New Estimates of the Matching Function[*]

Daniel Borowczyk-Martins
University of Bristol

Grégory Jolivet
University of Bristol

Fabien Postel-Vinay
University of Bristol,
Paris School of Economics,
CEPR, and IZA

April 2011 — PRELIMINARY AND INCOMPLETE

Abstract

We show that equilibrium matching models imply that standard estimates of the matching function elasticities are exposed to an endogeneity bias, which arises from endogenous search behavior on either side of the market. We offer an estimation method which, under certain assumptions, is immune from that bias. Application of our method to the estimation of a basic version of the matching function using aggregate U.S. data from the Job Openings and Labor Turnover Survey (JOLTS) suggests that the bias is quantitatively important.

Keywords:
JEL codes:

1   Introduction

The matching function is a modeling device designed to capture the process through which the supply and demand sides are brought together in a frictional market. In a labor market context, the matching function maps the stock of job seekers and the stock of vacant jobs at any given date into the number of jobs (or "matches" between vacant jobs and job seekers) formed at that date (Pissarides, 2000). The matching function is the centerpiece of countless quantitative contributions to the broad field of macro-labor, some aiming to explain aggregate fluctuations in hours, wages, and other macro variables, others aiming to evaluate some policy tool, others still focusing on the allocation of the workforce between different regions or industries...

All those quantitative contributions have to rely on values of the matching function elasticities with respect to the numbers of vacant jobs and job seekers. Those elasticities have been and continue to be the focus of a large body of empirical work, which continues to expand as better data on job vacancies become available. In this paper, we argue that existing estimates of the matching function elasticities are likely to be exposed to an endogeneity bias. We offer an estimation method which, under certain assumptions, is immune from that bias. We apply our method to the estimation of a very simple version of the matching function using aggregate U.S. data from the Job Openings and Labor Turnover Survey (JOLTS). Results suggest that the bias is quantitatively important. For example, under the (common) assumption of constant returns to scale, an OLS estimate of the aggregate matching function elasticity w.r.t. vacancies based on the JOLTS series available at the time of writing is around 0.82. Our proposed estimate is around 0.68.[1]

From a theoretical standpoint, the source of bias that we highlight is very straightforward. The matching function takes job vacancies as one of its inputs. Vacancies are posted by profit maximizing firms. The returns to posting a vacancy depend on the efficiency of the matching process. Therefore, random shocks to matching efficiency affect the number matches formed both directly through the matching technology and indirectly through firms' vacancy-posting behavior — very much like TFP shocks affect aggregate production both directly and indirectly through the demand for inputs. Hence, job vacancies are endogenous, and an estimation strategy consisting of, say, running OLS regressions of the number of new matches on measures of the numbers of job seekers and job vacancies (a common strategy in the literature) fails to account for that endogeneity. Yet, numbers based on such OLS estimates are routinely used to calibrate matching models in which labor demand is explicitly endogenous, typically through a free entry condition (firms post job vacancies as long as the expected value of doing so is positive).[2]

Perhaps surprisingly, the source of bias we identify in this paper has been largely ignored in

---

[1]So the bias is positive in this case. While that particular OLS estimator is very commonly used in the literature, other estimators have also been implemented, leading to different biases with different signs. We provide an extensive discussion of those different estimators below.

[2]Many of those papers revolve around the model developed by Mortensen and Pissarides (1994).

the matching function estimation literature, which has mainly been concerned with other potential sources of bias such as time aggregation or imperfect input measurement, or with fundamental specification issues (such as the relevance of stock-flow matching).[3] While we do recognize the importance of those various issues, we set them aside in this paper and focus on the source of endogeneity described above.

Recent papers have used lags of the matching function's inputs as instruments for their own current values. Depending on the assumptions made about the process of matching efficiency shocks, some of the resulting estimates may coincide with ours. However our focus is different from that of these papers as we highlight the role of the free entry condition, or more generally of endogenous search behavior on both sides of the market, as potential sources of simultaneity.[4] By explicitly modeling the response of labor demand to matching efficiency shocks, we make the source of endogeneity explicit which allows us to justify our instrumentation strategy within the structure of a general search and matching model.

The paper is organized as follows. In Section 2 we give a brief formal account of the endogeneity of vacancies using a stripped-down, standard labor-matching model. In Section 3 we show how to consistently estimate the matching function within the model of Section 2, imposing some structure on the matching efficiency shock. Section 4 gives a brief description of the data. Results are then set out in Section 5 and further discussed in Section 6. Section 7 concludes.

