Review of Economic Dynamics 16 (2013) 440–451

Contents lists available at SciVerse ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

Accounting for endogeneity in matching function estimation ☆

Daniel Borowczyk-Martins [a], Grégory Jolivet [a], Fabien Postel-Vinay [a,b,c,d,*]

[a] Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, United Kingdom
[b] Sciences Po, Paris, France
[c] CEPR, United Kingdom
[d] IZA, Germany

A R T I C L E   I N F O

Article history:
Received 26 August 2011
Revised 7 July 2012
Available online 14 July 2012

JEL classification:
J63
J64
E24

Keywords:
Matching function estimation
Unemployment
Vacancies
Job finding

A B S T R A C T

We show that equilibrium matching models imply that standard estimates of the matching function elasticities are exposed to an endogeneity bias, which arises from the search behavior of agents on either side of the market. We offer an estimation method which, under certain structural assumptions about the process driving shocks to matching efficiency, 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 can be quantitatively important.

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 flow 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 macroeconomic variables, others aiming to evaluate some policy, others still focusing on the allocation of the workforce between different regions or industries. In most of those models, the matching function is a key parameter to assess constrained efficiency of labor market equilibria — and therefore the scope for policy intervention.[1]

All such quantitative analysis has 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 keeps expanding as better and more abundant data on job vacancies become available. In this paper, we argue that

☆ We thank the editor, Gianluca Violante, and two anonymous referees for constructive comments. We also thank Simon Burgess, Juan Dolado, Jon Temple, Frank Windmeijer, David Winter, Eran Yashiv and seminar/conference audiences at Bristol, Essex, Marseille, the 2011 IZA Summer School and the 2011 SED meetings for feedback on early versions of this paper. We take full responsibility for any error or shortcoming in this paper. Daniel Borowczyk-Martins acknowledges financial support from Fundação para a Ciência e a Tecnologia under grant SFRH/BD/38968/2007, co-funded by the European Social Fund.
* Corresponding author at: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK. Fax: +44(0)117 928 8577.
E-mail address: Fabien.Postel-Vinay@bristol.ac.uk (F. Postel-Vinay).

[1] Under constant returns to scale in the matching function and ruling out employed job search, constrained efficiency obtains when the employer's share of the match surplus is equal to the elasticity of the job finding rate w.r.t. labor market tightness (the so-called Hosios condition).

1094-2025/$ – see front matter
http://dx.doi.org/10.1016/j.red.2012.07.001

existing estimates of the matching function elasticities are exposed to an endogeneity bias arising from the search behavior of agents on either side of the market. 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).

Our 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 0.84. Our proposed estimate is around 0.7.[2]

From a theoretical standpoint, the source of bias that we highlight is 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 of matches formed both directly through the matching technology and indirectly through firms' vacancy-posting behavior — very much like total factor productivity (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.[3] Yet, numbers based on such OLS estimates are routinely used to calibrate matching models that typically feature a free entry condition (firms post job vacancies as long as the expected value of doing so is positive), which explicitly makes vacancies a function of current matching efficiency.[4]

Perhaps surprisingly, the source of bias we identify in this paper has been largely ignored in 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).[5] 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.

Our estimation strategy consists of imposing an ARMA structure on matching efficiency, transforming the model using pseudo differences to account for the autoregressive component of matching efficiency, and then estimating the transformed equation by the generalized method of moments (GMM), using lags of the labor market tightness and/or the job finding rate as instrumental variables. We implement this procedure for a wide range of ARMA processes and propose some explicit, practical criteria for specification selection.

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 one or both sides of the market, as potential sources of simultaneity.[6] By 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 describes the data. Results are then presented in Section 5. In Section 6, we discuss the specification of matching efficiency shocks, we relax the standard assumption of constant returns to scale in the matching function and we investigate the role of aggregate recessions in creating the endogeneity bias studied in this paper. 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 determines the number of matches formed between *unemployed* job seekers and

[2] 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.

[3] A similar problem would arise on the supply side if job seekers chose their search effort according to the efficiency of the matching process. Our approach can be adapted to that case.

[4] Most of those papers build upon the Mortensen and Pissarides (1994) model.

[5] See the surveys by Petrongolo and Pissarides (2001) and Yashiv (2007). Burdett et al. (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), Sunde (2007) and Jolivet (2009) 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.

[6] Yashiv (2000) conducts a structural estimation of an equilibrium search and matching model. Sedláček (2011) 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.

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:

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

where to further fix ideas we follow the vast majority of empirical studies of the matching function in assuming a Cobb–Douglas functional form.[7] 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 information and communication technologies 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. A is variously referred to in the literature as a 'reallocation shock', a 'measure of mismatch' or one of 'matching efficiency'. We will use the latter option in this paper.

We further assume that the matching function exhibits Constant Returns to Scale (CRS) so that $\eta = 1 - \delta$.[8] In this case, and with random matching — where 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.[9]

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:

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

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 part 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 will fail to produce a consistent estimate of η as (4) clearly implies that θ is correlated with a.[10] Rewriting (4) in log terms yields:

---

[7] 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.

[8] 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. We briefly investigate the consequences of relaxing the CRS assumption in Section 6.2.

[9] 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 take on values that are consistent with $F \leqslant 1$.

[10] 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 FDOLS estimator is also exposed to a simultaneity bias.

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

so that $\mathrm{Cov}(\theta, a) \neq 0$ in general. Intuitively, matching efficiency 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.

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 now show how the endogeneity of θ can be overcome by imposing some structure on the process of the matching efficiency shock.

