throughout to measure changes in employment, but we do not write down another stock-flow equation for these changes, as this equation would simply combine equations (13) and (14) and would thus not be independent.

To close the model, one can also write down a stock-flow equation for changes in non-participation. If we neglected for a moment that ΔN_t does not correspond to the flows in our model because N_t does not include the self-employed, we could write

$$\Delta N_t = \Delta P_t + (EN_t + UN_t) - (NE_t + NU_t)$$

where ΔP_t denotes the growth of the working-age population. However, this equation would not add an independent equation either. Since it should then hold that

$$\Delta U_t^{QS+} + \Delta U_t^{LR+} + H_t - TS_t + \Delta N_t = \Delta P_t$$

the equation for ΔN_t would be a combination of equations (11), (12), (13) and (14). Instead, we obtain another independent equation using the CPS series on the stock of work-

[11] Given that changes in the population only affect state N in our model, deaths and emigration (taken into account by TS_t) are thought of as a part of EN_t.

[12] A CPS series for employment without self-employed is apparently not publicly available. Discrepancies between $H_t - TS_t$ and ΔE_t will also arise if they are measured on different days of the month.

ers who have been unemployed for less than five weeks, denoted $U_t^{<5}$. As our model disregards very short unemployment spells for the reasons mentioned above, this stock of newly unemployed represents the inflows into unemployment during the month:

$$U_t^{<5} = EU_t + NU_t \tag{15}$$

Equations (11) through (15) form a system of five equations in the following seven unknowns: EE_t, UE_t, NE_t, EN_t, EU_t, NU_t, and UN_t. To turn this into an exactly identified system, we make an identifying assumption. From the discussion in the last subsection, recall the CPS gross flow data that are plagued by measurement errors. Denote the CPS series for the flow from employment to non-participation by $\mathcal{EN}_t$ and the reverse flow by $\mathcal{NE}_t$. Due to measurement error, they differ by proportions p_t^{EN} and p_t^{NE} from the true flows EN_t and NE_t, respectively. Thus we have

$$\frac{\mathcal{EN}_t(1-p_t^{EN})}{\mathcal{NE}_t(1-p_t^{NE})} = \frac{EN_t}{NE_t}$$

Our identifying assumption is that $\mathcal{EN}_t$ and $\mathcal{NE}_t$ respectively differ from EN_t and NE_t by roughly equal average proportions, $\mathbb{E}(p_t^{EN}) \approx \mathbb{E}(p_t^{NE})$. Then the ratio of the averages of $\mathcal{EN}_t$ and $\mathcal{NE}_t$ will tend to approximately the ratio of the averages of EN_t and NE_t:

$$\operatorname*{plim}_{T\to\infty} \frac{\frac{1}{T}\sum_{t=1}^{T}\mathcal{EN}_t}{\frac{1}{T}\sum_{t=1}^{T}\mathcal{NE}_t} \approx \frac{\frac{1}{T}\sum_{t=1}^{T}EN_t}{\frac{1}{T}\sum_{t=1}^{T}NE_t} \tag{16}$$

Panel (a) in figure 3 shows how the ratio of $\mathcal{EN}_t$ to $\mathcal{NE}_t$ has evolved from 2001 to mid-2011. The ratio has fluctuated relatively little around a particularly stable average just above 1, and the ratio of the averages over this period is accordingly 1.05. Hence, one would expect the convergence in equation (16) to happen quickly, so that we can rely on the approximation already at comparatively small values of T. We therefore find that $\frac{1}{T}\sum_{t=1}^{T}EN_t \approx \frac{1}{T}\sum_{t=1}^{T}NE_t$. As nothing can be inferred about the variation of EN_t and NE_t, we use these averages in each period.

Under the identifying assumption, subtracting equation (14) from equation (13) leads the averages of EN_t and NE_t to cancel out:

$$H_t - TS_t = UE_t - EU_t \tag{17}$$

A single step has thus eliminated the three unknowns EE_t, NE_t, and EN_t. Together with equations (11), (12), and (15), we now have a system of four equations in the four unknowns UE_t, EU_t, NU_t, and UN_t. In turn, it will be impossible to separately identify the three unknowns just eliminated, which is in fact not necessary: we will only need the sum $NE_t + EE_t$, which we denote by SE_t for short, to estimate a general matching function. This sum can be obtained from equation (13) once UE_t has been identified.[13]

[13] Using a CPS series for the stock of discouraged workers (i.e. workers who are considered non-participating after a spell in unemployment that did not lead to a job), one could write down an additional independent equation that would allow to identify NE_t separately. However, we found that this approach produces most implausible results. One problem with this CPS series might be that it is unclear when those remaining in non-participation cease to be discouraged workers.

To solve the system, we sum equations (17) and (15) and then use equation (12) to substitute for NU_t, which leads to

$$H_t - TS_t + U_t^{<5} = UE_t + \Delta U_t^{LR+} + \frac{UE_t + UN_t}{U_{t-1}} U_{t-1}^{LR+} \tag{18}$$

Next, we sum equations (17) and (11) to obtain

$$H_t - TS_t + \Delta U_t^{QS+} = UE_t - \frac{UE_t + UN_t}{U_{t-1}} U_{t-1}^{QS+} \tag{19}$$

We then solve equations (18) and (19) respectively for UN_t and equate the resulting expressions. Then one can solve for the only remaining unknown:

$$UE_t = H_t - TS_t + \frac{1}{1 + \frac{U_{t-1}^{LR+}}{U_{t-1}^{QS+}}} \left[U_t^{<5} - \Delta U_t^{LR+} + \frac{U_{t-1}^{LR+}}{U_{t-1}^{QS+}} \Delta U_t^{QS+} \right] \tag{20}$$

We next obtain EU_t from equation (17) as

$$EU_t = UE_t - H_t + TS_t = \frac{1}{1 + \frac{U_{t-1}^{LR+}}{U_{t-1}^{QS+}}} \left[U_t^{<5} - \Delta U_t^{LR+} + \frac{U_{t-1}^{LR+}}{U_{t-1}^{QS+}} \Delta U_t^{QS+} \right] \tag{21}$$

and subsequently NU_t from equation (15) as

$$NU_t = U_t^{<5} - EU_t = \frac{U_t^{<5}}{1 + \frac{U_{t-1}^{QS+}}{U_{t-1}^{LR+}}} + \frac{1}{1 + \frac{U_{t-1}^{LR+}}{U_{t-1}^{QS+}}} \left[\Delta U_t^{LR+} - \frac{U_{t-1}^{LR+}}{U_{t-1}^{QS+}} \Delta U_t^{QS+} \right] \tag{22}$$

From equation (11) (or equivalently from equation (12)), we find UN_t:

$$\begin{aligned} UN_t &= \frac{U_{t-1}}{U_{t-1}^{QS+}} \left[EU_t - \Delta U_t^{QS+} \right] - UE_t \quad = \quad TS_t - H_t \\ &+ \frac{1}{U_{t-1}^{QS+} + U_{t-1}^{LR+}} \left[(U_{t-1} - U_{t-1}^{QS+})(U_t^{<5} - \Delta U_t^{LR+}) - (U_{t-1} + U_{t-1}^{LR+}) \Delta U_t^{QS+} \right] \end{aligned}$$

Finally, given the results for UE_t and EU_t, equations (13) and (14) respectively return

$$EE_t + NE_t = H_t - UE_t, \quad EE_t + EN_t = TS_t - EU_t \tag{23}$$

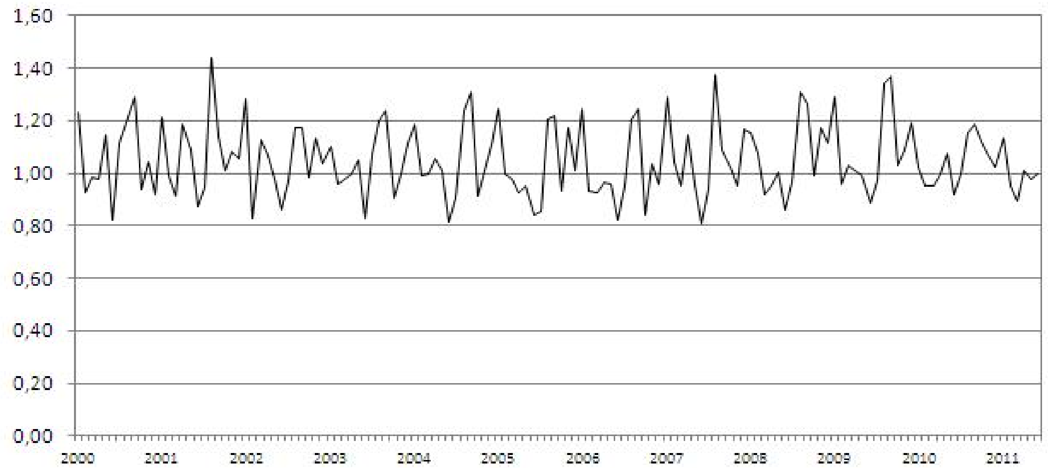
Table 4 gives the descriptive statistics of the time series that we can thus construct for the flows. Panel (b) in figure 3 depicts the series for UE_t, EU_t, and SE_t. It is worth noting that the correlation between SE_t and market tightness is 0.73, in line with theoretical results on procyclical endogenous search by employed job seekers (see Burgess (1993), for example).