2 Statement of the Problem

2.1 A Simple Matching Model

Although in principle the argument that we make in this paper applies to any matching model, in order to make our point with minimal peripheral complication we shall focus on the simplest — and perhaps most widely considered — case of an aggregate matching function $m(\cdot)$ that summarizes the number of matches formed between *unemployed* job seekers and vacant jobs. Specifically, the number M of such matches formed in a given month is related to the number of unemployed workers, U, and job vacancies, V, at the beginning of that month, in the following way:

$$M = m(U, V) = AU^\delta V^\eta, \tag{1}$$

---

[3]See the surveys by Petrongolo and Pissarides (2001) and Yashiv (2007). Burdett, Coles and van Ours (1994) offer a very clear and insightful discussion of time aggregation in matching models. An empirical analysis of the time aggregation bias is conducted by Berman (1997). Anderson and Burgess (2000), Fahr and Sunde (2005) and Sunde (2007) quantify the bias arising from incomplete or imperfect input measurement. Gregg and Petrongolo (2005) and Coles and Petrongolo (2008) offer an empirical investigation of the stock-flow matching hypothesis, whereby the number of matches formed at any date is jointly determined by the stock of job seekers and the *inflow* of new job vacancies into the search market.

[4]Yashiv (2000) conducts a structural estimation of an equilibrium search and matching model. Sedláček (2010) analyzes the efficiency of the matching function while accounting for unobserved job vacancies. Lastly, Barnichon and Figura (2011) study the effect of unemployment composition and dispersion of labor market conditions on matching efficiency.

where to further fix ideas we follow the vast majority of empirical studies of the matching function in assuming a Cobb-Douglas functional form.[5] An important feature of (1) is the presence of a shifter, A, which has a random component capturing random shocks to the matching technology. Those shocks parallel TFP shocks hitting the aggregate production function: they can be interpreted as recruitment-sector specific productivity shocks caused, for example, by changes in ICT affecting the way jobs are advertised or applied for, or by policy shocks affecting the functioning of employment agencies, or by changes in the geographic mobility of the workforce... In this paper we will refer to A as the "reallocation shock", for want of a better term.

We further assume that the matching function exhibits Constant Returns to Scale (CRS) so that $\eta = 1 - \delta$.[6] In this case, and with random matching — whereby all job seekers (vacant jobs) have equal sampling probability — the matching function can be redefined in terms of a job seeker's *job finding rate*, F, as follows:

$$F = \frac{M}{U} = A\Theta^{\eta}, \tag{2}$$

where $\Theta = V/U$ is labor market tightness. The job finding rate is the probability for any unemployed worker to find a job in the current month.[7]

The standard matching model (see e.g. Pissarides, 2000) is closed by assuming free entry and exit of firms in the search market. While there are alternatives to the free entry assumption as a way to model labor demand (mostly involving some adjustment cost of vacancies), we choose to focus on the free entry assumption as it is used in the overwhelming majority of applications. Under free entry, firms post vacancies at a flow cost of C per month until profit opportunities from doing so are exhausted. Labor demand is then determined by the free entry condition:

$$C = \frac{M}{V} \cdot \Pi, \tag{3}$$

where Π is the present discounted value (PDV) of a filled and producing job in the typical firm. The interpretation of (3) is that employers equate the marginal flow cost of posting a vacancy (the constant C) to the expected marginal return of doing so, which equals the value of a filled job, Π, times the probability of filling the job, which from a firm's perspective equals M/V under random matching. Further note that, with CRS in matching, that probability is also a function of labor market tightness only as $M/V = F/\Theta = A\Theta^{\eta-1}$. Substituting into the free entry condition (3) yields:

$$\Theta^{1-\eta} = \frac{\Pi A}{C}. \tag{4}$$

---

[5]Implicit in (1) is the additional assumption that unemployed workers all look for jobs with the same fixed intensity. It is conceptually straightforward to extend our point to the case of endogenous search intensity.

[6]CRS is a theoretically desirable property for the aggregate matching function, and is indeed assumed in a vast majority of theoretical applications, as well as in many empirical studies of the matching function. Yet an important body of empirical literature has been concerned with testing the assumption of CRS.

[7]Note that, in a discrete time model as the one considered in this paper, a constraint should be added to (2) to ensure that F is always less than one. We follow conventional practice and ignore that constraint, assuming that A and Θ takes on values that are consistent with $F \leq 1$.

Given the number of unemployed job seekers U, firms post more vacancies if the PDV of employing a worker, Π, is higher, or if the cost of posting a vacancy, C, is lower, or if the efficiency of the matching technology, A, is currently higher.