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 matching efficiency 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 Eq. (5), we obtain:

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

We further assume that the stochastic component of matching efficiency ϵ_t follows an ARMA(p,q) process:

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

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

3.2. Estimation

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

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

with the common factor restriction:

$$\nu = \left(1 - \sum_{\ell=1}^{p} \rho_\ell\right)\mu \quad \text{and} \quad \forall \ell\colon \lambda_\ell = \eta\rho_\ell. \tag{10}$$

Estimation of this latter model can be based on the moment conditions $\mathbf{E}(\omega_t\theta_{t-\ell}) = 0$ and $\mathbf{E}(\omega_t f_{t-\ell}) = 0$ for all $\ell \geqslant q+1$. While $(\theta_t, \ldots, \theta_{t-\min\{p,q\}})$ and, if $q \geqslant 1$, $(f_{t-1}, \ldots, f_{t-\min\{p,q\}})$ are endogenous in (9) as a consequence of free entry and the persistence in the MA(q) component of ϵ_t, $Q(\mathbf{L})\omega_t$, the structure imposed on the matching efficiency shock implies that lags of θ_t (and/or f_t) beyond order $q+1$ are valid instruments for $(\theta_t, \ldots, \theta_{t-\min\{p,q\}})$ and $(f_{t-1}, \ldots, f_{t-\min\{p,q\}})$. How strong those instruments are will depend on the amount of persistence in the various components of θ — see Eq. (6) — and will be assessed in the estimation.

The number of coefficients to be estimated is $p+1$ (the elasticity η and the p autoregressive coefficients ρ_ℓ), so that in principle only $p+1$ moment conditions are needed for identification.[11] Because all lags of θ_t (and/or f_t) beyond order $q+1$ are valid instruments, the model is potentially overidentified.

Estimation of (7) finally requires a choice of values for (p,q), for which we adopt the following method. First, we estimate (7) for all values of (p,q) in a grid, $(p,q) \in \{1, \ldots, \bar{p}\} \times \{0, \ldots, \bar{q}\}$. Next, we select the highest value of p for which we estimate statistically significant autoregressive coefficients (ρ_ℓ) up to order p. Finally, we select q by inspection of the autocorrelogram of the residual $Q(\mathbf{L})\omega_t$ from (7), estimated with the value of p previously selected. While informal, this model selection procedure has the twofold advantage of being intuitive and straightforward to implement. Estimation of (7) over a whole grid of values of (p,q) further allows us to assess the sensitivity of our estimate of η, the matching function elasticity, to different assumptions about the persistence of the matching efficiency shock.

[11] This count ignores the constant and the eleven month dummies, which are exogenous and included in the set of instruments.

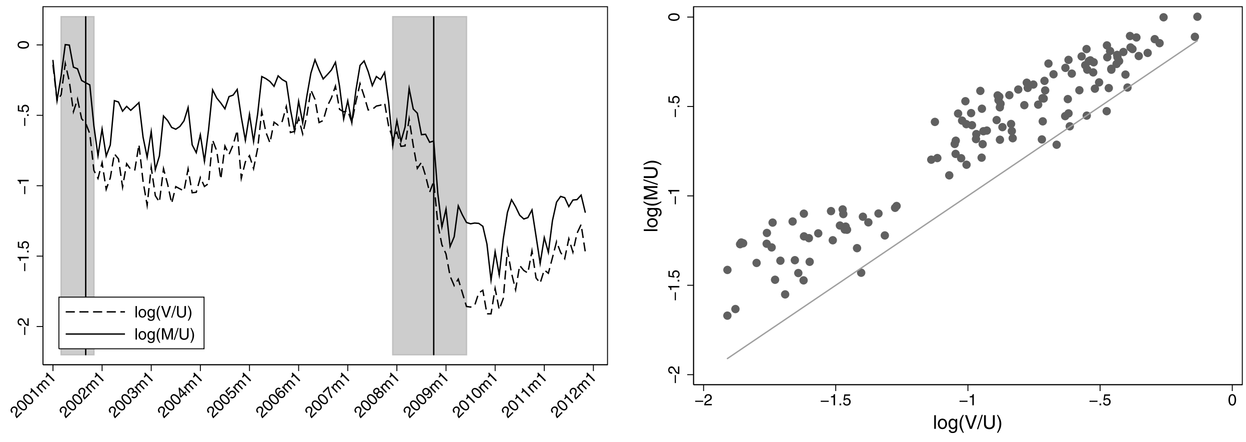


Fig. 1. The (logarithm of) job finding rate and labor market tightness.

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 January 2012 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 in Fig. 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. 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). We discuss the econometric contents of those breaks in some detail in Section 6.3. The right panel of Fig. 1 is a scatter plot of the job finding rate against labor market tightness, both in log terms. Those two series co-vary very closely and the figure provides *prima facie* evidence of an affine relationship between f_t and θ_t.

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 GMM estimation of Eq. (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).[14] 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, GMM on (9) is consistent.

5. Estimation results

Our baseline specification of the matching efficiency process assumes that ϵ_t follows an ARMA(3, 3) process. Our choice of values for (p, q) was guided by the protocol described in Section 3: three is the highest autoregression order which we find to be statistically significant in our data, and, given $p = 3$, inspection of the autocorrelogram of the residual from (9) suggests an MA(3) structure. (We provide a more complete discussion of our choice of specification in the next section.) 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 are shown in Table 1. Columns 1–3 show estimates obtained using benchmark specifications taken from the literature and are displayed for comparison purposes. Our preferred estimator, accounting for endogeneity of labor market tightness and for the dynamics of efficiency shocks, is in columns 4 and 5 and places the elasticity at about 0.7 over the period under consideration. We now go over those results in detail.