We can now investigate how it matters that the JOLTS series for H_t and TS_t also count very short spells while the CPS series might not: the analytic solutions for UE_t and EU_t include the term $H_t - TS_t$, so that such spells should cancel out, as every transitory spell in employment counts equally towards H_t and TS_t. Only the values for $EE_t + NE_t$ and $EE_t + EN_t$ are thus affected and might be overestimated relative to the other flows.

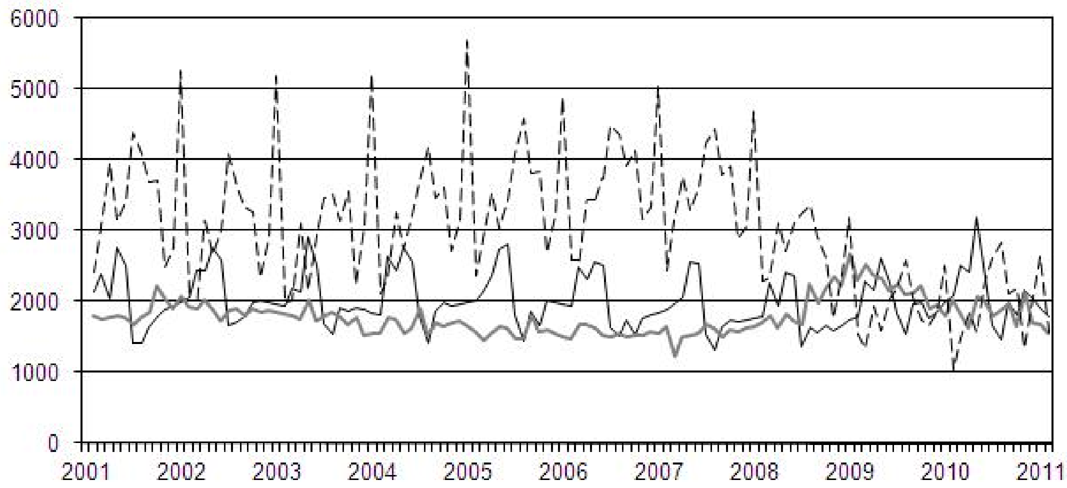
Table 4: Descriptive statistics for the constructed series

Variable	mean	st. dev.	min.	max.
U_t^{QS+}	5997	2229	3615	11038
U_t^{LR+}	3265	692	2260	4937
UE_t	1779	695	-110	3182
EU_t	1786	251	1235	2668
NU_t	1002	115	691	1148
UN_t	946	746	-332	3205
SE_t	3042	945	1036	5680

All figures are monthly levels in thousands.



(a) The ratio of CPS gross flows from employment to non-participation ($\mathcal{EN}_t$) and from non-participation to employment ($\mathcal{NE}_t$)



(b) The constructed series: the flows from non-participation and employment to employment (SE_t; top black line), from unemployment to employment (UE_t with interpolation of low outliers; bottom black line), and from employment to unemployment (EU_t; grey line). Levels in thousands

Figure 3: Assumption and result of the empirical model of labour market flows

3.4 Empirical evaluation

The series constructed from our empirical model enable us to estimate matching functions that account for all job seekers. We thereby offer a new empirical strategy to make such an estimation possible for the U.S. labour market. By using data on CPS stocks and JOLTS flows, this strategy is not limited to a rare data set, as in Anderson/Burgess (2000), nor does it, as in Jolivet (2009), rely on the CPS gross flow data (unadjusted for the problems discussed above). Like these contributions, however, we assume that the relative rates at which different job seekers transit into employment depend on their relative search intensities. Recall that ϕ is defined as the average search intensity of employed and non-participating job seekers relative to unemployed job seekers, whose search intensity is normalised to 1. Hence we assume that this ratio is mirrored by the ratio of transition rates:

$$\phi_t = \frac{SE_t}{E_{t-1} + N_{t-1}} \left(\frac{UE_t}{U_{t-1}} \right)^{-1} \tag{24}$$

With the series constructed from our model, we can now calculate ϕ_t and assemble the measure J_t of total job seekers, which allows us to estimate the log-linear model in equation (7). However, UE_t and SE_t together make up hirings H_t, so that the definition of ϕ_t makes J_t a potentially endogenous explanatory variable when we try to estimate equation (7). Line (11) in table 5 reports the OLS results.[14] The insignificance of the coefficient for j_t is counter-intuitive; this might reflect endogeneity bias. To avoid such bias, we instrument v_t and j_t respectively by v_{t-1} and j_{t-1} in line (12), which produces plausible significant coefficients (the correlation between j_t and its instrument being $r_{j_t, j_{t-1}} = 0.65$).[15] These results support the hypothesis of CRS: the Wald test statistic is 0.20 from $\chi^2(1)$, while the critical value at the 5% significance level is 3.84. Therefore, we can impose the restriction $\beta + \alpha = 1$ and estimate the model

$$\ln \frac{H_t}{J_t} = K + \gamma t + \beta \ln \frac{V_t}{J_t} + \epsilon_t$$

which is done in line (13) of table 5 using OLS. In line (14), the same model is estimated using $\ln(V_{t-1}/J_{t-1})$ as an instrument for $\ln(V_t/J_t)$ (where the correlation coefficient is 0.89). The results of these two regressions do not differ much; they are significant and plausible in both cases.

From the plausible estimation results in lines (12) through (14), the estimate that emerges for γ is exactly the same as we found using only unemployed job seekers (see section 2.2). In particular, regressions (13) and (4) are both plausible and directly comparable, yet also produce the same estimated time trend -0.0017. Regressions (14) and (10) are likewise plausible and comparable and also give -0.0017 as the estimate for γ. Such comparisons suggest that the omission of employed and non-participating job seekers does not generate

[14] Since ϕ_t is calculated using lagged values of E_t and N_t, the timing is only aligned in the IV regressions in table 5. However, aligning the timing in the OLS regressions produces only marginally different results. Further, while most of the following regressions include constructed data, we do not adjust the standard errors, as any such attempt would inevitably increase them. Then estimated time trends might be insignificant only due to the higher standard errors, while they might be significant if collected data were used.

[15] To replicate these results, insignificant dummies for July, October, and November should be dropped.

Table 5: Matching function regressions with all job seekers

estimator and T	dependent variable	v_t	j_t	$\ln(V_t/J_t)$	t	const.	seas. dum.	R^2
(11) OLS	h_t	0.498*	0.060		-0.0011*	3.801*	yes	0.93
101		0.034	0.061		0.0002	0.810		
(12) IV	h_t	0.612*	0.324*		-0.0016*	0.295	yes	0.90
90		0.050	0.095		0.0002	1.321		
(13) OLS	$\ln(H_t/J_t)$			0.647*	-0.0017*	-0.294*	yes	0.97
101				0.018	0.0002	0.037		
(14) IV	$\ln(H_t/J_t)$			0.634*	-0.0017*	-0.292*	yes	0.97
90				0.018	0.0002	0.037		

GMM with a robust weight matrix was used in the case of IV estimation.
* Significant at 1% level.
** Significant at 5% level.

the time trend. This conclusion is confirmed when we estimate equation (8) with seasonal dummies and still obtain $\hat{\gamma} = -0.0017$. Using the result $1-\hat{\beta} = 0.2647$ from this regression to quantify the bias, equation (10) gives the expectation of the estimated time trend in the specification without x_t as -0.0015. Hence there is practically no bias from the omission of other job seekers.