2.2 Endogeneity of Labor Market Tightness

Taking logs in (2) and using lower case letters to denote logarithms, one obtains a convenient linear relationship between f and θ:

$$f = \eta\theta + a. \tag{5}$$

The focus of a large fraction of the empirical literature on the matching function — and that of this paper — is to obtain an estimate of η. The common approach to this problem is to use measures of f and θ to estimate (5) by OLS. If free entry holds, however, this approach will fail to deliver a consistent estimate of η as (4) clearly implies that θ is correlated with a.[8] Rewriting (4) in log terms yields:

$$\theta = \frac{\pi - c + a}{1 - \eta}, \tag{6}$$

so that $\text{Cov}\,(\theta, a) \neq 0$ in general. Intuitively, the reallocation shocks affects the job finding rate both directly by changing the efficiency of the matching process, and indirectly by affecting the employers' incentives to post vacant jobs. In spite of this potential source of bias originating from the free entry condition, estimates of η based on OLS regressions of f on θ are routinely used to calibrate matching models in which the free entry condition is assumed to hold.[9]

We now show how the endogeneity of θ can be overcome by imposing some structure on the process of reallocation shocks.

3 The Statistical Model

3.1 Specification

We propose to estimate the matching function using monthly time series observations of the job finding rate and labor market tightness. Introducing a time index t, which becomes necessary at this juncture, we now decompose the reallocation shock a_t as follows: $a_t = \mu + \tau_t + \epsilon_t$, where μ is a constant, τ_t is a seasonal dummy, and ϵ_t is an unobserved component. Rewriting equation (5), we obtain:

$$f_t = \mu + \eta\theta_t + \tau_t + \epsilon_t. \tag{7}$$

---

[8]Notable exceptions to the OLS-in-levels approach are discussed below. For example, some authors have estimated a first-differenced version of (5) by OLS. As we show below, this approach is also exposed to a simultaneity bias.

[9]As briefly mentioned in the Introduction, many papers in the matching function estimation literature have addressed potential simultaneity biases originating from measurement problems, as well as temporal aggregation biases. The problem we address in this paper is clearly distinct and, in principle, cannot be solved by recourse to better or higher frequency data.

We further assume that the stochastic component of the reallocation shock ϵ_t follows an AR(p) process:

$$\epsilon_t = \sum_{\ell=1}^{p} \rho_\ell \epsilon_{t-\ell} + \omega_t \quad \Leftrightarrow \quad P(\mathbf{L}) \epsilon_t = \omega_t, \tag{8}$$

where $\mathbf{L}$ is the lag operator, $P(\mathbf{L}) := 1 - \sum_{\ell=1}^{p} \rho_\ell \mathbf{L}^\ell$, and ω_t is a serially uncorrelated disturbance.[10]

3.2 Estimation

Applying the transformation $P(\mathbf{L})$ to (7), using (8) and re-arranging we obtain our main equation of interest for the estimation:

$$f_t = \nu + \sum_{\ell=1}^{p} \rho_\ell f_{t-\ell} + \eta \theta_t - \sum_{\ell=1}^{p} \lambda_\ell \theta_{t-\ell} + P(\mathbf{L}) \tau_t + \omega_t, \tag{9}$$

where $\nu = \left(1 - \sum_{\ell=1}^{p} \rho_\ell\right) \mu$ and $\lambda_\ell = \eta \rho_\ell$ for all ℓ. With serially uncorrelated ω_t's, estimation of this latter model can be based on the moment conditions $\mathbf{E}(\omega_t \theta_{t-\ell}) = 0$ for all $\ell \geq 1$. While θ_t is still endogenous in (9) as a consequence of free entry, the structure imposed on the reallocation shock implies that its own lags are valid instruments for θ_t.[11] How strong those instruments are will depend on the amount of persistence in the various components of θ — see equation (6) — and will be assessed in the estimation.

In this simple case of serially uncorrelated ω_t's, only θ_t is endogenous in (9). Thus in principle we only need one excluded instrument for identification. Because all lags of θ_t are valid instruments, the model is overidentified.

4 Data

We take our measures of job vacancies and matches formed from the Job Openings and Labor Turnover Survey (JOLTS). JOLTS offers an aggregate time series of job openings and hires covering the U.S. non-farm sector starting December 2000 and ending in February 2011 at the time of writing. The "job openings" variable (our measure of vacancies) is a count of all positions that are open on the last business day of the month. The "hires" variable (our measure of matches formed) counts all additions to the payroll during the month.[12] Finally, we use data on the number of unemployed aged 16 or over from the Bureau of Labor Statistics (BLS).

