UNEMPLOYMENT AND WORKER-FIRM MATCHING
IN POST-COMMUNIST ECONOMIES:
TRANSITION, POLICIES OR STRUCTURAL PROBLEMS?\(^1\)

Daniel Münich *
and
Jan Svejnar **

February 2006

JEL Number: P2, J4, J6, C33

\(^1\)In preparing the paper, the authors thank the EBRD & the Japan Europe Cooperation Fund for support. The authors would like to thank Elie Applebaum, Olivier Blanchard, Simon Commander, Robert Dunn, Randall Filer, Stepan Jurajda, Jan Kmenta, Janos Kollo, Miroslav Lizal, Dale Mortensen, Katherine Terrell, and participants of the William Davidson Institute Conferences on Labor Markets in Transition Economies for valuable comments. The usual disclaimer applies.

* CERGE-EI, Prague.
** University of Michigan; CERGE-EI, Prague.
IN POST-COMMUNIST ECONOMIES: TRANSITION, POLICIES OR STRUCTURAL PROBLEMS?

1. Introduction

More than fifteen years after the fall of the Berlin Wall, unemployment remains a major, and in many respects the most fundamental, problem in many of the post communist economies of the former Soviet bloc and Yugoslavia. From the academic and policy standpoint, the question that arises is whether high unemployment is the result of (a) ongoing (unfinished) transition from plan to market, (b) inappropriate macroeconomic policies and major external shocks, (c) problems related to the economic structures of these countries. The nature of appropriate policies for alleviating unemployment obviously depends on the answer to this question, with Hypothesis 1 (H1) implying the need to complete the transition process, Hypothesis 2 (H2) suggesting that macroeconomic policies are the key and Hypothsis 3 (H3) indicating that the policy should focus on labor market institutions and measures to stimulate labor mobility and create appropriate skills.²

In this paper, we use district-level panel data on the unemployed $U$, vacancies $V$, inflow $S$ into unemployment, and outflow $O$ from unemployment in the Czech Republic, Hungary, Poland, Slovakia, and Eastern and Western parts of Germany (hereafter East and West Germany) to provide evidence on the above question.³ In particular, we examine the three hypotheses in the context of

² By way of historical background, a fundamental systemic feature of the Soviet-type economies was the nonexistence of open unemployment. An equally distinguishing feature of the transition during the early-to-mid 1990s was the emergence of double digit unemployment rates in all the rapidly transforming economies except for the Czech Republic.

³ Apart from having the data needed for our analysis, these countries constitute an appropriate set of economies in which to examine these issues. In the Czech Republic, the unemployment rate remained at mere 3-4 percent throughout the (transformation) recession of first half of the 1990s and only rose to 6-9 percent during the second recession of 1997-99. In the 2000s, the unemployment rate remained in the very high 14-20 percent range in the rapidly growing economy of Slovakia (as well as Poland), and stabilized in the high 7-10 percent range in the moderately growing Czech Republic and Hungary. In Western and Eastern Germany, which we examine as
inflows into unemployment and the efficiency of matching of the unemployed and vacancies in the process of job creation. H1, namely that unemployment means that transition is still at work, is consistent with the observation that inflow $S$ (presumably from old, “socialist” jobs) is high, with outflow $O$ (job creation in the new “market” economy) being fine by some accepted norms. The manifestation of this situation would be that $U$ is high due to high inflow $S$.

H2 is that matching between the unemployed and vacancies is fine but that high unemployment is caused by low demand for labor (e.g., due to restrictive macroeconomic policies, overvalued exchange rate, or globalization shocks). The manifestation of this would be low $V$ relative to $S$, irrespective of $U$.

H3 is that high unemployment is caused by inefficient matching, brought about for example by inadequate labor market institutions or geographical or skill mismatch (see also Jurajda and Terrell, 2006). Here we would observe both $U$ and $V$ being high, but not necessarily in the same districts or skill groups.

In Table 1, we provide basic time series statistics on unemployment, inflow, outflow and vacancies in the six economies. As may be seen from the table, the six economies differ markedly in terms of their unemployment, flows and vacancy rates. West Germany (our benchmark economy) is

comparison economies, the unemployment rate has since the early 1990s fluctuated around 10 percent and … percent, respectively. An important part of the answer to the above questions is that from the time unemployment started appearing in CEE in the early 1990s, the Czech Republic has had a much higher outflow rate of individuals from the unemployment state to employment than did the other CEE economies (see e.g., Boeri, 1994, Boeri and Scarpetta, 1995 and Ham, Svejnar and Terrell (HST), 1998, 1999). Using OECD data, HST (1998), for instance report 1991 outflow rates to be 17.1 in the Czech Republic but only 4.8 in Slovakia. Similarly, their data for 1992 indicate the outflow rate to be 26.6 in the Czech Republic, 10.2 in Slovakia, 6.6 in Poland, and 4.3 in Hungary. Moreover, other possible causes of the less rapid rise of unemployment in the Czech Republic in the early 1990s, such as lower inflow rates into unemployment due to higher government subsidies to Czech firms or to greater declines in Czech labor force participation, are not borne out by the data. These basic findings suggest that one needs a better understanding of the determinants of outflow from unemployment and matching of the unemployed and vacancies in the Czech Republic and the other CEE countries. While HST (1998,1999) examine the outflow issue using individual unemployment duration data, in the present paper we analyze the process of matching using long
in the intermediate range, displaying between 1991 and 2005 an unemployment rate increasing from 5% to 10%, inflow as a share of the labor force rising from 0.9% to 1.5%, outflow as a share of the labor force fluctuating around 1-1.4%, and a vacancy rate (vacancies as a share of the labor force) fluctuating between 0.7% and 1.4%. East Germany, in contrast, registers an unemployment rate rising from 11.5% to 18.6%, inflow as a share of the labor force rate rising from about 1.5% to well above 2%, outflow as a share of the labor force rising to about 2-2.3% by the mid 1990s and fluctuating around this level ever since, and a vacancy rate rising from 0.4% in 1991 to about 1% in the late 1990s and remaining at that level in the 2000s. For most of the 1991-2005 period, the East German part of the German economy hence displays a much higher unemployment rate, accompanied by higher inflow, lower outflow rates (relative to the number of unemployed), and a similar vacancy rate. By 2005, the two economies converge somewhat in that the unemployment and inflow rates are higher in East Germany, but the outflow and vacancy rates are similar. The East German economy therefore operates with a higher unemployment rate in the presence of very sizable active labor market policies that unfortunately do not prevent high inflows into unemployment. Slovakia and Poland (for which we have only partial data so far) represent two transition economies that, like East Germany, operate with very high unemployment rates but, unlike East Germany, have not experienced an administratively set high wage level and cross-border subsidies. For most of the 1990s and 2000s, these two economies have experienced an unemployment rate in the 14% to 20% range, accompanied by relatively high inflow (1.2-1.3% of the labor force) and low outflow (1.0-1.4% of the labor force). In most years, they have also had vacancy rates significantly below 1%. The Czech Republic is an intermediate case, with unemployment rising from the low rate of 3-4% in the monthly panels of district-level data.
early to mid 1990s to 8-10% range since then. Its inflow and outflow as a share of the labor force have both risen from about 0.5-0.7% in the early-to-mid 1990s to a still relatively low level of 1.0-1.1% since then. Its vacancy rate has declined from 1.4-1.9% in the early-to-mid 1990s to 0.8-1.1% since then. Finally, Hungary has achieved the lowest level of unemployment. After reaching an unemployment rate of 10-12% on the mid-to-late 1990s, Hungary has succeeded to lower the rate to around 8% in the mid 2000s, reduced its inflow rate to 1.4%, raised the outflow rate to 14-16% and kept the vacancy rate at 1.0-1.1%. Hungary’s success is hence brought about by keeping the outflow rate relatively high and inflow rate relatively low.