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

[13] The standard strategy is then to consider a first-differenced version of Eq. (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.

Table 1
Estimation results — baseline specification.

	1 OLS on (7)	2 OLS on (7) in FD	3 IV on (7) in FD	4 GMM on (9) (just identified)	5 GMM on (9) (with overidentification)
η	0.842 (0.013)	0.331 (0.070)	0.412 (0.107)	0.706 (0.095)	0.692 (0.062)
ρ_1	–	–	–	0.517 (0.383)	0.461 (0.294)
ρ_2	–	–	–	−0.057 (0.270)	−0.038 (0.256)
ρ_3	–	–	–	0.568 (0.290)	0.591 (0.270)
μ	0.042 (0.027)	–	–	0.073 (0.474)	0.231 (1.025)
Obs.	133	132	129	129	129
Sargan test (df : stat : p-value)	–	–	3 : 14.62 : 0.002	–	1 : 0.056 : 0.813

Month dummies included as regressors in all specifications. Instruments in column 3 are $\theta_{t-\ell}$ for $\ell = 2$ to $\ell = 5$ (following Yashiv, 2000). Excluded instruments in column 4 are $\theta_{t-\ell}$ for $\ell = q + 1 = 4$ to $\ell = q + p + 1 = 7$. Instruments in column 5 are the same as in column 4, plus $f_{t-\ell}$ for $\ell = q + 1 = 4$.

Column 1 reports estimates obtained from OLS applied to Eq. (7). Our OLS point estimate of η, equal to 0.842, 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 point estimate of η based on the first-differenced version of (7), equal to 0.332, is indeed less than 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 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 θ_t of order 2 and above. Note that this estimator would coincide with our preferred estimator if ϵ_t followed a random walk. The point estimate reported in column 3 is markedly higher than the OLS estimate from the model in first differences (0.412 vs. 0.332), although still about only half as high as the OLS estimate from the model in levels (0.842). 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.

Estimates using our proposed strategy are reported in columns 4 and 5. As per the estimation protocol laid out in Section 3, Eq. (9) is estimated by two-step GMM with $(\theta_t, \dots, \theta_{t-3})$ and $(f_{t-1}, \dots, f_{t-3})$ instrumented by lags of θ_t, imposing the common factor restriction (10). Column 4 reports estimates from the just-identified case, where exactly $p + 1$ instruments are used to identify the $p + 1$ parameters $(\eta, \rho_1, \dots, \rho_p)$, which, in the $(p, q) = (3, 3)$ case, means using $(\theta_{t-4}, \dots, \theta_{t-7})$ as instruments. The resulting point estimate of the matching function elasticity is reasonably precise, and still markedly lower (0.706 vs 0.842) than the simple OLS estimate based on (5) in levels (column 1), indicating that the simultaneity bias affecting OLS estimates has a positive sign. The autoregression coefficients $(\rho_1, \dots, \rho_3)$ are estimated with somewhat disappointing precision, with only ρ_3 being statistically significant in column 4.

In order to gain some precision, we exploit the overidentified nature of the model. As noted in Section 3, all lags of θ_t and f_t beyond order $q + 1$ are valid instruments. Column 5 reports GMM estimates based on the same 'minimal' vector of instruments as in column 4, $(\theta_{t-4}, \dots, \theta_{t-7})$, supplemented by one additional instrument, f_{t-q-1}.[17] Point estimates are virtually the same as in the just-identified case, and precision is slightly improved (ρ_1 is now borderline statistically significant). Finally, consistency of our set of instruments is overwhelmingly accepted by a Sargan test (reported at the bottom of column 5).

[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 matching efficiency shock a and on its correlation with π and c (see Section 2).

[17] Formally, estimation is based on the moment conditions $\mathbf{E}(\omega_t \theta_{t-\ell}) = 0$ for all $\ell \in \{q + 1, \dots, q + p + 1\}$ and $\mathbf{E}(\omega_t f_{t-q-1}) = 0$.

Table 2
Test of the common factor restriction (10).

θ_t [η]	f_{t-1} [ρ_1]	f_{t-2} [ρ_2]	f_{t-3} [ρ_3]	θ_{t-1} [$\eta\rho_1$]	θ_{t-2} [$\eta\rho_2$]	θ_{t-3} [$\eta\rho_3$]	Const.
0.685	0.386	0.048	0.593	−0.304	0.039	−0.439	−0.031
(0.309)	(0.494)	(0.457)	(0.610)	(0.270)	(0.326)	(0.169)	(0.201)
Sargan test: df = 1, stat = 1.505, p-value = 0.220							
Wald test of common factor restriction: df = 3, stat = 0.03, p-value = 0.998							

Month dummies included as regressors. Excluded instruments are $\theta_{t-\ell}$ for $\ell = q+1$ to $q+2p+1$ and $\theta_{t-\ell}$ for $\ell = q+1$.