Finally, note that regressions with all job seekers are based on fewer observations because recurrent outliers in the JOLTS data (December and January) generate extreme outliers in UE_t and thus in ϕ_t. Using an interpolated series for UE_t led results to deteriorate, so that these outliers were simply dropped. To ensure that this does not affect our conclusions about the time trend, we repeated the regressions in lines (3) and (4) of table 2 using the same 101 observations as for the estimation with all job seekers. The estimated time trends were essentially the same as with all 122 observations (-0.0015 in both cases), so that our comparisons are not invalidated by the difference in observations.

4 Bias from the omission of flows

4.1 Theory

A growing number of papers suggest that the flows into job seekers and vacancies should also feature in the matching function, not just the stocks. The reason for the coexistence of vacancies and unmatched job seekers might be that no mutually acceptable matches can be formed among these vacancies and job seekers. If this is the case, then the inflows of new vacancies and job seekers are central to the matching process: existing job seekers match with the flow of new vacancies, while existing vacancies match with the flow of new job seekers. Where such stock-flow matching happens, a canonical matching function as in equation (1) is likely to suffer from omitted variable bias because it neglects the flows of new vacancies and job seekers, and this might affect the estimated time trend.

From their model of stock-flow matching, Coles/Smith (1998) obtain a log-linear matching function that includes stocks as well as inflows. Their analysis suggests the model

$$h_t = K + \beta v_t + \alpha u_t + \gamma t + \eta \tilde{v}_t + \zeta \tilde{u}_t + \epsilon_t \tag{25}$$

where $\tilde{v}_t$ and $\tilde{u}_t$ respectively denote the logarithm of the inflow into vacancies and unemployed job seekers. If $\eta \neq 0$ or $\zeta \neq 0$, the estimate for γ when flows are omitted will be biased and inconsistent:

$$\mathbb{E}(\hat{\gamma}') = \gamma + \eta \tau_{\tilde{v}} + \zeta \tau_{\tilde{u}} \tag{26}$$

where $\tau_{\tilde{v}}$ and $\tau_{\tilde{u}}$ are the coefficients of t in regressions of $\tilde{v}_t$ and $\tilde{u}_t$, respectively, on all included explanatory variables.

To the best of our knowledge, only Gregg/Petrongolo (2005) discuss time trends in the context of stock-flow matching. Their somewhat different approach is to estimate separate models for the outflows from the stocks of vacancies and unemployment. They argue that stock-flow matching implies

$$UE_t = g(U_t, V_t, \tilde{V}_t, t) \tag{27}$$

for some function $g(\cdot)$ that is increasing in U_t, V_t, and $\tilde{V}_t$. As explained in Coles/Petrongolo (2008), $\tilde{U}_t$ is not included because newly unemployed job seekers are expected to match primarily with existing vacancies, thus hardly affecting job seekers in the stock of unemployment who match primarily with the flow of vacancies. We can estimate equation (27) using the series for UE_t from our model, while Gregg/Petrongolo (2005) can apparently not distinguish between UE_t and UN_t in their data and therefore have to use the sum as dependent variable.

We can also extend the stock-flow reasoning to non-participating and employed job seekers here, using our constructed series. Let $O_t = \phi_t(E_t + N_t)$ denote these other job seekers. Those in O_t who are at risk of becoming unemployed can be thought of as an inflow into unemployment; they match primarily with existing vacancies in V_t, typically before they are counted towards the stock of unemployed. To the extent that they compete with unemployed job seekers for vacancies in V_t, U_t might also play a role. Others in O_t remain in their current status and wait for a suitable vacancy to appear; they thus match primarily with vacancies in $\tilde{V}_t$. The flow of hirings from O_t, which is SE_t, should obey

$$SE_t = q(O_t, U_t, V_t, \tilde{V}_t, t) \tag{28}$$

for some function $q(\cdot)$ that is non-increasing in U_t but increasing in O_t, V_t, and $\tilde{V}_t$.

Gregg/Petrongolo (2005) find significant negative time trends only in the outflow from unemployment, and the magnitude appears to halve as they switch from a model of random matching to a model of stock-flow matching. They attribute the negative time trend in standard matching functions to a rise over time of V_t relative to $\tilde{V}_t$. This view thus seems analogous to the familiar argument about the share of long-term unemployed (see section 2.3): V_t largely consists of vacancies that have not been taken up before and are thus less likely to be taken up than those in $\tilde{V}_t$. A growing share of 'long-term vacancies' would bias the time trend downwards in a canonical matching function that, by omitting $\tilde{V}_t$, does not distinguish between short-term and long-term vacancies.

4.2 Empirical evaluation

In this section, we estimate equations (25) through (28) for the U.S. labour market. The empirical model we built in section 3.3 allows us to do this without additional data. Recall from equation (15) that we identified the total inflow into unemployment, $\tilde{U}_t$, with $U_t^{<5}$. Next, the spirit of the model implies a further accounting equation:

$$\Delta V_t = \tilde{V}_t - H_t \tag{29}$$

Assuming that vacancies are only closed when filled, the change in vacancies is given by the difference in the inflow into vacancies and hires. As V_t and H_t are known from the JOLTS data, this implies a series for $\tilde{V}_t$. We can now estimate equation (25), and the OLS results are reported in line (15) of table 6. While the estimated coefficients of u_t and $\tilde{u}_t$ are negative and significant, that of v_t is insignificant, and that of $\tilde{v}_t$ is positive and significant. With the exception of the insignificance of v_t, these results accord qualitatively with the results in Coles/Smith (1998). They interpret the finding of a negative coefficient for $\tilde{u}_t$, which appears recurrent in estimated stock-flow matching functions, as a crowding-out effect among the unemployed.

The estimated time trend in line (15) is very low and only significant at the 5% significance level. However, this regression is not necessarily reliable because it ignores the endogeneity of $\tilde{v}_t$, which is constructed from equation (29) and therefore depends itself on hirings. We employ three-stage least squares (3SLS) to account for this endogeneity. The second equation of the simultaneous equations system is a linear regression model of $\tilde{v}_t$ determined by h_t, v_t, and v_{t-1} (plus a time trend and seasonal dummies). Because logarithms and levels are very highly correlated, this regression model would reflect the relationship between levels in equation (29). This empirical formulation of equation (29) appears to be borne out by the data, as we discuss in detail in section 5.2 below.

The results of the 3SLS estimation in line (16) of table 6 suggest a significant positive coefficient for v_t. As in line (15), the estimated coefficients of u_t and $\tilde{u}_t$ are both negative and significant, while the estimated coefficient of $\tilde{v}_t$ is positive and significant. We obtain analogous results in line (17) where we include v_{t-1} and u_{t-1} rather than v_t and u_t. To estimate the exact specification employed by Coles/Smith (1998), one can repeat the regression in line (17) with $\ln(H_t/U_{t-1})$ as dependent variable and obtain the same qualitative results (with a coefficient near -1 for u_{t-1}). All these results are thus in exact qualitative accordance with the results in Table 3 in Coles/Smith (1998).