In terms of our hypotheses, the raw data in Table 1 suggest that the West German economy behaves consistently with all three hypotheses. The unemployment rate has risen with increasing inflows (H1), the vacancy rate has declined while inflow has risen (H2) and both the unemployment and vacancy rates are relatively high (H3). The East German economy seems to conform to H1, having high unemployment and inflows (the latter being to a significant extent inflow from training programs), as well as H2 in that its vacancy rate is low relative to inflows into unemployment. Slovakia and Poland behave consistently with H1 (high unemployment and high inflows) and H2 (low vacancies relative to inflow) throughout the 1990s and 2000s, while the Czech economy starts with virtually no unemployment problem but increasingly conforms to H1 and H2. Finally, in part because it has relatively low unemployment, Hungary does not fit clearly into any of the three hypothesized scenarios. It has an element of all three hypotheses in that its inflow is relatively sizable (H1), the vacancy rate is low relative to inflow (H2) and unemployment and vacancies are relatively high (H3). In our analysis, we focus on identifying the extent to which the different countries exhibit different levels of (in)efficiency in matching.
The paper is structured as follows: We start in Section 2 by presenting our conceptual framework of a matching function model and a brief survey of the literature. In Section 3 we discuss our estimating framework and explain how we overcome some of the principal problems of the existing studies. In Section 4 we describe our data and the implementation of our econometric model. In Section 5 we present basic statistics and our econometric estimates. We conclude in Section 6.

2. Conceptual Framework and Existing Literature

The unemployment rate is determined by the inflow rate into and outflow rate from unemployment. In a steady state, inflow equals outflow so that $S = O$ and the steady state unemployment rate, $u$, is given by $u = U/LF = s/(s + o)$, where $s = S/E$ is the steady state inflow rate and $o = O/U$ is the steady state outflow rate. Assuming that $s$ is given by some exogenous job destruction, the only other determinant of the steady state unemployment rate is the matching function, $O = M(U,V)$, since $u = U/LF = s/[s + M(U,V)/U]$. The fact that the unemployment rate has remained high in many transition economies points to the need to understand the properties and shifts of the matching function in these economies.

The literature on matching functions usually assumes that outflow from unemployment to employment $O$ is a function of the number of unemployed $U$ and the number of posted vacancies $V$:

$$O = M(U,V)$$  \hspace{1cm} (1)

Matching is seen as a search process, with both the unemployed and employers with vacant positions striving to find the best acceptable match, given exogenous factors such as skill and spatial mismatch, as well as costly access to information. Some authors (e.g., Blanchard and Diamond, 1989, Pissarides, 1990, and Storer, 1994) expect the matching function $M$ to display constant returns
to scale, while others have identified reasons such as externalities in the search process, heterogeneity in the unemployed and vacancies and lags between matching and hiring, why increasing returns may prevail (e.g., Diamond, 1982, Coles and Smith, 1994, Profit, 1996, and Mortensen 1997). Increasing returns are conceptually important because in many models they constitute a necessary condition for multiple equilibria and provide a rationale for government intervention. In this paper we show that increasing returns appear to be an important phenomenon, especially in the later (1997-2003) than the earlier (1993-96) period, and that it is more pronounced in some of the economies than others.

In view of the serious unemployment problem in the transition economies, the literature on the matching of unemployed and vacancies in these economies has grown rapidly. It has also produced contradictory results, in part because the studies use different methodologies and data. Methodologically, the studies differ especially with respect to the specification of the matching function and treatment of returns to scale, the inclusion in equation (1) of other variables that might affect outflows and the extent to which they use static or dynamic models. In terms of data, the studies differ in whether they use annual, quarterly or monthly panels of district-level or more aggregate (regional) data and whether they cover short or long time periods. None of the studies adjusts the data for the varying size of the unit of observation (district or region) which, as we show presently, may generate biased estimates of the returns to scale in many studies.4

---

Unlike other studies, we use a more up-to-date empirical methodology and superior data. In particular, unlike other studies we a) control for the endogeneity of explanatory variables, b) account for the presence of a spurious scale effect introduced by the varying size across units of observation (districts), and c) use long panels of comparable monthly data from all districts in the countries that we analyze. Unlike most studies, we also employ both static and dynamic specifications and estimate on contiguous panels to allow for dynamic adjustment and regime changes. Like other studies, we do not address the issue of the matching of vacancies with employed individuals (job-to-job mobility).