The left panel on Figure 1 plots the non-seasonally adjusted time series of the job finding rate and labor market tightness, both in log terms. The shaded areas indicate NBER-dated recessions.

---

[10]Extending the method to allow for some degree of persistence in ω_t by assuming it to follow an MA(q) process is straightforward. However the data do not appear to call for this complication (see below).

[11]θ_t is the only endogenous regressor under the assumption of no serial correlation in ω_t. Had we assumed that ω_t followed an MA(q) process, we would have needed to instrument all lags of θ_t and f_t up to q and base estimation on $\mathbf{E}(\omega_t \theta_{t-\ell}) = 0$ holding for all $\ell \geq q + 1$. Our proposed estimation method remains valid, however, up to this simple modification.

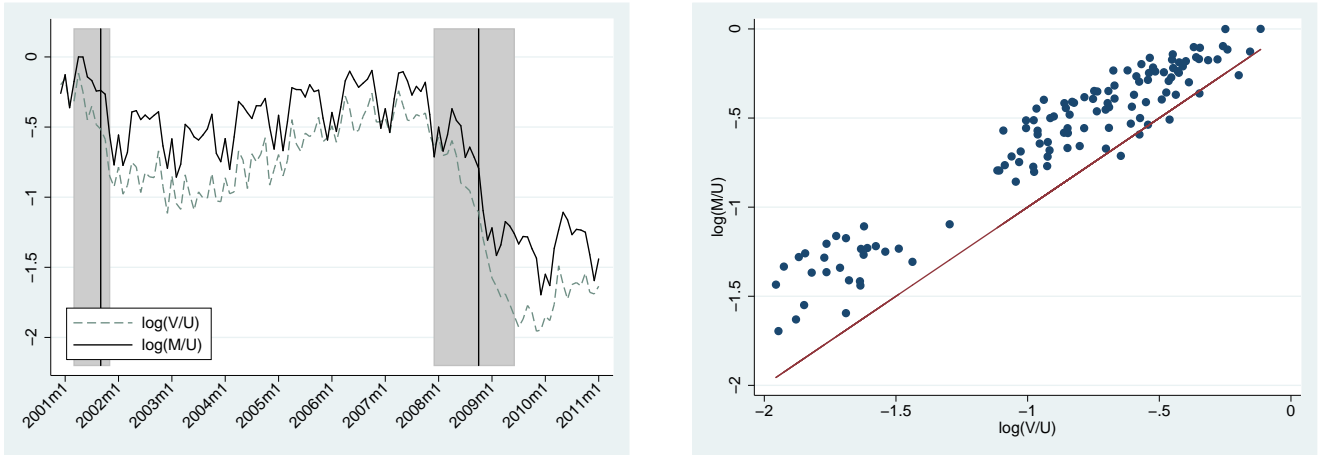[12]See www.bls.gov/jlt/ for details.

Figure 1: The job finding rate and labor market tightness

Apart from strong seasonality in both variables, the graph suggests the presence of two breaks occurring around the middle of the two recessions covered by the data (September 2001 and, more markedly, October 2008 — both breaks are materialized on the figure by vertical solid lines). The right panel of Figure 1 is a scatter plot of the (log) job finding rate against log tightness. Those two series co-vary very closely and the figure provides *prima facie* evidence of an affine relationship between f_t and θ_t. The right panel of Figure 1 also shows that labor market tightness is slightly more volatile than the job finding rate.

Dickey-Fuller tests do not reject the hypothesis of a unit root in both the job finding rate and labor market tightness. The possibly nonstationary nature of f_t and θ_t has led some authors to be concerned about the spuriousness of the correlation between those two variables.[13] As shown by Hsiao (1997), these concerns do not apply to our structural estimation, based on a 2SLS regression of equation (9). More specifically, if f_t and θ_t are nonstationary then OLS on (7) will yield a super-consistent estimator of a cointegrating vector for (f_t, θ_t) (see Phillips and Durlauf, 1986, or Stock, 1987). However, what we are after when estimating a matching function is *not* a cointegrating relation between f_t and θ_t. Rather, we are seeking to estimate the parameter of a *structural* relationship between f_t and θ_t (the matching function). In other words, at the true value of the matching function elasticity η, the residual ϵ_t in (7) may not be stationary (our estimation results will indeed show that we cannot reject nonstationarity). Hsiao (1997) shows that, in this context, 2SLS on (9) is consistent.[14]

---

[13]The standard strategy is then to consider a first-differenced version of equation (7) (see e.g. Yashiv, 2000).