Before we move on to a further assessment of the robustness of our results to alternative specifications, a few remarks are in order. First, in principle we could have included many more lags of both θ_t and f_t into the set of instruments. However, because of the persistence in θ_t and the high correlation between f_t and θ_t, those additional instruments would not add much information and only contribute to bias estimates toward least squares. Application of (nonlinear) least squares to (9) — still imposing the common factor restriction (10) — produces an estimate of η of 0.595, which is markedly different from the value of 0.706 produced by GMM.[18] However we have observed in experiments not reported here that adding too many consecutive lags of θ_t and/or f_t to the set of instruments brings the GMM estimate of η close to the NLS value of 0.595.[19]

Second, based on the results shown in column 5, the hypothesis that $\sum_{\ell=1}^{3} \rho_\ell = 1$ cannot be rejected, meaning that the matching efficiency shock ϵ_t may be nonstationary. This restriction can be imposed in (9), which leads to a similar equation except that all the f's and θ's are taken in first differences (and summations only go up to $p-1$ lags). We estimate this equation by two-step GMM, imposing the common factor restriction (10), and obtain an estimate of 0.652 for η (s.e. of 0.079), close to the one reported in columns 4 and 5.

Lastly, we should check the validity of the common factor restriction (10) and assess how much identification power it carries. To that end, we estimate (9) *not* imposing the restriction (10).[20] Results for the $(p,q) = (3,3)$ specification are reported in the top panel of Table 2. Point estimates thus obtained are very close to the corresponding estimates obtained when (10) is imposed (see columns 4 and 5 in Table 3). While condition (10) does not drive the point estimates, it seems however to play an important part in improving the precision of our estimates. The common factor restriction (10) can further be tested using a Wald test, which is also reported in Table 2.[21] It is overwhelmingly accepted.

6. Discussion

In Section 6.1 we discuss model selection based on the amount of persistence in matching efficiency ϵ_t. In Section 6.2, we attempt to relax the assumption of constant returns to scale in the matching function. Finally, in Section 6.3 we discuss the nature of matching efficiency shocks and the possible use of an alternate estimator for η, namely OLS including structural breaks in September 2001 and October 2008.

6.1. Model selection

Table 3 contains estimates of the model parameters under the maintained assumption that matching efficiency ϵ_t follows an ARMA(p,q) process for all values of (p,q) in the grid $\{1,\dots,4\} \times \{0,\dots,6\}$. In each case, we expand the 'minimal' set of instruments by adding f_{t-q-1} to provide some overidentification, as we did in column 5 of Table 1. Table 3 suggests a number of remarks.[22]

First, $p = 3$ is the highest AR order for which we were able to find a statistically significant coefficient ρ_p. Table 3 only reports estimation results for values of p up to $p = 4$ because of space constraints, but results for higher values of p are available on request, and confirm that the ρ_ℓ's are never found to be statistically significant beyond $\ell = 3$. Following our model selection protocol (see Section 3), we choose $p = 3$ as our preferred specification.

Second, while the point estimate of the matching function elasticity η is somewhat variable across specifications for low values of p, it becomes reasonably stable (around 0.68) as soon as one allows for an autoregression order of at least three (once again, estimates obtained under higher values of p are available on request). The estimate of $\eta = 0.692$ under our preferred specification of $(p,q) = (3,3)$ seems 'typical' in the set of results displayed in Table 3.

[18] Note that NLS applied to (9) can only be consistent for η if ω_t is serially uncorrelated ($q = 0$) and uncorrelated with θ_t, which would be the case if, for example, it took some time for employers to adjust vacancies, so that the free-entry condition (3) applied with a lag of at least one period.

[19] For example, using 24 lags of both θ_t and f_t produces an estimate of 0.62 for η. Other examples are available on request.

[20] Note that, if (10) is not imposed, (9) becomes linear.

[21] 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.

[22] Note that in theory the point estimates may or may not be constant across (p,q) specifications. Indeed identification, and thus the choice of the instrument set, is driven by the structure imposed on the shock ϵ. Values of (p,q) too far below the true ones will not produce estimates fully cleansed of the endogeneity bias whereas values of (p,q) which are too high may lead to imprecise estimates, potentially getting closer to the NLS ones (cf. discussion in Section 5).

Table 3
Different model specifications.