3. The Estimating Framework

Theories of search and matching generally do not imply a particular functional form of the matching function. Like most studies, we use the Cobb-Douglas function which may be written in a deterministic form as

\[ \ln O_{i,t} = \alpha \ln U_{i,t-1} + \beta \ln V_{i,t-1} + a \]  

(2)

where, \( U_{i,t-1} \) and \( V_{i,t-1} \) are the number of unemployed and vacancies in district \( i \) at the end of period \( t-1 \), respectively, \( O_{i,t} \) denotes the outflow to jobs during period \( t \) (the number of successful matches

---

5 There are of course exceptions. Pissarides (1990) for instance shows that in his theoretical model the Cobb-Douglas function could represent a useful approximation. In the empirical work, Boeri (1994) estimates a Cobb-Douglas matching function of unemployment and vacancies, with unemployment entering as a CES function of short
between the currently unemployed and current vacancies) and constant $a$ captures the efficiency of matching. Using lowercase letters for logarithms of variables and introducing unobserved (time invariant) district specific effects $a_i$ and an idiosyncratic error term $e$, we can write (2) as

$$o_{i,t} = \alpha u_{i,t} + \beta v_{i,t-1} + a_i + \epsilon_{i,t-1}.$$  \hspace{1cm} (3)

In estimating (3), one has to take into account the specific features of the matching model. Estimating by the ordinary least squares (OLS) method is not appropriate if the unobserved district specific effects $a_i$s are correlated with explanatory variables $u$ and $v$. This correlation is likely to exist on account of structural differences between districts caused by factors which affect $a_i$, $u$ and $v$. One particular factor is district size, although this factor is observable and its impact can be eliminated by adjusting all variables in equation (3) by some measure of district size (we discuss the removal of the spurious size effect below).

If panel data are available, as in our case, suitable within transformations of (3) can be used to remove the unobserved $a_i$s. The most common within transformations are (i) deviations from district specific means (fixed effects) and (ii) first differences. Both transformations remove the $a_i$s, but the mean deviations transformation is not suitable if the model contains regressors that are only weakly exogenous. This is the case for matching functions because the explanatory variables (unemployment and vacancies) are predetermined by previous matching processes through flow identities. In particular, omitting district subscripts for simplicity, note that the stock-flow identities require that\(^6\)

---

\(^6\) These identities assume that all matches are brought about by the reported unemployed and vacancies (there being
Lagged outflows in (4) are described by lagged version of (3) as

\[
\begin{align*}
\log O_{t-1} & = \alpha u_{t-2} + \beta v_{t-2} + \epsilon_{t-1} \\
\log O_{t-2} & = \alpha u_{t-3} + \beta v_{t-3} + \epsilon_{t-2}
\end{align*}
\]  

(5)

Since district means are computed from all district observations, the means contain all values of the error term \( \{e_t: t = 1, 2, 3, \ldots, T\} \). This creates correlation between transformed explanatory variables and transformed error term, thus leading to biased estimates.

The first difference transformation also removes the district-specific effects and it contaminates the transformed variables only with recent error terms \( \{e_t: t = T-2, T-1\} \). To see this, rewrite (5) in a first difference form as

\[
\Delta o_t = o_t - o_{t-1} = \alpha (u_{t-1} - u_{t-2}) + \beta (v_{t-1} - v_{t-2}) + \epsilon_t - \epsilon_{t-1},
\]  

(6)

which can be expressed in a simplified notation as

\[
\Delta o_t = \alpha \Delta u_{t-1} + \beta \Delta v_{t-1} + \Delta \epsilon_t.
\]  

(7)

From equation (4) it follows that \( u_{t-1} \) and \( u_{t-2} \) contain \( e_{t-1} \) and \( e_{t-2} \), respectively, through outflows as no out-of-register matching). Other forms of matching may create more complicated identities but will not eliminate
in equation (5).  

The first difference transformation thus leaves further lags of $\Delta u_t$ \{u$_t$; t = 2, ..., T-4, T-3\}$^8$ uncorrelated with $\Delta e_{t-1}$ (i.e., with $e_{t-1}$ and $e_{t-2}$) and these further lags of $\Delta u_t$ can be used as instrumental variables. Vacancies in (6) are predetermined in the same statistical sense and may be treated the same way. Available instruments are therefore given by \{$\Delta u_t$, $\Delta v_t$; t = T-2, T-3, ...,2\}.

There are a number of additional specific features of the matching function model that need to be taken into account in estimation. First, identities in (4) show that lagged changes in inflows \{$\Delta s_t$; t = 2, 3, ...,T-2, T-1) are available as additional instruments because they codetermine the explanatory variables but do not directly affect outflow.

Second, rather than using changes in lagged values as instruments, we can use lagged levels because differences are simply linear combinations of levels.

Third, it is desirable to include month and year specific dummy variables as regressors to control for the seasonality in unemployment flows.

Fourth, although idiosyncratic errors \{e$_t$; t = 1, 2, ..., T) may in principle be uncorrelated, their first differences $\Delta e_t$ will by definition be autocorrelated. To see this, note that $\Delta e_t = e_t - e_{t-1}$ and $\Delta e_{t-1} = e_{t-1} - e_{t-2}$, and both term contain $e_{t-1}$. To warrant unbiased coefficients’ standard errors, robust variance-covariance matrix has to be used.

Fifth, further complications arise if $e_t$ has an autoregressive form. In this case the problem is that current $e_t$ and $e_{t-1}$ (and therefore $\Delta e_t$) contain all previous values \{e$_t$; t = 2, 3, ...,T-2, T-1) and

---

$^7$ It does not matter that (4) is defined in levels and (5) in logs.

$^8$ Note that the first observation (for t=1) of the first differenced variables is not available because observations for t=0 do not exist.

$^9$ It can be easily shown that $\text{CORR}(\Delta e_t, \Delta e_{t-1}) = \frac{\text{COV}(e_t - e_{t-1}, e_{t-1} - e_{t-2})}{\text{VAR}(e_t - e_{t-1}) \cdot \text{VAR}(e_{t-1} - e_{t-2})} = -0.5$ so that $\Delta e_t = -0.5 \cdot \Delta e_{t-1} + \text{error}_t$. 


this disqualifies past lags as instruments. This problem, if it appears, is fortunately substantially weakened by first differencing. Moreover, the correlation of instruments and error term declines quickly with the length of lags. Further lags of the explanatory variables therefore remain reasonably good instruments even in the case of autoregressive $e_t$.

Sixth, equation (3) represents a full adjustment model, which assumes that outflow reacts immediately and fully to changes in the stocks of unemployed and vacancies. Since in practice there may be numerous frictions, it is useful to consider also a partial adjustment model of matching, where outflow reacts to changes in RHS only partially during each period:

$$o_{i,t} = \gamma o_{i,t-1} + \alpha u_{i,t-1} + \beta v_{i,t-1} + a_i + \varepsilon_{i,t}. \quad (8)$$

In this model, the lagged value of outflow is endogenous by definition and contains $e_{t-1}$ and $e_{t-2}$. It is therefore correlated with $\varepsilon_t$ and leads to the same problem as that caused by the endogeneity of $u$ and $v$. Lagged outflow hence has to be instrumented and one can consider further lags $\{o_{i,t}: t=T-2, T-3, \ldots\}$ or their levels $\{o_{i,t}: t=T-3, T-4, \ldots\}$ as instruments.