[14]A few further subtleties arise here. The statement that the OLS estimator of the regression coefficient of a nonstationary variable y_t on a nonstationary vector $\mathbf{x}_t$ is super-consistent for a cointegrating vector for $(y_t, \mathbf{x}'_t)$ is only true if the elements of $\mathbf{x}_t$ are not themselves cointegrated. Strictly speaking, this fails to hold in our case as the r.h.s. in (7) comprises θ *and* month dummies, which are stationary. Following from that, note that if we take pseudo-differences and thus consider equation (9), we have a cointegration relation (ω is stationary by assumption) but the regressors are obviously cointegrated so OLS on (9) will also produce a biased estimate.

5 Estimation Results

Table 1 shows estimation results for our baseline specification of the reallocation shock process, which is to assume that ϵ_t follows an AR(3) process, three being the highest order which we find to be statistically significant in our data. All estimations are run on non-seasonally adjusted data and include month dummies to capture seasonality. Prior seasonal adjustment of the data would indeed create artificial serial correlation in all adjusted variables which would combine itself with the endogeneity issue we are tackling here. The results gathered in Table 1 include our preferred estimator (column 5), as well as for a number of benchmark specifications taken from the literature. We now go over those results.

Column 1 reports estimates obtained from OLS applied to equation (7). Our OLS estimate of η is on the high side of estimates previously obtained by other authors based on JOLTS data, probably owing to the combined facts that the JOLTS series now covers a longer period and has undergone a substantial revision in April 2011.[15] Hall (2005) finds an elasticity of 0.77 based on one year of JOLTS data (2002). Nagypál (2009) finds an elasticity of total hires with respect to vacancies (*not* imposing CRS) of 0.668 on seasonally adjusted data and 0.531 on non-seasonally adjusted data. (She rejects CRS in the latter case.) Nagypál's sample stops in November 2004. Rogerson and Shimer (2010) find an elasticity of 0.42 (imposing CRS) on a sample going up to mid-2009, although they use MA-smoothed seasonally adjusted data in the regression.

For Column 2, we took first differences (FD) of (7) and then ran OLS. Some authors have advocated estimating (7) in first differences, based on the worry that OLS estimates from the model in levels might be spurious owing to the nonstationary nature of f_t and θ_t. The OLS estimate of η based on the first-differenced version of (7) is indeed about half of that from the model in levels. While at first blush this may reflect the spurious nature of the estimates in levels, our interpretation is that the difference in estimates between Columns 1 and 2 simply reflect different biases.[16]

We next implement the estimator succinctly described and used by Yashiv (2000) in an effort to "cater for nonstationarity [...] and endogeneity". Yashiv's estimator is again based on first differences of (7): it consists of a 2SLS regression of Δf_t on $\Delta \theta_t$, where $\Delta \theta_t$ is instrumented by lags of $\Delta \theta_t$ of order 2 and above. Note that this estimator coincide with our preferred estimator (see Section 3) if ϵ_t follows a random walk. The point estimate reported in Column 3 is markedly higher than the OLS estimate from the model in FD (Column 2), although not as high as the OLS estimate from the model in levels. The discrepancy between the OLS and IV estimates on model (7) in FD should arouse suspicion as to the consistency of OLS. Moreover, a Sargan test rejects the consistency of the set of instruments used in Column 3.

---

[15]See www.bls.gov/jlt/ for details.

[16]Asymptotically (and ignoring month dummies), the bias in Column 1 converges to $\mathrm{Cov}\,(\theta, a)\,/\mathrm{Var}\,(\theta)$, while the bias in Column 2 converges to $\mathrm{Cov}\,(\Delta\theta, \Delta a)\,/\mathrm{Var}\,(\Delta\theta)$. Those expressions are impossible to sign in general as they depend on the dynamic structure of the reallocation shock a and on its correlation with π and c (see Section 2).

Column 4 reports OLS estimates of (9) and is there for comparison with Column 5 (our proposed estimator) to gauge the extent of the simultaneity bias arising from correlation between θ_t and the contemporaneous innovation ω_t. Estimates using our proposed strategy are reported in Column 5: equation (9) is estimated by 2SLS with θ_t instrumented by lags of $\Delta\theta_t$ ($\Delta\theta_{t-4}$ to $\Delta\theta_{t-7}$). While our preferred estimate of the matching function elasticity is closer to the simple OLS estimate based on the (5) in levels (Column 1) than to the one based on OLS in first differences (Column 2), it is still markedly lower, meaning that the simultaneity bias affecting OLS estimates has a positive sign.