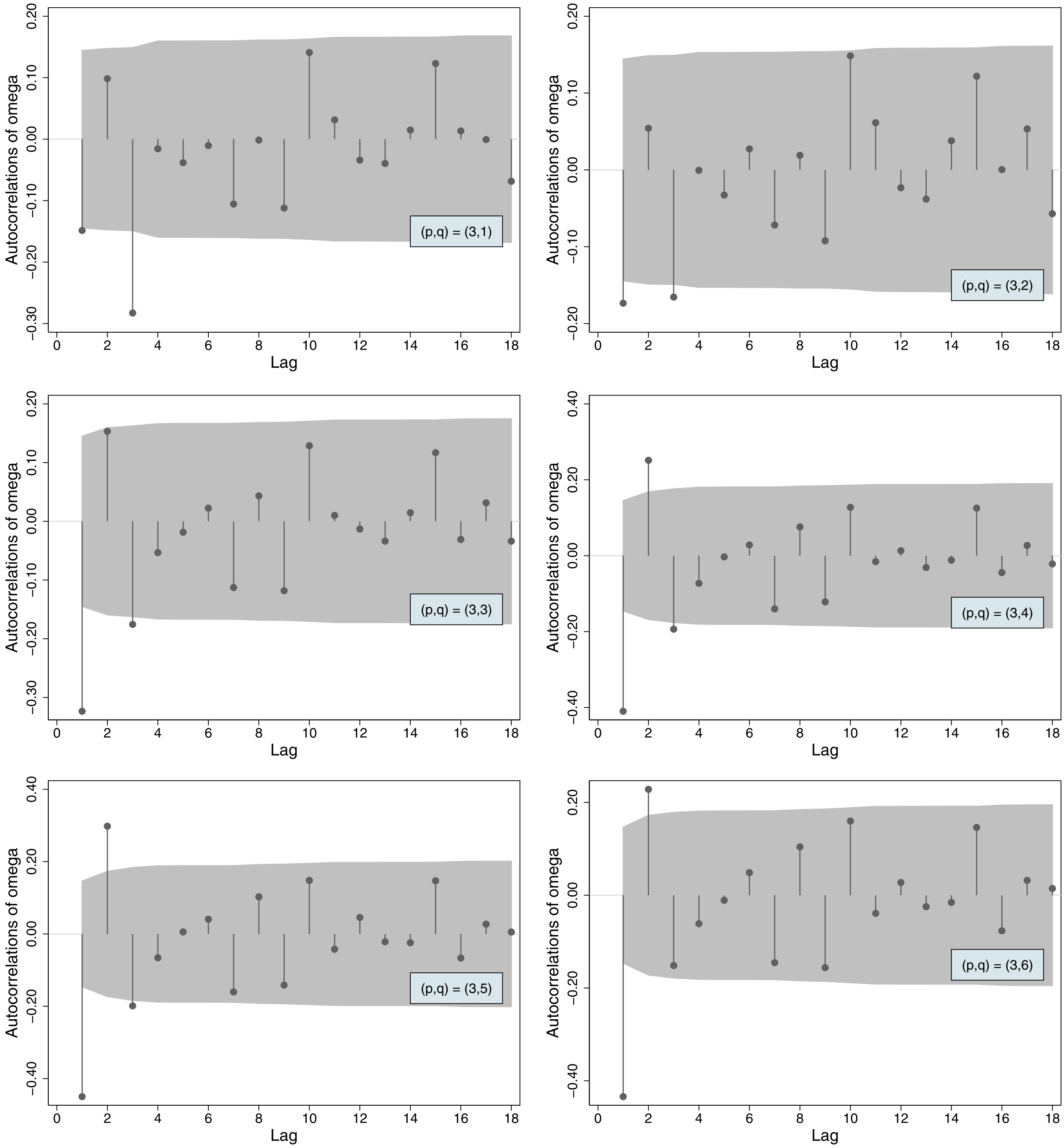
(p,q)	μ	η	ρ_1	ρ_2	ρ_3	ρ_4	Sargan
(1, 0)	0.069	0.866	0.706				0.082 : 0.774
	(0.045)	(0.038)	(0.063)				
(1, 1)	2.987	2.986	0.982				0.037 : 0.848
	(4.106)	(2.714)	(0.017)				
(1, 2)	−1.215	0.338	0.993				4.622 : 0.032
	(2.158)	(0.126)	(0.017)				
(1, 3)	−0.763	0.529	0.993				0.042 : 0.839
	(2.222)	(0.182)	(0.028)				
(1, 4)	0.326	1.107	0.926				0.349 : 0.554
	(0.367)	(0.320)	(0.065)				
(1, 5)	−0.030	0.695	1.031				0.459 : 0.498
	(0.210)	(0.165)	(0.062)				
(1, 6)	−0.682	0.558	0.993				1.133 : 0.287
	(2.126)	(0.209)	(0.033)				
(2, 0)	0.074	0.870	0.712	−0.009			5.448 : 0.020
	(0.046)	(0.038)	(0.116)	(0.144)			
(2, 1)	0.083	0.880	−0.442	1.009			5.756 : 0.016
	(0.043)	(0.034)	(0.544)	(0.392)			
(2, 2)	1.671	0.493	0.678	0.325			5.353 : 0.021
	(17.153)	(0.117)	(0.296)	(0.301)			
(2, 3)	0.115	0.910	1.026	−0.201			0.706 : 0.401
	(0.092)	(0.081)	(0.270)	(0.298)			
(2, 4)	0.133	0.930	−0.455	1.191			0.154 : 0.695
	(0.076)	(0.066)	(1.069)	(0.921)			
(2, 5)	0.025	0.612	1.018	−0.006			1.627 : 0.202
	(0.746)	(0.216)	(0.697)	(0.716)			
(2, 6)	1.636	0.632	0.891	0.111			1.819 : 0.177
	(46.735)	(0.160)	(0.661)	(0.676)			
(3, 0)	1.877	0.688	0.225	0.033	0.745		0.214 : 0.644
	(29.130)	(0.046)	(0.127)	(0.136)	(0.147)		
(3, 1)	0.092	0.698	0.268	0.036	0.726		0.156 : 0.692
	(0.299)	(0.049)	(0.198)	(0.156)	(0.141)		
(3, 2)	−0.763	0.643	0.348	0.101	0.542		0.296 : 0.587
	(2.335)	(0.059)	(0.208)	(0.233)	(0.161)		
(3, 3)	0.231	0.692	0.461	−0.038	0.591		0.056 : 0.813
	(1.025)	(0.062)	(0.294)	(0.256)	(0.270)		
(3, 4)	0.096	0.684	0.525	−0.177	0.677		0.533 : 0.465
	(0.393)	(0.066)	(0.473)	(0.478)	(0.285)		
(3, 5)	0.156	0.672	0.546	−0.274	0.746		0.004 : 0.950
	(0.748)	(0.067)	(0.410)	(0.523)	(0.471)		
(3, 6)	0.191	0.666	0.580	−0.173	0.608		0.316 : 0.574
	(0.945)	(0.076)	(0.592)	(0.557)	(0.433)		
(4, 0)	0.270	0.684	0.144	0.028	0.741	0.104	0.005 : 0.942
	(1.023)	(0.039)	(0.198)	(0.134)	(0.141)	(0.225)	
(4, 1)	0.031	0.820	1.339	−0.244	0.669	−0.687	0.004 : 0.950
	(0.173)	(0.152)	(0.660)	(0.279)	(0.198)	(0.428)	
(4, 2)	0.358	0.676	−1.073	0.475	0.604	1.023	0.001 : 0.980
	(1.254)	(0.036)	(2.117)	(0.687)	(0.249)	(1.542)	
(4, 3)	1.398	0.668	−0.312	0.226	0.675	0.417	0.420 : 0.517
	(11.804)	(0.038)	(1.165)	(0.476)	(0.276)	(0.698)	
(4, 4)	0.106	0.684	0.167	−0.033	0.655	0.245	0.545 : 0.461
	(0.284)	(0.043)	(1.051)	(0.469)	(0.269)	(0.760)	
(4, 5)	0.098	0.671	0.422	−0.079	0.559	0.123	0.423 : 0.515
	(0.381)	(0.057)	(1.150)	(0.632)	(0.404)	(0.860)	
(4, 6)	0.052	0.834	1.508	−0.683	0.906	−0.713	0.020 : 0.887
	(0.607)	(0.414)	(0.824)	(0.924)	(0.886)	(0.650)	