Finally, Model (3) assumes that all the unemployed individuals are identical in terms of their propensity to match. Yet, studies such as Coles and Smith (1994) suggest that the propensity to match is higher at the time of entry into unemployment when the newly unemployed search through all existing vacancies. Later on, those who remain unemployed do not search through existing vacancies (those already explored) and match only with the newly posted vacancies. To reflect this so called “stock-flow” matching, we also include inflow into

---

10 The newly unemployed may also have not yet experienced depreciations of skills and psychological scarring.
unemployment, s_n, as an additional explanatory variable in (3). Assuming that job destruction is exogenous with respect to actual matching o_n, no additional instruments are needed.

**Adjusting for Varying Size of Districts**

Since the literature on matching has not taken into account the variation in the size of the unit of observation (in our case the district), many of the existing studies have probably generated biased estimates. The reason for the bias, explained in detail in Appendix IV, relates to the fact that the size of a district, measured for instance by its labor force L_i, is positively correlated with the values of O_i, U_i, S_i and V_i simply due to different sizes of districts. In this situation, when district-level variables are not adjusted for the size of the district labor force, the intercorrelations among O_i, U_i, S_i and V_i tend to be biased upward on account of the variation in the size of districts (see Appendix Table A2 for an illustration). The usual Cobb-Douglas specification based on cross-section data then provides biased estimates of coefficients unless there are constant returns to scale or the unadjusted U_i and V_i are uncorrelated with the labor force L_i. The direction of the bias of a (\beta) is negative if U_i (V_i) is positively correlated with L_i and matching displays increasing returns to scale. Either decreasing returns to scale or negative correlation (but not both) in turn lead to a positive bias. Therefore, if the matching process does not exhibit constant returns, the bias is likely to cause an incorrect acceptance of the constant returns hypothesis. The bias, and therefore the likelihood of an incorrect acceptance of the constant returns hypothesis, is greater, the greater is the portion of the correlation of U_i and V_i with L_i that is due to differences in the size of the district labor force. As we show in Appendix IV, the bias is very similar to that stemming from an omitted variable problem. In what follows we call this phenomenon the *spurious*

---

11 There are several possible measures of district size. We use the district labour force, but the results would not be materially affected by using other measures.
scale effect.\textsuperscript{12}

It can be shown that the spurious scale effect is avoided if one uses panel data and estimates a Cobb-Douglas function by fixed effects. In this case, the within transformation removes the spurious scale effects together with all other unobserved district-specific time-invariant effects.

4. The Data and Variable Definitions

In order to produce the best possible parameter estimates, we have assembled an extensive panel of data on 76 Czech, 38(79) Slovak, 21 Hungarian, 34 East German and 140 West German districts. The data for all countries cover the period from January 1991\textsuperscript{13} - December 2003 and contain monthly observations for the following variables:

\[ O_{i,t} = \text{the number of individuals flowing from unemployment in district } i \text{ during period } t; \]
\[ U_{i,t} = \text{the number of unemployed in district } i \text{ the end of period } t; \]
\[ S_{i,t} = \text{the normalized number of individuals flowing into unemployment (the newly unemployed) in } \]

\textsuperscript{14} An interesting question for future research is whether the size of districts and regions, the usual units of observation in the matching function studies, tends to be determined by an arbitrary administrative fiat or an endogenous optimization process of population settlements, based on historical economic forces that are in principle similar to an optimization process determining the size of firms.

\textsuperscript{13} At this moment, Hungarian data at our disposal start in January 1995 and the German data end in 2001. The missing data will be added once we receive them from the providers. The structure of Slovak districts was thoroughly changed in 1997 and we use Slovak data as two separate panels. In January 1997, three new Czech districts were formed from two original districts. These three districts are excluded from the analysis.
district $i$ during period $t$;\(^{14}\)

$$V_{i,t} = \text{the number of vacancies in district } i \text{ at the end of period } t;$$

\(^{14}\) Although the individuals flow into unemployment in the same calendar month, they enter on different days within the month. This means that they face different probabilities of finding vacancies during the calendar month. Assuming, that the inflow is approximately uniform over the month, we multiply the total monthly inflow by .5.
Although outflow to jobs is a theoretically preferred variable to total outflow, the data on outflow to jobs are available only for the Czech Republic, while data on total outflow are available for all the countries in our study. We carry out the estimation for the Czech Republic using both measures and find that the estimates based on total outflow and outflow to jobs are similar.\footnote{Total outflows and outflows to jobs are positively correlated, with the latter representing about 75 percent of the former in the Czech Republic.} As a result, the lack of data on outflow to jobs in other countries should not have a dramatic impact on our results (see also Petrongolo and Pissarides, 2000, for similar evidence from other countries).

5. Basic Statistics and Econometric Estimates

5.1 Basic Statistics

The basic statistics are described in Table 1...

5.2 Econometric Estimates

We start our discussion by presenting in Table 2 estimates of the matching function (3) for the West German districts during the 1994-2000 period of economic growth. The West German estimates provide a benchmark for a mature market economy against which we compare the estimates from the four transition economies, three of which proceeded autonomously (CR, SR and Hungary) and one of which (East Germany) was integrated with West Germany. In the last two lines of Table 2, we present the estimates from what we consider to be the most appropriate estimation method (first-difference IV model). On the preceding lines, we present results generated by the other estimating techniques used in the literature.
The OLS estimates of coefficients on unemployment and vacancies in Table 2 are low and imply decreasing returns to scale \((a + \beta < 1)\). Including month or annual time dummy variables does not affect the estimates. However, the OLS estimates of \(a\) and \(\beta\) based on variables adjusted for district size are lower, which is surprising as one would expect the spurious scale effect to bias coefficients toward constant returns to scale. As discussed earlier, both sets of OLS estimates are inconsistent due to the presence of unobserved fixed effects and if, as is likely, these unobserved effects are negatively correlated with unemployment and vacancies levels, the estimated coefficients are downward biased. This appears to be the case for the estimates in Table 2.

We next present in Table 2 the panel data estimates. The random effects, district dummy variable and fixed effects estimators all yield coefficients that are similar to the OLS estimates reported above, with the returns to scale \((a + \beta)\) being in the 0.8-0.85 range. The estimates based on mean deviations yield lower coefficients, but they are higher than the OLS estimates on adjusted levels and the resulting returns to scale are 0.76. As discussed above, these coefficients are still biased. The OLS estimates based on first differences have a notably higher coefficient on unemployment \((a = 1.75)\), implying increasing returns to scale \((a + \beta > 1)\). However, these estimates are also biased because \(u_{t-1}\) contains \(-e_{t-1}\) in \(u_{t-1}\) through the stock-flow identity and \(-e_{t-1}\) is contained also in \(\epsilon_t\). This causes positive correlation between transformed error term \(\epsilon_t\) and both explanatory variables \(u_{t-1}\) and \(v_{t-1}\), and brings about a positive bias observed in our estimates.