Consistency of the set of instruments that we use to produce the results in Column 5 is accepted by a Sargan test (reported at the bottom of Column 5). The sizeable differences between IV and OLS estimates of (9) (Column 4) as well as the reasonably large first-stage partial R-squared and F statistic all further suggest that we are not confronted with a weak instrument problem. Most importantly, estimation of (9) by OLS produces a substantially lower value of η, suggesting that simultaneity is indeed there and quantitatively important. Finally, the common factor restriction (CFR) $\lambda_\ell = \eta\rho_\ell$ was not imposed in our estimation of (9). The restriction can be tested using a Wald test, which is also reported in Table 1 and does not reject the restriction.[17]

Based on the results shown in Column 5, the hypothesis that $\sum_{\ell=1}^{3}\rho_\ell = 1$ cannot be rejected, i.e. the reallocation shock ϵ_t may be nonstationary.[18] This restriction can be imposed in (9), which leads to estimating the following model (with $p = 3$):

$$\Delta f_t = -\left(1 - \rho_1\right)\Delta f_{t-1} - \left(1 - \rho_1 - \rho_2\right)\Delta f_{t-2}$$
$$+ \eta\Delta\theta_t + \eta\left(1 - \rho_1\right)\Delta\theta_{t-1} + \eta\left(1 - \rho_1 - \rho_2\right)\Delta\theta_{t-2}$$
$$+ P\left(\mathbf{L}\right)\tau_t + \omega_t.$$

We estimate this equation by 2SLS, not imposing the common factor restriction, and report the results in Column 6. As expected, point estimates are very close to those obtained in Column 5 (with a slight gain in precision) and the common factor restriction is not rejected.

We investigate the robustness of the results to different assumptions about the amount of persistence in the reallocation shock ϵ_t, to the inclusion of structural breaks in September 2001 and October 2008, and to restricting the estimation sample to a different sub-period. Results are gathered in Table 2.[19] We note that the conventional OLS estimate of model (7) in levels is sensitive both to the inclusion of a structural break at the beginning of the 2008 Recession and to the inclusion of the post-2008 Recession period in the estimation sample. Our IV estimate is

---

[17]Because the CFR is a nonlinear restriction, the form under which we test it may matter for the Wald test (Gregory and Veall, 1986). Here we test equality to zero of $\lambda_\ell - \eta\rho_\ell$. Using other forms (e.g. $\lambda_\ell/\eta - \rho_\ell = 0$) leads to the same qualitative conclusion.

[18]Note that, consistently with (9), the constant term is estimated at 0 in Column 4.

[19]The table only reports a subset of the coefficients. A full table is available on request.

	1 OLS on (7)	**2** OLS on (7) in FD	**3** IV on (7) in FD	**4** OLS on (9)	**5** IV on (9)	**6** IV on (9) in FD
θ_t	0.824 (.015)	0.369 (.063)	0.557 (.104)	0.494 (.060)	0.675 (.116)	0.709 (.080)
f_{t-1}	—	—	—	0.411 (.099)	0.294 (.114)	0.334 (.064)
f_{t-2}	—	—	—	0.269 (.104)	0.291 (.100)	0.306 (.100)
f_{t-3}	—	—	—	0.336 (.102)	0.388 (.102)	—
θ_{t-1}	—	—	—	-0.150 (.079)	-0.242 (.091)	-0.300 (.091)
θ_{t-2}	—	—	—	-0.036 (.081)	-0.067 (.079)	-0.062 (.072)
θ_{t-3}	—	—	—	-0.318 (.068)	-0.338 (.066)	—
constant	0.043 (.028)	—	—	0.027 (.034)	-0.009 (.038)	—
Obs.	119	119	119	119	119	119
Sargan test (df : stat : p-value)	—	—	$2 : 7.6 : 0.02$	—	$3 : 1.9 : 0.59$	$3 : 1.9 : 0.58$
Wald test of CFR (df : stat : p-value)	—	—	—	—	$3 : 3.1 : 0.38$	$2 : 4.3 : 0.12$
First stage:						
Partial R^2	—	—	0.36	—	0.25	0.42
F test (df : stat : p-value)	—	—	$(3, 105) : 19.8 : 0.0$	—	$(4, 97) : 8.1 : 0.0$	$(4, 100) : 17.9 : 0.0$

Month dummies included as regressors in all specifications. Excluded instruments in column 3 are $\Delta\theta_{t-\ell}$ for $\ell = 2$ to 5 (following Yashiv, 2000). Excluded instruments in column 5 are $\Delta\theta_{t-\ell}$ for $\ell = 4$ to 7. Excluded instruments in column 6 are $\Delta\theta_{t-\ell}$ for $\ell = 3$ to 6.