Month dummies included as regressors in all specifications. Instruments are $\theta_{t-\ell}$ for $\ell = q+1$ to $\ell = p+q+1$ and $f_{t-\ell}$ for $\ell = q+1$. The 'Sargan' column reports the Sargan test statistic and p-value.

Third, while estimates of the autoregressive coefficients ρ_ℓ are slightly less stable, the unit-root hypothesis $\sum_{\ell=1}^{p}\rho_\ell = 1$ is never rejected based on a simple Student test (not reported here, available on request).[23] In other words, matching efficiency shocks are consistently found to be very persistent.

The next step is to select a value of q. Again following our proposed model selection protocol, we base our choice of $q = 3$ on an inspection of the autocorrelograms of the residual of (9) for $p = 3$ and all values of $q \in \{0, \ldots, 6\}$, shown in

[23] One manifestation of this is the occasionally very poor precision with which the constant, μ, is estimated in the second column of Table 3. This is because, as is apparent in (10), the estimating Eq. (9) only depends on $\nu = (1 - \sum_{\ell=1}^{p}\rho_\ell)\mu$, which is zero if $\sum_{\ell=1}^{p}\rho_\ell = 1$.



Shaded areas indicate 90% confidence bands.

Fig. 2. Autocorrelograms of estimated residuals from (7).

Fig. 2, where we look for 'consistent' specifications in the sense that the residual should exhibit no statistically significant autocorrelation beyond order q. It appears in Fig. 2 that only specifications with $q \geqslant 3$ are consistent in that sense. Autocorrelograms constructed using higher values of q suggest the same thing. For parsimony, we thus adopt the minimum value of q in the consistent set, $q = 3$.

6.2. Non-constant returns to scale

Estimation has so far been conducted under the maintained assumption of a constant-return to scale (CRS) matching function. While, as mentioned before, CRS is a theoretically appealing property for an aggregate matching function to have,

Table 4
Relaxing CRS.

	1 OLS on (7)	**2** GMM on (12) (with overidentification)
η	0.164	0.736
	(0.031)	(0.374)
$\tilde{\delta}$	−1.118	0.117
	(0.051)	(0.793)
ρ_1	–	0.367
		(0.264)
ρ_2	–	−0.010
		(0.245)
ρ_3	–	0.661
		(0.248)
μ	9.685	−0.813
	(0.442)	(6.743)
Sargan test (df : stat : p-value)	–	1 : 0.775 : 0.379

Month dummies included as regressors in all specifications. Instruments in column 2 are $\theta_{t-\ell}$ for $\ell = q+1 = 4$ to $\ell = q+p+2 = 8$, plus $f_{t-\ell}$ for $\ell = q+1 = 4$.

it is natural to investigate its empirical validity. The main problem one has to confront when relaxing the CRS assumption is that it is difficult to separately identify the elasticities of the matching function w.r.t. unemployment and vacancies, since those two variables are highly correlated. This is especially true when using a relatively short sample such as JOLTS, which only covers a dozen years.

Going back to the basic specification of the matching function (1), we thus relax the assumption that $\eta = 1 - \delta$. Taking logs in (1), one obtains:

$$m = a + \eta v + \delta u \quad \Leftrightarrow \quad f = a + \eta\theta + (\delta + \eta - 1)u. \tag{11}$$

Defining $\tilde{\delta} = \delta + \eta - 1$, simply testing $\tilde{\delta} = 0$ then provides a straightforward test of CRS.

To obtain consistent estimates of η and $\tilde{\delta}$ under our specification of the matching efficiency shock a_t, we proceed again by applying the $P(\mathbf{L})$ transform to (11):

$$f_t = \left(1 - \sum_{\ell=1}^{p} \rho_\ell\right)\mu + \sum_{\ell=1}^{p} \rho_\ell f_{t-\ell} + \eta\left(\theta_t - \sum_{\ell=1}^{p} \rho_\ell \theta_{t-\ell}\right) + \tilde{\delta}\left(u_t - \sum_{\ell=1}^{p} \rho_\ell u_{t-\ell}\right) + P(\mathbf{L})\tau_t + Q(\mathbf{L})\omega_{t-\ell}. \tag{12}$$

As in the CRS case, this can be estimated based on moment conditions of the form $\mathbf{E}(\omega_t\theta_{t-\ell}) = 0$, $\mathbf{E}(\omega_t f_{t-\ell}) = 0$, or $\mathbf{E}(\omega_t u_{t-\ell}) = 0$, for all $\ell \geqslant q+1$. Note that, compared to the CRS case, we now have one extra parameter to estimate (namely, $\tilde{\delta}$), and therefore require one extra moment condition for identification.