Our final set of estimates in Table 2 comes from an IV estimation based on first differences of variables. Since these are our preferred estimates, we report them in two versions: with and without the newly unemployed being included as a regressor. The model without the newly unemployed yields coefficients \(a = 1.32\) and \(\beta = 0.14\). These estimates are consistent. The instruments used are lagged
levels of explanatory variables plus lagged inflows, but close lags for T-2 and T-3 are excluded to guarantee strict exogeneity. The issue is whether the instruments used have adequate explanatory power. Moreover, first differencing decreases significantly the variance of the explanatory variables while doubling the variance of the error term and potential measurement error. This in turn leads to higher standard errors of estimated parameters. In all of our empirical work, we find that the explanatory power of the proposed instruments is adequate. When the newly unemployed are included in the regression (last row in Table 2), we find that their coefficient is significant at 0.12, indicating that the newly unemployed indeed display a high propensity to match. The last coefficients (not shown in Table 2) are estimates of partial adjustment model estimated by IV on first differences. Estimated partial adjustment coefficient \( \beta \) is insignificant and matching function parameters are very similar to the estimates based on the full adjustment model.

In Table 3, we present IV first-difference estimates of the parameters of the matching function for the Czech Republic, Hungary, Poland, Slovak Republic, East Germany, and West Germany. In order to capture the potential differences in the functioning of the labor markets in the early-to-mid 1990s and the late 1990s to early 2000s, respectively, we provide separate estimates for the 1994-96 and 1997-2005 periods. The earlier period corresponds to the early transition in the post-communist countries and a period of relatively slow economic growth in West Germany. The latter period captures the late transition in the post-communist economies and a period of relative boom and later slowdown in West Germany. For the earlier period, we do not have data for Hungary and Poland, but for the latter period we have data on all six economies. In all cases, we present coefficients from the

---

16 The instruments explain 20% to 75% of variation in explanatory variables. The lowest explanatory power of instruments was for vacancies (20-30%), and the highest power was for outflows (60-70%).
IV first difference extended model with unemployment, vacancies and the inflow into unemployment (newly unemployed) as regressors.

As may be seen from Table 3, the estimated coefficients on unemployment, vacancies and newly unemployed vary considerably across these economies and across the two time periods. During the 1994-96 period, we observe relatively precisely estimated coefficients in the Czech Republic (except for vacancies) and West Germany, pointing to constant (0.98) and increasing (1.66) returns to matching, respectively. The main difference between the two countries lies in a much higher coefficient on unemployment in West Germany than the Czech Republic (1.24 v. 0.58). In both East Germany and Slovakia, the coefficients on unemployment and vacancies are very imprecisely estimated, suggesting that there was considerable diversity of matching patterns across the districts in these two economies.

During the more recent period of 1997-2005 we obtain precisely estimated coefficients in all countries except Poland (where we have only few years of data). The returns to scale are increasing in all the countries, with the highest returns being observed in Hungary (2.28) and East Germany (2.18), with Czech Republic, West Germany and Slovakia coming in next with increasing returns of 1.73, 1.54 and 1.41, respectively. The Czech and West German estimates are significantly different from 1.0 at all conventional levels of significance, while the Slovak estimate is significantly different from unity only at a 6% test level. The Polish results also point to increasing returns, but the coefficient on unemployment (1.34) is imprecisely estimated, the coefficient on vacancies is actually negative (-0.25) and one cannot reject the hypothesis that the estimated returns to scale are constant or even decreasing. Since we have only data for a few years on Poland, these results are tentative.

The results in Table 3 indicate that in the early (1994-96) transition period posted vacancies...
played a negligible part in outflow from unemployment (job creation) in all the post-communist countries for which we have data. In contrast, unemployment was an important determinant of outflow in the Czech Republic, which maintained a low unemployment rate, and in East Germany, which had a high unemployment rate but also a high inflow rate and very sizable active labor market programs. Unemployment is statistically unimportant in Slovakia, which experienced a rapid rise in unemployment during this period. Interestingly, inflow into unemployment generates a similar (0.2-0.3) and statistically significant coefficient in all these economies, suggesting that job creation involved significantly the newly unemployed. During the late 1990s and early-mid 2000s, the efficiency of matching rises in all the transition economies, while declining somewhat in West Germany. Leaving Poland aside, we observe that East Germany and Hungary have the highest returns to scale, driven by relatively high coefficients on all three explanatory variables – unemployment, vacancies and inflow into unemployment. In East Germany, this may be in part generated by the very sizable active labor market policies. These are present in Hungary as well, but not on the same scale. The Hungarian results hence suggest that the underlying feature is a relatively efficiently functioning matching system in a free labor market. These findings are consistent with the high unemployment in East Germany and low unemployment in Hungary. The Czech economy generates lower returns to scale than Hungary and East Germany, but higher than those of Slovakia. The Czech-Slovak difference is driven primarily by higher coefficients on vacancies and inflow in the Czech estimates, suggesting that these aspects of the matching process account for its higher efficiency in the Czech lands. These findings are consistent with the relatively low unemployment in the Czech Republic as compared to Slovakia.

In Figures 1-4 we present estimates of the matching function parameters from repeated estimations
on two years of data, from which we consecutively remove the oldest quarter and add the newest one. The dots in the figures are the individual point estimates and the lines provide one standard error confidence interval bands around these point estimates. In Figure 1, we compare the estimates from East and West Germany, thus observing the difference between a market and a transition economy that share the same historical and cultural background, but with the transition economy being affected by administratively set subsidies and increased wages as a result of unification. In Figure 2, we in turn compare West Germany and the Czech Republic – thus eliminating the effect of unification but also common history and culture.

As may be seen from Figures 1 and 2, throughout the 1994-2005 period, the Czech and West German estimates are more similar than those of East and West Germany. In particular, both the levels and variation over time in the estimated coefficients on \( u, v, \) and \( s \) are very similar in the Czech and West German data. The East German estimates are also broadly similar to the West German ones, but they display a much greater variation over time and their dynamics is different. Thus the East German coefficient on \( u \) is higher than the West German one around 1994-95 and then again around 1997. Similarly, while the West German coefficient on \( v \) varies mostly between 0 and 0.3, the East German coefficient fluctuates widely and reaches into the 0.5-0.8 range in the 2000-02 period.