Table 1: Estimation results — baseline specification

far more robust to these changes especially when considering our preferred specification of the residuals, AR(3).

	1 OLS on (7)	**2** OLS on (7) in FD	**3** IV on (9) assuming ϵ_t is... AR(1)	**4** AR(2)	**5** AR(3)
Whole sample (2000:m12 - 2011:m1)					
θ_t	0.824 (.015)	0.369 (.063)	0.612 (.117)	0.672 (.105)	0.675 (.116)
f_{t-1}	—	—	0.749 (.079)	0.430 (.107)	0.294 (.114)
f_{t-2}	—	—	—	0.441 (.102)	0.291 (.100)
f_{t-3}	—	—	—	—	0.388 (.102)
Whole sample, with structural breaks 2001:m9 and 2008:m10					
θ_t	0.623 (.027)	0.371 (.064)	0.610 (.112)	0.660 (.103)	0.711 (.129)
f_{t-1}	—	—	0.561 (.089)	0.341 (.110)	0.212 (.122)
f_{t-2}	—	—	—	0.406 (.102)	0.312 (.101)
f_{t-3}	—	—	—	—	0.332 (.100)
Period up to 2008:m9					
θ_t	0.646 (.026)	0.362 (.074)	0.722 (.187)	0.722 (.181)	0.677 (.167)
f_{t-1}	—	—	0.538 (.151)	0.322 (.166)	0.264 (.159)
f_{t-2}	—	—	—	0.414 (.166)	0.322 (.121)
f_{t-3}	—	—	—	—	0.233 (.120)

Month dummies included as regressors in all specifications. Structural breaks at 2001:m9 and 2008:m10 included in bottom two panels. Excluded instruments in columns 2-4 are $\Delta\theta_{t-p-\ell}$ for $\ell = 1$ to 3 and p is the order of the autoregressive component of ϵ.

Table 2: Robustness checks

6 Further Quantitative Assessment of the Bias

Whether unemployment fluctuations are caused by fluctuations in labor demand induced by shocks to firms' revenues (induced by productivity or demand shocks), or by shifts in the matching function itself (i.e., in the parlance of equation (2), by shifts in A), is a classic question in macro-labor

(dating back to at least Blanchard and Diamond, 1989).[20] The standard approach to addressing that question is to pick a matching function elasticity $\widehat{\eta}$ (usually obtained from calibration or OLS regression), back out the implied time series of reallocation shocks as $a_t = f_t - \widehat{\eta}\theta_t$, and look at the cyclical behavior of that shock. In this section, we gauge the quantitative consequences of using OLS estimates of the matching function elasticity for that exercise, as opposed to using our preferred, consistent IV estimates.
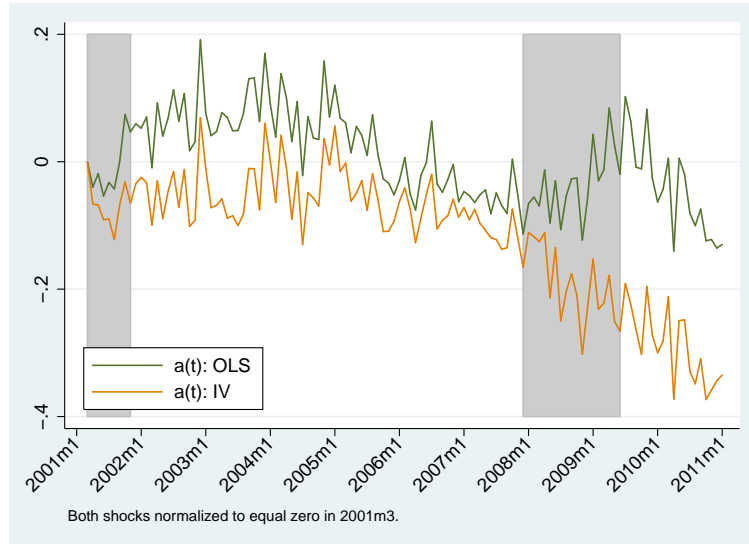

Figure 2: Matching efficiency

Figure 2 plots time series of the reallocation shock constructed from standard OLS estimates (Table 1, Column 1) and from IV estimates (Table 1, Column 5). The OLS-based matching efficiency shows no particular trend and is somewhat *countercyclical* (it has a correlation of 0.31 with detrended unemployment). The IV-based matching efficiency, on the other hand, does not appear to follow any cyclical pattern, but declines steadily since the second semester of 2007. Those two time series suggest radically different conclusions as to the causes of unemployment fluctuations. The OLS estimates suggests that, if anything, matching efficiency *improves* in periods of high unemployment, so that based on the model, high unemployment can only be due to depressed labor demand caused by adverse productivity or demand shocks. The IV estimates, on the other hand, suggests that matching efficiency has started to slowly and steadily deteriorate round the mid 2000s, which may partly account for the slow recovery that the U.S. labor market is experiencing at the time of writing.