Results obtained under the specification of ϵ_t following an ARMA(3, 3) are reported in Table 4, which also shows OLS estimates of η and $\tilde{\delta}$ for comparison. Focusing on GMM estimates (column 2), we see that, as expected, a substantial amount of precision is lost by relaxing CRS ($\tilde{\delta}$, in particular, is very imprecisely estimated). However, three conclusions can be drawn from the results in Table 4. First, at the low level of precision at which $\tilde{\delta}$ is estimated, CRS cannot be rejected. Second, the point estimates of η and the ρ_ℓ's obtained without assuming CRS are very close indeed to our benchmark results (obtained under the CRS assumption) shown in Table 1. Third, OLS estimates suggest a very low elasticity of the matching function w.r.t. vacancies ($\eta = 0.164$) and a *negative* elasticity w.r.t. unemployment ($\delta = -0.282$), which does not seem credible.

6.3. Structural breaks

As we discussed in Section 4, inspection of the raw series of f_t and θ_t plotted in Fig. 1 suggests the presence of two breaks occurring around September 2001 and, more markedly, October 2008. This observation in turn prompts the following possible interpretation about the nature of matching efficiency shocks: one could think of ϵ_t as resulting from the combination of rare, large, persistent shocks (the breaks observed in Fig. 1), and smaller, higher-frequency shocks.[24] Under that interpretation, one can further argue that occurrences of the big, low-frequency shocks are (approximately) observed in the data: they coincide with the breaks that one can locate by simply eye-balling the data as we did in Section 4. Assuming that those large shocks are permanent shocks, one can then control for their occurrences by simply adding post-break dummies in the right-hand side of our equation of interest (7).[25]

[24] We are grateful to Gianluca Violante for suggesting this interpretation, and indeed prompting the discussion in this section.

[25] A peripheral, yet important question when one tries to explicitly capture the time variation in matching efficiency is that of the specification of the time trend. We use post-break dummies in this section because it is consistent (in the sense of being a special case of) our preferred ARMA specification.

Table 5
OLS with structural breaks.

	Imposing CRS			Not imposing CRS		
	1 OLS	**2** OLS with struct. breaks	**3** GMM	**4** OLS	**5** OLS with struct. breaks	**6** GMM
η	0.842	0.583	0.692	0.164	0.215	0.736
	(0.013)	(0.023)	(0.062)	(0.031)	(0.042)	(0.374)
$\tilde{\delta}$	–	–	–	−1.118	−0.942	0.117
				(0.051)	(0.111)	(0.793)

Month dummies included as regressors in all specifications. Structural breaks placed in September 2001 and October 2008. GMM with CRS uses instruments as indicated in Table 1, column 5. GMM without CRS uses instruments as indicated in Table 4.

We interpreted the discrepancy between the OLS and GMM estimates of the matching function elasticity reported in Table 1 as a consequence of the fact that the job finding rate f_t and labor market tightness θ_t are simultaneously affected by the matching efficiency shock. A natural question to ask is whether that simultaneity bias is still present when one controls for 'observed' large shocks in the way just described. Again, under the interpretation of shocks outlined above, one might think that the response of labor market tightness to high-frequency (small) matching efficiency shocks is negligible compared to its response to low-frequency, large shocks.

The left panel in Table 5 (columns 1–3) shows estimates of η obtained under the assumption of a CRS matching function from simple OLS (column 1), OLS including indicators of structural breaks in September 2001 and October 2008 (column 2), and from our proposed GMM estimator (column 3; columns 1 and 3 are taken up from Table 1). The right panel in Table 5 (columns 4–6) repeats the same comparison without imposing CRS. Estimates of η and $\tilde{\delta}$ are reported in that case (columns 4 and 6 are taken up from Table 4).

Clearly, the inclusion of those structural breaks in the set of regressors has a large impact on the OLS estimate of η. First, looking at the right panel of Table 5, it appears that, if one does not impose CRS, the only credible estimates in the set reported are the ones obtained from GMM, as both OLS estimators predict a negative elasticity of the matching function w.r.t. unemployment. As discussed in the previous section, GMM seems robust to relaxing the CRS assumption.

Next focusing on the (perhaps more standard) CRS case, we see that the inclusion of structural breaks as regressors brings the OLS estimator slightly closer to, although now markedly below, our GMM estimate. Focusing on point estimates, OLS with breaks produces an estimate of 0.583 for η, while GMM produces 0.692. Whether that difference is quantitatively important or not is open to question — the answer probably varies between applications. Yet the fact that we find such a sizeable difference based on our JOLTS sample, which only has 133 monthly observations and straddles the single deepest recession since World War Two (a shock of unusually large magnitude), leads us to conclude that the idea that the endogeneity of labor market tightness in matching function estimation is mostly driven by low-frequency, large and persistent shocks that are observable, seems somewhat fragile.

Whether that idea is true or not, however, we want to advocate our GMM estimator and model selection protocol as a systematic and robust way of addressing the endogeneity of θ_t in matching function estimation: not only is it robust to relaxing the assumption of constant returns to scale, but it also obviates the need to 'detect' structural breaks by eye-balling the data prior to estimation.