In Figure 3, we present the time specific country intercepts \( p_t \), whose estimation is explained in equation (7) of Appendix B. These coefficients give an estimate of the base efficiency of matching in each country, holding \( u, v \) and \( s \) constant. As may be seen from the figure, Hungary and East Germany, which have relatively high estimated coefficients on \( u, v \) and \( s \) (and therefore also high returns to scale) have low base efficiency. West Germany and the Czech Republic, which have intermediate coefficients and returns to scale, register intermediate values of \( p_t \). Finally, the Slovak Republic and Poland, the two countries with the lowest estimated coefficients (and returns to scale) have the highest time specific intercepts.
In Figure 4 we depict the estimated returns to scale in matching for the Czech Republic and East and West Germany. The figure shows that the returns are most volatile and relatively high in East Germany, with the difference being the most pronounced in the earlier (1994-98) and the most recent (2001-05) periods. The Czech and West German returns to scale are similar, with the Czech returns being below the West German ones in 1994-98 and exceeding the West German ones in the 2002-04 period. Overall, Figure 4 supports our earlier findings that in terms of matching, the Czech and West German labor markets seem to be more similar than either one of them is to the East German market.

6. Concluding Observations

Our paper is motivated by the three alternative hypotheses about the causes of unemployment in the Central European transition economies. The first hypothesis (H1) stipulates that transition is still unfinished and is consistent with the observation that inflow (presumably from old, “socialist” jobs) is high, outflow (job creation in the new “market” economy) is fine and unemployment is high due to the high inflow. The second hypothesis (H2) states that matching between the unemployed and vacancies is fine, with high unemployment being caused by low demand for labor (e.g., due to restrictive macroeconomic policies, overvalued exchange rate, or globalization shocks). The manifestation of this would be low level of vacancies relative to inflow, irrespective of unemployment. The third hypothesis (H3) is that high unemployment is caused by inefficient matching, brought about for example by inadequate labor market institutions or geographical or skill mismatch. Here we would observe both unemployment and vacancies being high, but not necessarily in the same districts or skill groups.

Our results suggest that the situation differs across the sampled economies and different
hypotheses receive support in different countries. Our benchmark market economy, namely the West German part of Germany, is an economy with rising unemployment and inflow, declining vacancies and relatively efficient matching (high returns to scale). Its outcome is hence most consistent with H1 and H2. Czech Republic appears to be in a similar situation, and with rising unemployment, as well as inflow and outflow, and a declining vacancy rate and high returns to matching, it increasingly gives support to H1 and H2. However, since Czech Republic has increasingly pursued a policy of low interest rates and fiscal deficits, the support for H2 implies the presence of negative exogenous demand shocks. East Germany’s results are also in line with H1 and H2, in that the region has relatively high unemployment and inflows, a low vacancy rate and very efficient matching (including outflow into the training programs). Finally, the Slovak and Hungarian economies provide the two extremes. The Slovak economy (and probably also the Polish one, to be confirmed with more data) displays relatively low returns to scale in matching, and until recently it has suffered from very high unemployment, rising inflow rates and a low vacancy rate. It has also had a relatively loose monetary and fiscal policies and a floating exchange rate. Its outcome is hence consistent with a combination of H1, H2 and H3. Finally, over the last several years Hungary has managed to lower its unemployment rate to around 8% and it displays the highest estimated returns to matching. Given its low vacancy rate relative to inflows, the existing unemployment seems to be consistent with H1 and H2.

Overall, our findings suggest that the transition economies contain two broad groups of countries. The first group comprises the Czech Republic, Hungary and (possibly) East Germany and it resembles the West German (benchmark) case. These countries display efficient matching of the unemployed and vacancies, and their unemployment appears to be driven by restructuring and low demand for labor. The East German case is complex because of its major active labor market policies,
and in some sense it resembles more the second group of countries, exemplified by Slovakia. These countries, in addition to restructuring and low demand for labor, appear to suffer from a structural mismatch (i.e., display less efficient matching).
REFERENCES


Schaffer, M., “Government Subsidies to Enterprises in Central and Eastern Europe: Budgetary


Table 2
Matching Function Estimates for West Germany

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cross-sectional estimators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLS</td>
<td>0.680</td>
<td>0.003</td>
<td>0.148</td>
<td>0.003</td>
<td>-</td>
<td>-</td>
<td>0.828</td>
<td>0.853</td>
<td>14734</td>
</tr>
<tr>
<td>OLS_Month_Dummies</td>
<td>0.687</td>
<td>0.003</td>
<td>0.129</td>
<td>0.002</td>
<td>-</td>
<td>-</td>
<td>0.817</td>
<td>0.896</td>
<td>14734</td>
</tr>
<tr>
<td>OLS_Year_Dummies</td>
<td>0.687</td>
<td>0.017</td>
<td>0.130</td>
<td>0.014</td>
<td>-</td>
<td>-</td>
<td>0.817</td>
<td>0.896</td>
<td>14734</td>
</tr>
<tr>
<td>OLS_Adjusted</td>
<td>0.547</td>
<td>0.028</td>
<td>0.029</td>
<td>0.021</td>
<td>-</td>
<td>-</td>
<td>0.576</td>
<td>0.616</td>
<td>14734</td>
</tr>
<tr>
<td><strong>Panel estimators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_Random_Coeff</td>
<td>0.736</td>
<td>0.009</td>
<td>0.075</td>
<td>0.004</td>
<td>-</td>
<td>-</td>
<td>0.811</td>
<td></td>
<td>14734</td>
</tr>
<tr>
<td>P_District_Dummies</td>
<td>0.742</td>
<td>0.011</td>
<td>0.072</td>
<td>0.004</td>
<td>-</td>
<td>-</td>
<td>0.815</td>
<td>0.949</td>
<td>14734</td>
</tr>
<tr>
<td>P_FE</td>
<td>0.742</td>
<td>0.011</td>
<td>0.072</td>
<td>0.004</td>
<td>-</td>
<td>-</td>
<td>0.815</td>
<td>0.664</td>
<td>14734</td>
</tr>
<tr>
<td>P_mean_Deviations</td>
<td>0.687</td>
<td>0.039</td>
<td>0.073</td>
<td>0.009</td>
<td>-</td>
<td>-</td>
<td>0.760</td>
<td>0.658</td>
<td>14734</td>
</tr>
<tr>
<td>P_1st_Dev</td>
<td>1.643</td>
<td>0.064</td>
<td>0.067</td>
<td>0.013</td>
<td>-</td>
<td>-</td>
<td>1.710</td>
<td>0.638</td>
<td>14734</td>
</tr>
<tr>
<td>P_1st_Dev_IVa</td>
<td>1.315</td>
<td>0.042</td>
<td>0.140</td>
<td>0.031</td>
<td>-</td>
<td>-</td>
<td>1.455</td>
<td>0.632</td>
<td>14734</td>
</tr>
</tbody>
</table>
**Figure 1**