Lines (16) and (17) report negative estimated time trends that are again significant at the 1% significance level. Their magnitude of -0.0007 is much lower than before, around half of the magnitude we found using either a standard approach or a model with all job seekers. This indicates that a substantial part of the estimated time trend in standard matching functions may be due to the omission of inflow measures. We reach the same

Table 6: Stock-flow matching function regressions

dep. var. est., T	v_t	u_t	v_{t-1}	u_{t-1}	θ_t	o_t	$\tilde{v}_t$	$\tilde{u}_t$	$\ln \frac{\tilde{V}_t}{U_t}$	$\ln \frac{\tilde{U}_t}{U_t}$	t	const.	seas. dum.	R^2
(15) h_t	-0.038	-0.252*					0.441*	-0.120*			-0.0003**	8.142*	yes	0.97
OLS, 121	0.048	0.037					0.041	0.043			0.0001	0.770		
(16) h_t	0.140**	-0.200*					0.155*	-0.205*			-0.0007*	9.270*	yes	0.97
3SLS, 121	0.057	0.041					0.058	0.047			0.0002	0.876		
(17) h_t			0.128*	-0.150*			0.293*	-0.208*			-0.0007*	7.809*	yes	0.97
3SLS, 121			0.042	0.043			0.043	0.048			0.0001	1.231		
(18) f_t					0.920*				-0.234	0.022	-0.0020*	-0.039	yes	0.97
3SLS, 121					0.129				0.151	0.077	0.0004	0.058		
(19) $\ln UE_t$			0.199**	0.476*			0.342*				-0.0015*		yes	0.89
3SLS, 101			0.093	0.094			0.080				0.0003			
(20) $\ln SE_t$	-0.086	-0.748*				0.377*	0.469*				0.0000		yes	0.98
3SLS, 101	0.065	0.047				0.034	0.060				0.0002			

* Significant at 1% level.
** Significant at 5% level.

conclusion based on equation (26): taking the estimate -0.00028 in line (15) as the true value of γ, we obtain $\mathbb{E}(\hat{\gamma}') = -0.00078$ for the time trend in a standard matching function. That is, if inflow measures are omitted, the resulting bias will increase a negative time trend of very small magnitude in the stock-flow matching function to roughly half the magnitude estimated for the standard matching function. However, the other half of the time trend in the standard matching function remains as unexplained as the fact that there is apparently a negative time trend also in the stock-flow matching function.

The hypothesis of CRS (here $\beta + \alpha + \eta + \zeta = 1$) is soundly rejected for the regressions in line (16) and (17). We nevertheless include a CRS specification here because it has been considered in the literature (see for example model 2 in Gregg/Petrongolo (2005)). We therefore want to estimate the model

$$f_t = K + \beta\theta_t + \gamma t + \eta \ln \frac{\tilde{V}_t}{U_t} + \zeta \ln \frac{\tilde{U}_t}{U_t} + \epsilon_t$$

while accounting for the endogeneity of $\ln(\tilde{V}_t/U_t)$. To this end, we divide equation (29) by U_t and make this the basis for a linear second equation that now endogenously determines $\ln(\tilde{V}_t/U_t)$. When we then apply 3SLS to the system, we obtain the results in line (18) of table 6. These results are poor, a likely consequence of the invalid CRS assumption.

In lines (19) and (20), we estimate the models in equation (27) and (28), respectively. For simplicity, we assume that $g(\cdot)$ and $q(\cdot)$ can also be written in a log-linear form. In line (19), we employ 3SLS as before, while the simultaneous equations system estimated for line (20) includes a third equation that determines o_t, which is also potentially endogenous by construction.[16] The estimates in line (19) are all significant and signed as expected. Interestingly, the estimated time trend is strongly negative and significant. In line (20), a significant positive estimate might have been expected for the coefficient of v_t where the results indicate a zero coefficient instead. The other estimates from this regression, all significant at the 1% significance level, are in line with our expectations. The constant has in both cases been dropped by the estimation procedure.

In conclusion, it appears that the magnitude of the estimated negative time trend roughly halves when the flows into unemployment and vacancies are included as explanatory variables in the matching function. At the same time, the evidence in this section clearly suggests that a significant negative time trend remains. In other words, the approach based on stock-flow matching does not fully account for the time trend in standard matching functions. The next section thus goes a step further in order to explain the entire time trend.

[16] O_t is itself determined by SE_t through ϕ_t. Using equation (24),

$$O_t = \phi_t(E_t + N_t) = \frac{SE_t}{UE_t}\frac{E_t + N_t}{E_{t-1} + N_{t-1}}U_{t-1}$$

A logarithmic transformation of the latter formulation returns the third equation of the system.

5 Bias from ignoring vacancy dynamics

5.1 Theory

The estimation of the matching function might be affected by vacancy dynamics that mechanically link H to V and $\tilde{V}$, independently of any structural relationships $H = m(V, U, t)$ or, as in the previous section, $H = m(V, U, \tilde{V}, \tilde{U}, t)$. To see this, let us recall the law of motion for vacancies that we have specified in equation (29) and rewrite it as

$$H_t = \tilde{V}_t - V_t + V_{t-1} \tag{30}$$

Hence the law of motion for vacancies implies a competing model for the determination of hires, and the model shares one explanatory variable with a standard matching function and two explanatory variables with a stock-flow matching function. This raises the possibility that, when we intend to estimate a matching function, our results are in fact driven by the relationship in equation (30). Recall that this relationship roughly holds as an accounting identity provided only few vacancies disappear without being filled, so that the relationship might well manifest strongly in empirical results.

Suppose the extreme case that equation (30) holds exactly while the standard matching function simply does not exist, so that there is no structural relationship $H = m(V, U, t)$. Then the estimation of a standard matching function would in fact be an estimation of equation (30) that omits $\tilde{V}_t$ and either V_t or V_{t-1}. Since one of the latter is always included and V_t is highly autocorrelated, the coefficients of the included variables would always be biased by these omissions. In particular, the variables t and U_t with a true coefficient 0 in equation (30) could appear to have significant explanatory power.

To abstract from the difference between levels and logarithmic values, we note again that the coefficient of the correlation between them is almost 1 and take

$$h_t = \rho_{\tilde{v}_t} \tilde{v}_t - \rho_{v_t} v_t + \rho_{v_{t-1}} v_{t-1} + \gamma t + \epsilon_t \tag{31}$$

to be the true model instead of equation (30). We include a time trend here because, if e.g. V_t grows over time, so will the discrepancy $V_t - v_t$. This tendency might lead to a time trend that does not reflect any real changes, but only the inappropriate logarithmic transformation. For the same reason, the coefficients $\rho_{\tilde{v}_t}$, ρ_{v_t}, and $\rho_{v_{t-1}}$ will not necessarily equal 1, -1, and 1, respectively. When we then estimate a standard matching function as

$$h_t = K + \beta v_t + \alpha u_t + \gamma' t + \epsilon_t$$

under the maintained assumption that this function does not exist ($K = \beta = \alpha = \gamma' = 0$), we expect to obtain the following coefficients:

$$\begin{aligned} \mathbb{E}(\hat{\beta}) &= -\rho_{v_t} + \rho_{\tilde{v}_t} \beta_{\tilde{v}_t} + \rho_{v_{t-1}} \beta_{v_{t-1}} && (32) \\ \mathbb{E}(\hat{\alpha}) &= \rho_{\tilde{v}_t} \alpha_{\tilde{v}_t} + \rho_{v_{t-1}} \alpha_{v_{t-1}} && (33) \\ \mathbb{E}(\hat{\gamma}') &= \gamma + \rho_{\tilde{v}_t} \tau_{\tilde{v}_t} + \rho_{v_{t-1}} \tau_{v_{t-1}} && (34) \end{aligned}$$

where $\beta_{\tilde{v}_t}$ and $\beta_{v_{t-1}}$ are the respective coefficients of v_t in auxiliary regressions of $\tilde{v}_t$ and v_{t-1} on all included explanatory variables; $\alpha_{\tilde{v}_t}$, $\alpha_{v_{t-1}}$, $\tau_{\tilde{v}_t}$, and $\tau_{v_{t-1}}$ are defined analogously. The expected coefficients are thus the sum of the 'true' coefficient in equation (31) and the bias induced by the omission of $\tilde{v}_t$ and v_{t-1}. For the case that a stock-flow matching function as in equation (25) is estimated instead, we can express the expected coefficients in the same fashion, noting that equation (25) does not omit $\tilde{v}_t$.

5.2 Empirical evaluation

To evaluate the role that vacancy dynamics play for our previous empirical results, let us first explore to what extent equation (30) is supported by our data. In order to account for the endogeneity of $\tilde{v}_t$, we instrument it by $\tilde{v}_{t-3}$, as the correlation coefficient $r_{\tilde{v}_t,\tilde{v}_{t-3}} = 0.56$ is somewhat higher than that for other lagged values. The results in line (21) of table 7 suggest that the coefficients of v_t, v_{t-1}, and $\tilde{v}_t$ are all significant at the 1% significance level and that v_t enters negatively, exactly as one would expect given equation (30).