---

[20]Of course a further source of confusion resides in the fact that, if labor demand is determined by a free entry condition, then any shock to matching efficiency A will induce a simultaneous response of firms' labor demand, in addition to its direct impact on the matching function.

7 Conclusion

[IN PROGRESS]

References

[1] Anderson, P. M. and S. M. Burgess (2000) "Empirical Matching Functions: Estimation and Interpretation Using State-level Data", *Review of Economics and Statistics*, 82, 90-102.

[2] Barnichon, R. and A. Figura (2011) "What Drives Matching Efficiency? A Tale of Composition and Dispersion", mimeo.

[3] Berman, E. (1997) "Help Wanted, Job Needed: Estimates of a Matching Function from Employment Service Data", *Journal of Labor Economics*, 15(1), S251-92.

[4] Blanchard, O. J. and P. A. Diamond (1989) "The Beveridge Curve", *Brookings Papers on Economic Activity*, 1989(1), 1-76.

[5] Burdett, K., M. G. Coles and J. Van Ours (1994) "Temporal Aggregation Bias in Stock-flow Models", CEPR Discussion Paper No. 967.

[6] Coles, M. G. and B. Petrongolo (2002) "A Test Between Unemployment Theories Using Matching Data", *International Economic Review*, 49, 1113-41.

[7] Fahr, R. and U. Sunde (2005) "Job and Vacancy Competition in Empirical Matching Functions", *Labour Economics*, 12, 773-80.

[8] Gregg, P. and B. Petrongolo (2005) "Stock-flow Matching and the Performance of the Labour Market", *European Economic Review*, 49, 1987-2011.

[9] Gregory, A. W. and M. R. Veall (1986) "Wald Tests of Common Factor Restrictions", *Economics Letters*, 22, 201-8.

[10] Hall, R. E. (2005) "Employment Fluctuations with Equilibrium Wage Stickiness", *American Economic Review*, 95(1), 50-65.

[11] Hsiao, C. (1997), "Statistical Properties of the Two-Stage Least Squares Estimator Under Cointegration", *Review of Economic Studies*, 64, 385-98.

[12] Mortensen, D. T. and C. A. Pissarides (1994), "Job Creation and Job Destruction in the Theory of Unemployment", *Review of Economic Studies*, 61(3), 397-415.

[13] Nagypál, E. (2009) "What Can We Learn About Firm Recruitment from the Job Openings and Labor Turnover Survey?", in: T. Dunne, J. B. Jensen, and M. J. Roberts, editors, *Producer Dynamics: New Evidence from Micro Data*, University of Chicago Press.

[14] Petrongolo, B. and C. A. Pissarides (2001) "Looking into the Black Box: An Empirical Investigation of the Matching Function", *Journal of Economics Literature*, 39, 390-431.

[15] Phillips, P. C. B. and S. N. Durlauf (1986), "Multiple Time Series Regression with Integrated Processes", *Review of Economic Studies*, 53(4), 473-95.

[16] Pissarides, C. A. (2000) *Equilibrium Unemployment Theory*, Cambridge, MA: the MIT Press.

[17] Rogerson, R. and R. Shimer (2010) "Search in Macroeconomic Models of the Labor Market", in: O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 4A, 619-700.

[18] Sedláček, P. (2010) "Match efficiency and the cyclical behavior of job finding rates", mimeo.

[19] Stock, J.H. (1987) "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", *Econometrica*, 55(5), 1035-56.

[20] Sunde, U. (2007) "Empirical Matching Functions: Searchers, Vacancies, and (Un-)biased Elasticities", *Economica*, 74, 537-60.

[21] Yashiv, E. (2000), "The Determinants of Equilibrium Unemployment", *American Economic Review*, 90(5), 1297-1322.

[22] Yashiv, E. (2007), "Labor Search and Matching in Macroeconomics", *European Economic Review*, 51(8), 1859-95.