**Table 3**


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>0.581</td>
<td>0.155</td>
<td>0.148</td>
<td>0.150</td>
<td>0.253</td>
<td>0.032</td>
<td>0.648</td>
<td>0.98</td>
<td>2661</td>
</tr>
<tr>
<td>SR</td>
<td>0.841</td>
<td>0.626</td>
<td>0.207</td>
<td>0.162</td>
<td>0.171</td>
<td>0.049</td>
<td>0.286</td>
<td>1.22</td>
<td>1292</td>
</tr>
<tr>
<td>EG</td>
<td>1.560</td>
<td>0.564</td>
<td>-0.133</td>
<td>0.100</td>
<td>0.296</td>
<td>0.069</td>
<td>0.505</td>
<td>1.72</td>
<td>1211</td>
</tr>
<tr>
<td>WG</td>
<td>1.236</td>
<td>0.066</td>
<td>0.224</td>
<td>0.041</td>
<td>0.200</td>
<td>0.022</td>
<td>0.665</td>
<td>1.66</td>
<td>5004</td>
</tr>
<tr>
<td>HU</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>PL</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1997-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>SR</td>
</tr>
<tr>
<td>EG</td>
</tr>
<tr>
<td>WG</td>
</tr>
<tr>
<td>HU</td>
</tr>
<tr>
<td>PL</td>
</tr>
</tbody>
</table>
Figure 1

Matching Function Coefficients
Figure 2

Matching Function Coefficients
Time Specific Intercepts
We estimate

\[ o_{i,t} = \beta u_{i,t} + \gamma v_{i,t} + \delta s_{i,t} + a_i + \pi_t + \varepsilon_{i,t} \]  

(1)

where our unit of observation, \( i \), is a district. We assume that district size does not affect matching (spurious scale effect) and 1st difference estimation removes scale effect. Our assumption in fact means that we assume constant RTS where scale is meant to be district size. On the other hand, we do not impose restrictions on RTS where scale means size of \( U, V, S \) not due to district size but due to economic factors.

Different countries differ in district sizes = units of observations. Therefore, we have to express intercept as if countries would have districts of similar size. It can be shown that it can be done as

\[ R_{i,t} = o_{i,t} - \left[ \beta u_{i,t} + \gamma v_{i,t} + \delta s_{i,t} \right] + \left[ 1 - \beta + \gamma + \delta \right] = a_i + \pi_t + \varepsilon_{i,t} \]

(2)

where \( l_i \) is \( \log(\text{LF}_i) \).
Appendix A

The Spurious Scale Effect

For the purposes of exposition, we present a simple case that demonstrates the impact of the spurious scale effect on estimation. Assume that the country is a homogeneous territory divided administratively into districts of different sizes with identical labor market conditions and characterized by a simple Cobb-Douglas matching function with increasing returns to scale ($\beta_u + \beta_v > 1$). As a result of homogeneity, the outflow,
unemployment and vacancies in each district, \( O_i, U_i \), and \( V_i \), are proportional to national aggregates \( O, U, \) and \( V \),

\[
O_i = l_i O, \quad U_i = l_i U, \quad \text{and} \quad V_i = l_i V, \quad (35)
\]

where \( l_i \) is the share of district \( i \) in the national labor force, defined as \( L_i / L \). Not taking the district size into account and estimating the matching function on unadjusted cross-sectional data amounts to

\[
\ln O_i = \alpha + \beta_u \ln U_i + \beta_v \ln V_i + \epsilon_i \nonumber
\]

estimating the model

\[
\ln( l_i O) = \alpha + \beta_u \ln( l_i U) + \beta_v \ln( l_i V) \nonumber
\]

Substituting (35) into (36), we get which in turn yields

\[
\ln l_i = (\beta_u + \beta_v) \ln l_i + (\alpha - \ln O + \beta_u \ln U + \beta_v \ln V) + \epsilon_i
\]

It is obvious that the estimation of (38) is identical to estimating (36). However, (38) represents a regression of \( l_i \) on itself plus a constant term. It will hence generate constant returns to scale \((\beta_u + \beta_v = 1)\) and a zero constant term \((a = \ln O - \beta_u \ln U - \beta_v \ln V)\). The estimation of model (36) therefore yields biased estimates since we have postulated increasing returns to scale.

A remedy for the problem is to adjust the variables by the district size in order to obtain the following model:

\[
\ln \left( \frac{O_i}{L_i} \right) = \alpha + \beta_u \ln \left( \frac{U_i}{L_i} \right) + \beta_v \ln \left( \frac{V_i}{L_i} \right) + \epsilon_i
\]

which can be rearranged as

\[
\ln O_i = \alpha + \beta_u \ln U_i + \beta_v \ln V_i + (\beta_u + \beta_v - 1) \ln L_i + \epsilon_i
\]

A comparison of the adjusted model (40) to the unadjusted model (36) indicates that they are equivalent if and only if at least one of the two following conditions is satisfied:

(i) \( \beta_u + \beta_v - 1 = 0 \) (the underlying matching displays constant returns to scale)

Note that for expositional purpose the variance in district level variables is brought about completely by the administrative variation in district sizes rather than by economic factors.
or

\[(ii) \ Cor(\ln L_i, \ln U_i) = Cor(\ln L_i, \ln V_i) = 0.\]

In our example, neither condition is satisfied because (i) we are assuming increasing returns to scale \((\beta_u + \beta_v - 1 > 0)\) and (ii) \(U_i = UL_i/L\) and \(V_i = VL_i/L\), resulting in \(Cor(\ln L_i, \ln U_i) = Cor(\ln L_i, \ln V_i) = 1\).

In general, one has no a priori information about the returns to scale since they represent a statistic that is to be estimated from the data. The intercorrelations among the unadjusted variables can of course be checked in advance. Judging from the data at our disposal, these intercorrelations are positive and significant.