These qualitative results appear to be robust over various similar specifications. For example, h_{t-1} is also strongly correlated with $\tilde{v}_t$ and can thus be used as an alternative instrument ($r_{\tilde{v}_t,h_{t-1}} = 0.44$). Line (22) reports the results for a regression that is otherwise the same as in line (21). The estimates are now even closer to the coefficients in equation (30): -1, 1 and 1 for v_t, v_{t-1}, and $\tilde{v}_t$, respectively. Indeed, when we test the hypothesis $\rho_{\tilde{v}_t} + \rho_{v_t} + \rho_{v_{t-1}} = 1$ for regressions (21) and (22), we respectively obtain Wald test statistics of 0.37 from $\chi^2(1)$ and 2.41 from $\chi^2(1)$. The critical value at the 5% significance level being 3.84, we fail to reject this hypothesis.

Also in line with equation (30), almost all seasonal dummies and the constants in lines (21) and (22) are insignificant. Hence we run a regression without them in line (23), using OLS in this case to prepare later results, and the pattern we found persists. Line (23) thus uses the exact specification in equation (31) and finds strongly supportive evidence for this specification. In line (24), we repeat this OLS regression with a constant and seasonal dummies included. In analogy to specifications with CRS, we finally divide through equation (30) by U_t, write down a modified model corresponding to equation (31), and estimate it without a constant and seasonal dummies. The estimated coefficients in line (25) fully confirm our previous results.

That regressions (21) to (25) bear out equation (30) is not wholly surprising, given that $\tilde{V}_t$ was constructed to fit this relationship. In appendix B, however, we also find strong evidence for it using series from Germany, none of which we construct. It thus appears indeed that few vacancies disappear without being filled, so that equation (30) is very likely to hold approximately. The fact that we can replicate the results in Coles/Smith (1998) using our constructed series on $\tilde{V}_t$ (see section 4) also leads us to trust our results. While a time trend cannot play a role in equation (30), the results in lines (23) and (24) include significant negative time trends, albeit of very small magnitude. We would attribute this to the correlation of time and such discrepancies as $V_t - v_t$ introduced by the (invalid) use of logarithms in equation (31). Another possible source of these trends is the endogeneity of $\tilde{v}_t$ that regressions (23) and (24) ignore.

Table 7: The law of motion for vacancies and comparisons to previous regressions

dep. var. est., T	v_t	u_t	v_{t-1}	u_{t-1}	θ_t	$\ln \frac{V_{t-1}}{U_t}$	$\tilde{v}_t$	$\tilde{u}_t$	$\ln \frac{\tilde{V}_t}{U_t}$	t	const.	seas. dum.	R^2
(21) h_t	-0.762*		0.746*				1.036*			-0.0001	-0.150	yes	0.99
IV, 118	0.078		0.052				0.061			0.0001	0.275		
(22) h_t	-0.911*		0.843*				1.151*			0.0000	-0.677	yes	0.99
IV, 121	0.121		0.078				0.099			0.0001	0.446		
(23) h_t	-0.682*		0.735*				0.950*			-0.0001*		no	1.00
OLS, 121	0.021		0.014				0.010			0.0000			
(24) h_t	-0.621*		0.665*				0.911*			-0.0003*	0.394*	yes	0.99
OLS, 121	0.039		0.028				0.028			0.0001	0.139		
(25) f_t					-0.728*	0.740*			0.969*	-0.0001		no	0.99
OLS, 121					0.024	0.018			0.013	0.0001			
(26) h_t	0.298*	-0.172*								-0.0008*	7.384*	yes	0.95
OLS, 122	0.059	0.055								0.0002	0.949		
(27) h_t	0.307	-0.164								-0.0008	7.234		
(28) h_t			0.215*	-0.017			0.490*	-0.076		-0.0007*	3.247*	yes	0.98
OLS, 121			0.039	0.039			0.034	0.042		0.0001	1.042		
(29) h_t			0.211	-0.022			0.490	-0.072		-0.0007	3.277		
(30) f_t					0.750*					-0.0017*	-0.039	yes	0.98
OLS, 122					0.013					0.0002	0.021		
(31) f_t					0.739					-0.0016	-0.051		

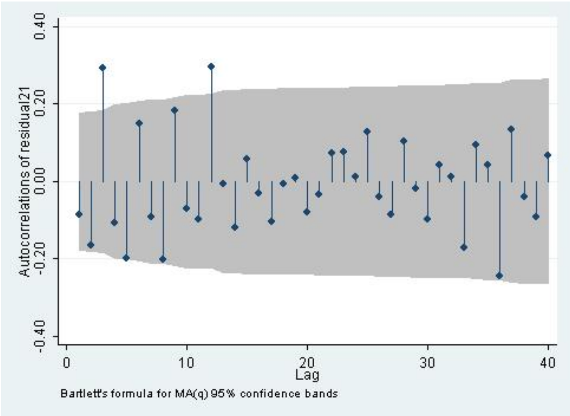
GMM with a robust weight matrix was used in all cases of IV estimation.
* Significant at 1% level.
** Significant at 5% level.

As in the previous sections, we ask whether the negative time trend in the standard matching function can be explained by the omission of variables that, by equation (30), also determine hirings. In contrast to previous sections, however, we do not compare different specifications of the matching function in this case. Rather, equation (30) is based on an altogether different theoretical motivation. The key question is therefore *whether the results obtained when we estimate a matching function do indeed reflect a matching function or the law of motion for vacancies instead*. An answer to this question should perhaps not be based on the investigation of just one estimated coefficient. Therefore, we explore how well each relevant coefficient can be explained when we regard equation (30) as the true data-generating process, the regression in line (24) as its empirical counterpart, and the coefficients estimated for the standard matching function accordingly as described by equations (32) through (34).

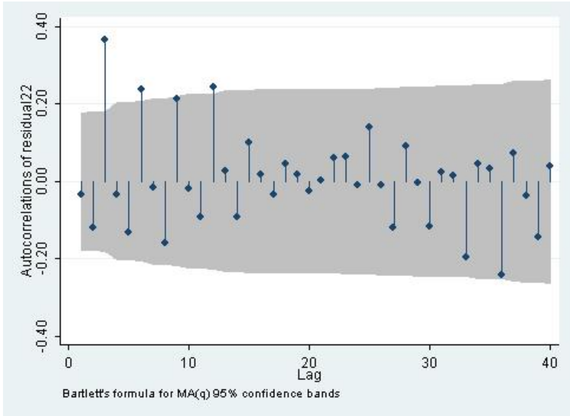
Line (26) of table 7 first recalls the OLS estimates for the standard matching function from line (2) in table 2. Line (27) reports the coefficients we would expect based on equations (32) through (34), and the expected constant is found analogously. The 'predicted' coefficients in line (27) are all very close to the coefficients actually obtained in line (26). There is even a virtual match for the time trend, suggesting that the trend only arises in the matching function because vacancy dynamics are ignored. We next repeat this exercise for the stock-flow matching function. In line (28) we first report the OLS results for a specification of the stock-flow matching function that is analogous to lines (16) and (17) in table 6. Line (29) then gives the coefficients 'predicted' in the same fashion as before. Again the 'predicted' and the actual coefficients hardly differ, while the time trends even match. The OLS results in line (15) of table 6 can also be predicted well.

Let us finally consider the standard matching function with CRS. Line (30) recalls the results in line (4) of table 2. In this context, we regard the regression in line (25) as the empirical counterpart of the true data-generating process and use it to 'predict' coefficients for the standard matching function with CRS, which are reported in line (31). They are close to the estimated coefficients but not as close as before. This is likely due to the absence of a constant and of seasonal dummies in line (25). Similarly, when we use the regression in line (23) instead of line (24) as the true benchmark above, the predictions for both the standard and the stock-flow matching functions become marginally worse.