7. Conclusion

This paper begins by pointing out a simple implication of equilibrium matching models: the search behavior of firms and/or job seekers implies that labor market tightness and the job finding rate are simultaneously determined as functions of the unobserved efficiency of the matching process. As a consequence, the standard practice of regressing the job finding rate on a measure of labor market tightness using, e.g., OLS, is exposed to a simultaneity bias. Imposing some structure on the process followed by matching efficiency allows us to offer a consistent estimator of the matching function elasticity. Application of our method to the estimation of a basic version of the matching function using JOLTS data suggests that the bias has potentially important quantitative consequences.

In order to make our point with minimal peripheral complication, we have focused on a very basic version of the equilibrium matching model, and deliberately abstracted from a number of important problems analyzed elsewhere in the literature (such as time aggregation, imperfect input measurement or stock-flow matching). Further work is needed to examine how those sources of bias interact with the 'structural' problem of endogeneity of labor market tightness that we emphasize in this paper.

Yet, obviously, other options exist: for example, Şahin et al. (2012) use a quartic trend instead of post-break dummies and find an elasticity of $\eta = 0.654$, very close to our own GMM estimate, on a JOLTS sample that stops in December 2010. Using polynomial trends on our extended data set produces similar, although somewhat unstable results: under CRS, point estimates for η using OLS with a polynomial trend start at 0.741 with a linear trend and decrease steadily as the order of the polynomial trend is increased (down to 0.695 for a quadratic trend, 0.65 for a trend of order 6, and 0.53 for order 10…). Relaxing the CRS assumption, estimates become very unstable, with the elasticity w.r.t. unemployment generally estimated a negative number. Details are available on request.

Another question is that of the practical quantitative importance of the structural bias studied in this paper, which can be measured by the extent to which it affects the quantitative predictions of various prominent equilibrium search and matching models. For instance the Mortensen and Pissarides (1994) model has been used extensively to predict the effects of labor market policy instruments such as firing costs, hiring subsidies, taxes or minimum wages. These predictions usually follow from calibrations based on estimates of the matching elasticity that potentially suffer from the endogeneity bias we studied in this paper. It would be interesting to see how sensitive those predictions are to variations in the value of the matching function elasticity.

References

Anderson, P.M., Burgess, S.M., 2000. Empirical matching functions: estimation and interpretation using state-level data. Review of Economics and Statistics 82, 90–102.

Barnichon, R., Figura, A., 2011. Labor market heterogeneities, matching efficiency and the cyclical behavior of the job finding rate. Mimeo.

Berman, E., 1997. Help wanted, job needed: estimates of a matching function from employment service data. Journal of Labor Economics 15 (1), S251–S292.

Burdett, K., Coles, M.G., Van Ours, J., 1994. Temporal aggregation bias in stock-flow models. CEPR discussion paper No. 967.

Coles, M.G., Petrongolo, B., 2008. A test between unemployment theories using matching data. International Economic Review 49, 1113–1141.

Fahr, R., Sunde, U., 2005. Job and vacancy competition in empirical matching functions. Labour Economics 12, 773–780.

Gregg, P., Petrongolo, B., 2005. Stock-flow matching and the performance of the labour market. European Economic Review 49, 1987–2011.

Gregory, A.W., Veall, M.R., 1986. Wald tests of common factor restrictions. Economics Letters 22, 201–208.

Hall, R.E., 2005. Employment fluctuations with equilibrium wage stickiness. American Economic Review 95 (1), 50–65.

Hsiao, C., 1997. Statistical properties of the two-stage least squares estimator under cointegration. Review of Economic Studies 64, 385–398.

Jolivet, G., 2009. A longitudinal analysis of search frictions and matching in the U.S. labor market. Labour Economics 16, 121–134.

Mortensen, D.T., Pissarides, C.A., 1994. Job creation and job destruction in the theory of unemployment. Review of Economic Studies 61 (3), 397–415.

Nagypál, E., 2009. What can we learn about firm recruitment from the job openings and labor turnover survey? In: Dunne, T., Jensen, J.B., Roberts, M.J. (Eds.), Producer Dynamics: New Evidence from Micro Data. University of Chicago Press, Chicago.

Petrongolo, B., Pissarides, C.A., 2001. Looking into the black box: an empirical investigation of the matching function. Journal of Economic Literature 39, 390–431.

Phillips, P.C.B., Durlauf, S.N., 1986. Multiple time series regression with integrated processes. Review of Economic Studies 53 (4), 473–495.

Pissarides, C.A., 2000. Equilibrium Unemployment Theory. The MIT Press, Cambridge, MA.

Rogerson, R., Shimer, R., 2010. Search in macroeconomic models of the labor market. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 4A. Elsevier, Amsterdam, pp. 619–700.

Şahin, A., Song, J., Topa, G., Violante, G., 2012. Mismatch unemployment. New York University. Manuscript.

Sedláček, P., 2011. Match efficiency and the cyclical behavior of job finding rates. Mimeo.

Stock, J.H., 1987. Asymptotic properties of least squares estimators of cointegrating vectors. Econometrica 55 (5), 1035–1056.

Sunde, U., 2007. Empirical matching functions: searchers, vacancies, and (un)biased elasticities. Economica 74, 537–560.

Yashiv, E., 2000. The determinants of equilibrium unemployment. American Economic Review 90 (5), 1297–1322.

Yashiv, E., 2007. Labor search and matching in macroeconomics. European Economic Review 51 (8), 1859–1895.