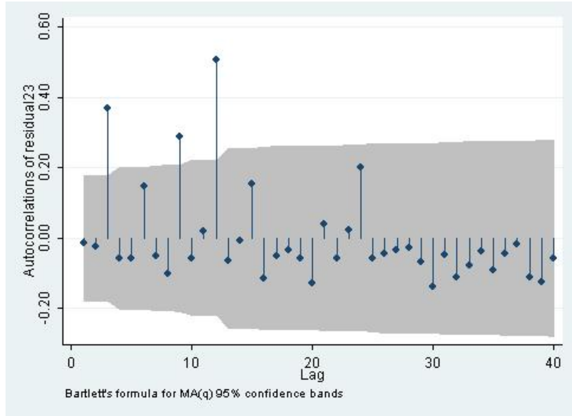
Across the specifications of the matching function considered here, the estimated time trends are very well 'predicted'. Therefore, they can apparently be interpreted as the sum of a time trend in the empirical law of motion for vacancies and omitted variable bias. The former is almost 0 and is typically insignificant in table 7. We have also argued that this time trend does not correspond to any real changes. The omitted variable bias arises for two reasons. Firstly, v_t, v_{t-1}, and $\tilde{v}_t$ all turn out to be important determinants of h_t. Secondly, time as an explanatory variable is correlated with all these determinants; the correlation coefficients for our data are $r_{t,v_t} = -0.45$, $r_{t,v_{t-1}} = -0.46$, and $r_{t,\tilde{v}_t} = -0.37$. Hence the estimated coefficient of time will be biased whenever any of them is omitted, as it happens in matching functions. Given that the biased estimate we expect in matching functions coincides with the estimate we obtain, we conclude that taking the law of motion for vacancies into account can fully explain the time trends in matching functions.



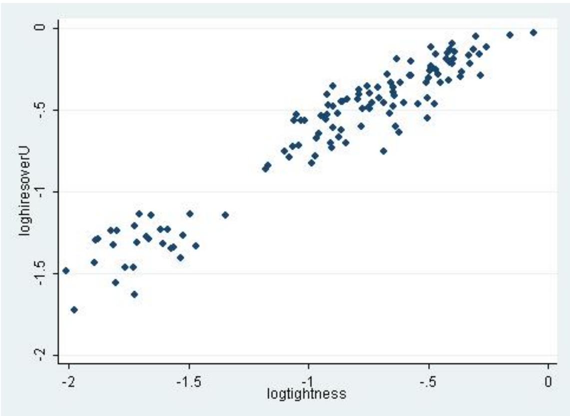
(a) Residuals of regression (21)



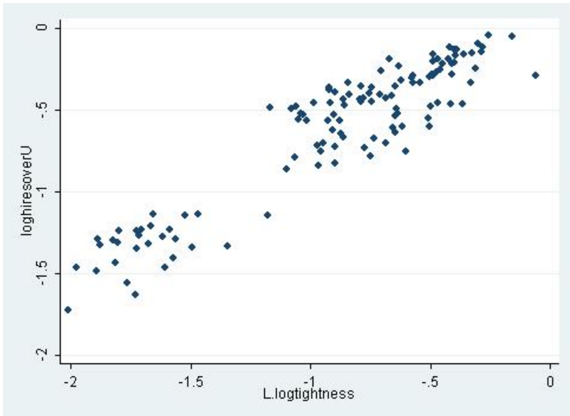
(b) Residuals of regression (22)



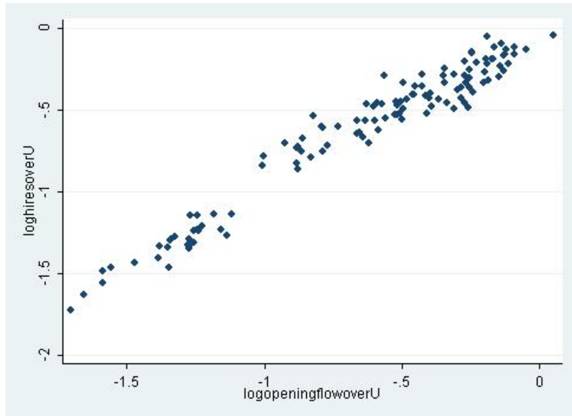
(c) Residuals of regression (23)



(d) f_t plotted against θ_t



(e) f_t plotted against θ_{t-1}



(f) f_t plotted against $\ln(\tilde{V}_t/U_t)$

Figure 4: Panels (a) to (c): autocorrelation of the regression residuals for vacancy dynamics. Panels (d) to (f): scatter plots for the job finding rate and measures of market tightness

The close match between virtually all estimated and 'predicted' coefficients further suggests that the equations used for the prediction, such as equations (32) through (34), are accurate. These equations were based on the hypothesis that $K = \beta = \alpha = \gamma' = 0$ in the standard matching function, and similarly for the stock-flow matching function. We find no evidence against this hypothesis. It thus appears that the variables in matching functions do not play any other role than reflecting the law of motion for vacancies. We reach the same conclusion where we perform the same exercise in appendix B with the data from Germany, so that our results here are likely not driven by the way $\tilde{V}_t$ is constructed.

It is worth noting that the reverse prediction, i.e. the coefficients in the law of motion predicted by the coefficients in the matching function, is very poor. For example, we could regard the stock-flow matching function estimated in line (28) as the true data-generating process and view the law of motion estimated in line (24) as a stock-flow matching function where u_{t-1} and $\tilde{u}_t$ are omitted while v_t is included instead. Applying the same type of formula as before, we obtain the following 'predictions' for the coefficients in line (24): $\mathbb{E}(\hat{\rho}_{v_t}) = -0.025$, $\mathbb{E}(\hat{\rho}_{v_{t-1}}) = 0.255$, $\mathbb{E}(\hat{\rho}_{\tilde{v}_t}) = 0.535$, and $\mathbb{E}(\hat{\gamma}) = -0.0006$. These values are not nearly close to the coefficients estimated in line (24). The prediction is also very poor when we regard the standard matching function as true data-generating process.

We can also use more conventional F-tests to investigate which variables determine hirings. Consider an OLS regression that includes v_t, v_{t-1}, $\tilde{v}_t$, u_{t-1}, and $\tilde{u}_t$, a time trend, a constant, and seasonal dummies. Call this the unrestricted model. Then we can test whether u_{t-1} and $\tilde{u}_t$ are irrelevant, as equation (30) implies, by testing a restriction on their coefficients α and ζ. The null hypothesis is $\alpha = \zeta = 0$; if it is true, the explanatory power of the unrestricted regression should not be significantly higher than that of the restricted regression where u_{t-1} and $\tilde{u}_t$ are dropped. In other words, if h_t is only determined by the law of motion for vacancies, dropping u_{t-1} and $\tilde{u}_t$ will not make a difference as long as v_t, v_{t-1}, and $\tilde{v}_t$ remain. The F-test statistic is 0.05 from $F_{(2,103)}$ and the critical value at the 5% significance level from a $F_{(2,100)}$ is 3.09. We therefore fail by a large margin to reject H_0; nothing appears to be missed for the explanation of h_t when u_{t-1} and $\tilde{u}_t$ are dropped. We can similarly test whether v_t has something to add even when v_{t-1} is also included. That is, we now consider the restriction that the coefficient of v_t in the unrestricted model is 0, so that it can be dropped and a stock-flow matching function remains. The F-test statistic is 234.01 from $F_{(1,103)}$, while the critical value at the 1% significance level from a $F_{(1,100)}$ is 6.90. We can therefore soundly reject the hypothesis that v_t may be dropped without a significant loss of explanatory power. These F-tests thus fully confirm our earlier conclusions: the variables in the law of motion for vacancies, and only these variables, determine h_t. Our 'prediction' exercise above has gone beyond coarse F-tests in so far as it discusses the magnitude of the coefficients, not just whether they differ from 0, and argues that the coefficients reflect exclusively the law of motion. We perform the same two F-tests also for the German data in appendix B, with the same test results.

Finally, there are some more regression diagnostics that better accord with or point to the law of motion for vacancies. Recall from figure 1 that the residuals of standard matching functions, also with CRS, exhibit systematic autocorrelation, which often indicates a misspecified regression model. By contrast, panels (a) to (c) in figure 4 do not seem to suggest as strong an autocorrelation in the residuals of the law of motion for vacancies (except for the autocorrelation across quarters and years). Panels (d) to (f) in this figure point to a linear relationship between f_t and market tightness or $\ln(\tilde{V}_t/U_t)$. While linearity makes perfect sense with equation (30), one would expect concavity in a matching function: matches always involve one vacancy and one job seeker, so that increasing only the number of vacancies should eventually run into decreasing returns. Further, equation (30) exhibits constant returns to scale: the coefficients necessarily sum to exactly 1. This would be in line with the numerous findings of CRS in estimated matching functions. Last, the very fact that matching functions feature an otherwise unexplained negative time trend while vacancy dynamics do not points to the latter as the driver of empirical results.

6 Conclusions

The common finding of a negative time trend in empirical matching functions suggests that labour market performance deteriorates over time. Especially for recent years, this result is at odds with improvements one would expect due to new information and communication technologies. Attempts to explain the time trend by real economic changes have had limited success. Therefore, we consider the possibility that the estimated time trend is negative as a result of downward bias. We investigate bias generated by the omission of job seekers beyond the unemployed, by the omission of inflow measures, or by ignoring vacancy dynamics.

Using recent U.S. labour market data and an empirical model that accounts for all labour market flows, we can construct series for employed and non-participating job seekers and for various flow measures, on which reliable data are unavailable. The constructed series enable us to estimate matching functions that do not make the omissions we examine here. Since these constructed series seem by and large plausible and deliver estimation results in line with comparable studies, they may also prove useful for future empirical analyses.

It turns out that the inclusion of employed and non-participating job seekers does not affect the magnitude of the estimated time trend. When inflow measures are included in the matching function, this magnitude drops by about 50%, but the time trend remains significant. When we account for vacancy dynamics, however, we can explain the entire magnitude of the estimated time trend by omitted variable bias. These results suggest that the finding of deteriorating labour market performance may well be a statistical illusion.

Our further examination of the role of vacancy dynamics even raises doubts about the empirical matching function as such: regarding the estimated coefficients in the law of motion for vacancies as true, we can precisely predict each estimated coefficient in the matching function as the sum of the true coefficient and omitted variable bias. By contrast, the reverse prediction is not nearly accurate. It thus appears possible that empirical matching functions ultimately only reflect the law of motion for vacancies. Given that the empirical matching function features as a central structural relation in many models relevant for policy, this possibility needs to be assessed by further research.

A Endogenous selectivity

While the following theoretical arguments apply very generally, we present them in the same set-up as Stevens (2007) (with exogenous search intensity) who points to the potential empirical consequences when endogenous selectivity is ignored. Time is continuous and all workers are either unemployed and searching or employed and not searching. Only when an unemployed worker meets a firm the productivity y of this potential match is realised as a random draw from the continuous distribution function $F(y)$ and observed by the firm and the worker. The value of search to an unemployed worker is given by

$$rY^U = b + \frac{m(z,V,U)}{U}\pi S(y) \tag{35}$$

where r is the common discount rate, b is the unemployment benefit as a flow payoff, $S(y)$ is the expected surplus generated by a match, $\pi \in (0,1)$ is the surplus share the worker obtains, and $m(z,V,U)$ is the matching function. Similarly for the value to a firm of offering a vacancy:

$$rY^V = -c + \frac{m(z,V,U)}{V}(1-\pi)\,S(y) \tag{36}$$

where c is the flow cost of the firm's recruitment efforts. The only non-standard feature in this set-up is the dependence of the matching function on the joint reservation productivity z, so that only matches with $y \geq z$ are actually concluded:

$$z = rY^U + rY^V$$

In words, the match is not concluded when y is so low that at least one agent prefers to continue searching. In the generalised matching function in equation (6), the probability that a contact leads to a match is therefore $p(z) = 1 - F(z)$. Assuming free entry of firms while the mass of workers is fixed, rY^V is driven to 0 so that, from equation (36),

$$S(y) = \frac{V}{m(z,V,U)}\frac{c}{1-\pi}$$

which allows us to substitute for $S(y)$ in equation (35). Noting that $z = rY^U$ when $rY^V = 0$, we thus obtain

$$z = b + \frac{V}{U}\frac{c\pi}{1-\pi} \tag{37}$$

Hence endogenous selectivity, captured by z, depends positively on market tightness.

B Vacancy dynamics in German data

To check the results we have obtained with U.S. data, here we look at German data provided by the Federal Employment Service (*Bundesagentur für Arbeit*). We have series on U_t and $\tilde{U}_t$ from registered unemployed job seekers. We use a series on the outflow of registered unemployed into employment as our dependent variable and denote it UE_t. This outflow is the appropriate dependent variable when only unemployed job seekers are considered, as hirings would also count other job seekers. However, some observations in this series have been estimated by the data provider. We obtain series on V_t and $\tilde{V}_t$ from the source data of the *BA-Stellenindex* or BA-X. This vacancy measure is designed to account for both registered and unregistered vacancies. The vacancy data have only been collected since January 2004, which limits our observations here to $T = 71$ (up to November 2009). All series are seasonally unadjusted.

Table 8: Analyses with data from Germany

dep. var. est., T	v_t	u_t	v_{t-1}	u_{t-1}	θ_t	$\ln \frac{V_{t-1}}{U_t}$	$\tilde{v}_t$	$\tilde{u}_t$	$\ln \frac{\tilde{V}_t}{U_t}$	t	const.	seas. dum.	R^2
(A1) $\ln UE_t$	-1.495*		0.863**				1.733*			0.0004		no	0.99
OLS, 70	0.462		0.402				0.162			0.0015			
(A2) $\ln \frac{UE_t}{U_t}$					-1.283*	0.840*			1.127*	0.0028*	-0.383	yes	0.96
OLS, 70					0.216	0.204			0.119	0.0005	0.256		
(A3) $\ln \frac{UE_t}{U_t}$					0.230*					0.0028*	-2.533*	yes	0.89
OLS, 70					0.042					0.0009	0.188		
(A4) $\ln \frac{UE_t}{U_t}$					0.228					0.0029	-2.548		
(A5) $\ln UE_t$			-0.262*	0.041			0.727*	0.007		0.0011	6.287**	yes	0.93
OLS, 70			0.077	0.086			0.124	0.120		0.0008	2.281		

* Significant at 1% level.
** Significant at 5% level.

In line (A1) of table 8, we first verify that equation (30) roughly holds for the logarithms of the series even though $\tilde{V}_t$ has not been constructed to fit it. Line (A2) finds particularly strong evidence after scaling all series by U_t, despite the inclusion of a constant and of seasonal dummies. We next estimate a standard matching function with CRS in line (A3). Line (A4) reports the coefficients we would expect based on the hypothesis that line (A2) is the empirical counterpart of the true data-generating process. These coefficients closely match the estimated coefficients in line (A3).

We finally estimate a stock-flow matching function in line (A5) and find insignificant estimated coefficients of u_{t-1} and $\tilde{u}_t$ as well as, implausibly, a significant negative estimated coefficient of v_{t-1}. This contrasts with the results in Fahr/Sunde (2009) who estimate stock-flow matching functions for Germany. The discrepancies might arise primarily from their use of registered vacancies that had weak explanatory power in our analyses, so that the estimated coefficient of v_t is insignificant in their regressions. We suspect that u_t and $\tilde{u}_t$ then proxy for actual vacancies in their regressions and only thereby become significant.

To test which variables are relevant, we first run an unrestricted regression that includes all the variables in regression (A5) as well as v_t. With the null hypothesis that u_{t-1} and $\tilde{u}_t$ may be excluded from the regression, the F-test statistic is 1.88 from $F_{(2,52)}$ and the critical value at the 5% significance level from a $F_{(2,50)}$ is 2.79. We thus fail to reject the null hypothesis; u_{t-1} and $\tilde{u}_t$ may be dropped. Let us now adopt the null hypothesis that v_t can be excluded. With the F-test statistic equal to 6.78 from $F_{(1,52)}$ and the critical value at the 5% significance level from a $F_{(1,50)}$ equal to 4.03 in this case, we reject the null hypothesis.

References

[1] Abowd, John M.; Zellner, Arnold (1985): Estimating Gross Labor-Force Flows, Journal of Business & Economic Statistics 3, 254-283

[2] Anderson, Patricia M.; Burgess, Simon M. (2000): Empirical Matching Functions: Estimation and Interpretation Using State-Level Data, Review of Economics and Statistics 82, 93-102

[3] Blanchard, Olivier J.; Diamond, Peter A. (1990): The Aggregate Matching Function, in Diamond, Peter A. (ed.) (1990): Growth, Productivity and Unemployment. Essays to Celebrate Bob Solow's Birthday, Cambridge, MA: MIT Press, 159-206.

[4] Bleakley, Hoyt; Fuhrer, Jeffrey C. (1997): Shifts in the Beveridge Curve, Job Matching, and Labor Market Dynamics New England Economic Review September/October 1997, 3-19

[5] Borowczyk-Martins, Daniel; Jolivet, Grégory; Postel-Vinay, Fabien (2011): Accounting for Endogenous Search Behavior in Matching Function Estimation IZA mimeo

[6] Broersma, Lourens; van Ours, Jan (1999): Job Searchers, Job Matches and the Elasticity of Matching, Labour Economics 6, 77-93

[7] Burdett, Kenneth; Coles, Melvyn G.; van Ours, Jan (1994): Temporal Aggregation Bias in Stock-Flow Models, CEPR mimeo

[8] Burgess, Simon M. (1993): A Model of Competition Between Unemployed and Employed Job Searchers: An Application to the Unemployment Outflow Rate in Britain, Economic Journal 103, 1190-1204

[9] Coles, Melvyn G.; Petrongolo, Barbara (2008): A Test Between Stock-Flow Matching and the Random Matching Function Approach, International Economic Review 49, 1113-1141

[10] Coles, Melvyn G.; Smith, Eric (1998): Marketplaces and Matching, International Economic Review 39, 239-254

[11] Fahr, René; Sunde, Uwe (2009): Did the Hartz Reforms Speed-Up the Matching Process? A Macro-Evaluation Using Empirical Matching Functions, German Economic Review 10, 284-316

[12] Fahr, René; Sunde, Uwe (2004): Occupational Job Creation: Patterns and Implications, Oxford Economic Papers 56, 407-435

[13] Fallick, Bruce; Fleischman, Charles A. (2004): Employer-to-Employer Flows in the U.S. Labor Market: The Complete Picture of Gross Worker Flows, Board of Governors of the Federal Reserve System mimeo

[14] Fujita, Shigeru; Ramey, Garey (2006): The Cyclicality of Job Loss and Hiring, Federal Reserve Bank of Philadelphia mimeo

[15] Gregg, Paul; Petrongolo, Barbara (2005): Stock-Flow Matching and the Performance of the Labor Market, European Economic Review 49, 1987-2011

[16] Jackman, Richard; Layard, Richard; Pissarides, Christopher A. (1989): On Vacancies, Oxford Bulletin of Economics and Statistics 51, 377-394

[17] Jolivet, Grégory (2009): A Longitudinal Analysis of Search Frictions and Matching in the U.S. Labor Market, Labour Economics 16, 121-134

[18] Kuhn, Peter; Mansour, Hani (2011): Is Internet Job Search Still Ineffective?, IZA mimeo

[19] Lindeboom, Maarten; van Ours, Jan; Renes, Gusta (1994): Matching Employers and Workers: An Empirical Analysis on the Effectiveness of Search, Oxford Economic Papers 46, 45-67

[20] Mumford, Karen; Smith, Peter N. (1999): The Hiring Function Reconsidered: On Closing the Circle, Oxford Bulletin of Economics and Statistics 61, 343-364

[21] Nagypál, Éva (2008): Worker Reallocation over the Business Cycle: The Importance of Employer-to-Employer Transitions, Northwestern University mimeo

[22] Nagypál, Éva (2005): On the Extent of Job-to-Job Transitions, Northwestern University mimeo

[23] Nickell, Stephen; Nunziata, Luca; Ochel, Wolfgang (2005): Unemployment in the OECD since the 1960s. What do we know?, Economic Journal 115, 1-27

[24] Petrongolo, Barbara; Pissarides, Christopher A. (2001): Looking into the Black Box: A Survey of the Matching Function, Journal of Economic Literature 39, 390-431

[25] Pissarides, Christopher A. (2000): Equilibrium Unemployment Theory, Second edition, Cambridge, MA: MIT Press (first edition, 1990, Oxford: Blackwell).

[26] Poterba, James M.; Summers, Lawrence H. (1986): Reporting Errors and Labor Market Dynamics, Econometrica 54, 1319-1338

[27] Sedláček, Petr (2010): Match Effciency and the Cyclical Behavior of Job Finding Rates, University of Amsterdam mimeo

[28] Shimer, Robert (2007): Reassessing the Ins and Outs of Unemployment, University of Chicago mimeo

[29] Shimer, Robert (2005): The Cyclicality of Hires, Separations, and Job-to-Job Transitions, Federal Reserve Bank of St. Louis Review 87, 493-507

[30] Stevens, Margaret (2007): New Microfoundations for the Aggregate Matching Function, International Economic Review 48, 847-868

[31] Stops, Michael; Mazzoni, Thomas (2010): Matchingprozesse auf beruflichen Teilarbeitsmärkten, Jahrbücher für Nationalökonomie und Statistik 230, 287-312

[32] Sunde, Uwe (2007): Empirical Matching Functions: Searchers, Vacancies, and (Un-)biased Elasticities, Economica 74, 537-560

Recently published

No.	Author(s)	Title	Date
15/2011	Osiander, Ch. Steinke, J.	Street-level bureaucrats in der Arbeitsverwaltung: Dienstleistungsprozesse und reformierte Arbeitsvermittlung aus Sicht der Vermittler published in: Zeitschrift für Sozialreform, Vol. 57, No. 2 (2011), p. 149-173	7/11
16/2011	Oberschachtsiek, D. Scioch, P.	The outcome of coaching and training for self-employment	8/11
17/2011	Heyer, G. Koch, S. Stephan, G. Wolff, J.	Evaluation der aktiven Arbeitsmarktpolitik: Ein Sachstandsbericht für die Instrumentenreform 2011	9/11
18/2011	Garloff, A. Pohl, C. Schanne, N.	Do small labor market entry cohorts reduce unemployment?	9/11
19/2011	Engelmann, S.	Trade Liberalisation, technical change and skill specific unemployment	10/11
20/2011	Mendolicchio, C. Rhein, T.	The gender gap of returns on education across West European countries	10/11
21/2011	Zapfel, St. Promberger, M.	Gemeinschaft, Gesellschaft und soziale Sicherung: Überlegungen zu Genese und Wandel des modernen Wohlfahrtsstaats	10/11
22/2011	Hohendanner, C. Klemm, M. Promberger, M. Sowa, Fr.	Vom Ein-Euro-Jobber zum ‚regulären' Mitarbeiter?: Eine Mixed-Methods-Evaluation zu innerbetrieblichen Übergängen aus öffentlich geförderter in sozialversicherungspflichtige Beschäftigung	11/11
23/2011	Krug, G. Rebien, M.	Job search via social networks: A analysis of monetary and non-monetary returns for low-skilled unemployed	11/11
24/2011	Caliendo, M. Hogenacker, J. Künn, St. Wießner, F.	Alte Idee, neues Programm: Der Gründungszuschuss als Nachfolger von Überbrückungsgeld und Ich-AG	11/11
25/2011	Schmelzer, P.	Unemployment and occupational mobility at the beginning of employment career in Germany and the UK	12/11
1/2012	Stephani, J.	Wage growth and career patterns of German low-wage workers	1/12
2/2012	Grunow, D. Müller, D.	Kulturelle und strukturelle Faktoren bei der Rückkehr in den Beruf	2/12

As per: 28.02.2012

For a full list, consult the IAB website
http://www.iab.de/de/publikationen/discussionpaper.aspx

Imprint

IAB-Discussion Paper 3/2012

Editorial address
Institute for Employment Research
of the Federal Employment Agency
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Regina Stoll, Jutta Palm-Nowak

Technical completion
Jutta Sebald

Website
http://www.iab.de

Download of this Discussion Paper
http://doku.iab.de/discussionpapers/2012/dp0312.pdf

For further inquiries contact the author:

Friedrich Poeschel
Phone +49.911.179 3105
E-mail friedrich.poeschel@iab.de