**Appendix B**

**Parameters of the Matching Function**

Since the matching function is usually viewed as a black box, in this appendix we discuss a possible interpretation of the parameters \(a, \beta, \gamma, \) and \(d\) of the matching function that we estimate.

Our log-linear approximation of the matching function is

\[ o_t = a + \beta u_t + \gamma v_t + \delta s_t \]  

(1)

In a discrete time framework, the accounting relationship between stocks and flows of unemployed can be written as

\[ \Delta u_t \equiv u_t - u_{t-1} = s_t - o_t \]  

(2)

and its corresponding form in continuous time framework

\[ u_t \equiv s_t - o_t \]  

18

and

\[ u_t \equiv s_t - o_t \]  

(3)

where the star marks steady-state values. Substituting for steady-state outflow to (1), steady-state unemployment can be expressed as

---

18 Net-flow identity holds for logs of levels only in a steady-state. Out of equilibrium, the identity is not fulfilled only for levels, and does not hold perfectly for logs.
Equation 4 describes the relationship between the steady state $u$ and $v$. It implies that steady-state unemployment increases with (steady-state) inflow rate, reflecting the contribution of natural employment turnover. The contribution of turnover is smaller if the newly unemployed match better (higher $d$). If $\gamma > 0$, the number of vacancies has a negative impact on steady-state unemployment via $\gamma$. The intercept of the matching function, $\alpha$, has a negative effect on steady-state unemployment. All 3 effects translate into steady-state unemployment via $1/\beta$, the inverse of the coefficient on unemployment.

Note that in the above, we implicitly assume that vacancies are exogenous and are fixed at $v^*$. Equation (4) describes steady-state unemployment as a function of exogenous variables $v^*$ and $s^*$. In particular, it describes a negatively sloped locus, $u^* = f(v^*)$, but does not identify a unique point $(u^*, v^*)$ on it. The number of posted vacancies in the steady state is determined by both the endogenous matching process (equation 4) and exogenously determined labor demand. To identify unique points on the $u^*-v^*$ locus, we would have to introduce vacancy supply equation. But this is not simple straightforward task (we would have to follow theoretical exposition in Jackman et al. (1990). 19

Finally, let me note that in a steady state of an ideal world where all matching is done through labor offices, the outflow of vacancies, $q$, would be equal to the number of unemployed-vacancy matches, $o$, and the inflow of new vacancies, $z$, would be equal to the inflow of unemployed, $s$ (pure case of restructuring where lay-offs result in equal number of vacancies which are being filled by new hires after some period of search). This implies that in a steady state

$$z_t^* = s_t^* = o_t^* = q_t^*.$$  

(5)

Going back to equation 4. Would not it be of interest to present actual values of those 3 terms (contributing to observed unemployment) in the bracket for each country? This can be done but before we have to decide on time-spans to estimate the matching function. Moreover, the 3rd term alpha in fact comprises district and time specific effects that we address next.

Decomposition of the Error Term

Stochastic version of the model for selected time-span we estimate is following

$$o_{i,t} = \beta u_{i,t} + \gamma v_{i,t} + \delta s_{i,t} + a_t + \pi_t + e_{i,t},$$  

(6)

Here, $a_i$ represents the district specific (fixed) effect, $p_t$ represents the time specific country effect (common to all districts), and $e_{i,t}$ is an

---

Figure 4
idiosyncratic error term. Doing so, there is implicit assumption that within the time span, parameters $\beta, \gamma$ and $d$ remain constant.20

Using our instrumented first differences approach, we estimate parameters of the deterministic part $\beta, \gamma$, and $d$. Substituting these back in levels into our model in (6), we express unexplained residuals as

$$R_{i,t} = o_{i,t} - \left[ \hat{\beta} u_{i,t} + \gamma \hat{\delta}_{i,t} + \delta \hat{\delta}_{i,t} \right] = a_i + \pi_t + \epsilon_{i,t}$$  \hspace{1cm} (7)

Since we are interested in $a_i$ and $p_t$, because their distributions provide important information, we estimate them as follows. First, we regress $R_{i,t}$ on a complete set of yearly and monthly dummy variables. Then we compute fitted values to get predicted time-effects $\pi_t$. Finally, from regression residuals $r_{i,t}$ representing $(a_i + \epsilon_{i,t})$, we estimate district-specific effects as

$$\bar{a}_i = \frac{1}{T} \sum_{t=1}^{T} r_{i,t} \text{ for each district } i.$$  \hspace{1cm} (8)

Note that $\sum_{t=1}^{T} \bar{a}_i = 0$ by definition, and country specific total factor productivity intercept $a$ is contained in $\pi_t$.

On pages bellow we present some basic preliminary findings.

---

20 It is also assumed that these parameters are common to all districts. Without this assumption, we would have hard times identifying sub-samples of districts with different matching function. But this is common limit of most other empirical studies like this.
I estimated matching function for the time-span 1997-2005 for each country. Below, According to the methodology above, I estimated unobserved components reviewed below.

This graph presents time specific fixed effects for all countries. It shows substantial differences across countries and weak trends in some of them. Small fluctuations have systemic monthly pattern.
Persistence of $r(i,t)$ 4/2004 vs. 4/1997

This graph shows the persistence $a_i$ in the residuals $r_{i,t} = (a_i + e_{i,t})$ over longer time span. W. Germany shows stable pattern (dominance of $a_i$) while post-communist countries (CR, SR, EG) show substantial changes (dominance of idiosyncratic $e_{i,t}$).
This graphs shows distribution of district fixed effects by countries. Reading these distributions, one should be aware that district size matters here. To understand it, imagine that district (district is identical with unit of observation) size in country A corresponds to the size of a region in another country B. If sub-districts of districts in country A (similar in size to districts in country B) are heterogeneous within each district of country A, it could be that all districts in A look alike while at the level of sub-districts country A is equally heterogeneous as country B. Simply said, country whose units of observation have bigger size are likely to show smaller heterogeneity across districts. Below are mean district labor forces (SR1 ~ before 1997, SR2 ~ after 1996). Taking this into account, it is clear that W. German districts are much more heterogeneous than E. German ones (note that district sizes in both German lands are similar). Very high is heterogeneity in Hungary (size comparable to German districts).

Czech Republic’s heterogeneity is similar to W. Germany but given that Czech district is ~3 times bigger, we can conclude that Czech Republic has very heterogeneous regions. In case of Poland, our unit of observation is very high (like Czech region). No wonder that the distribution for Poland is very narrow.
### Figure 4

<table>
<thead>
<tr>
<th>country</th>
<th>mean(th. of LF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>68360.91</td>
</tr>
<tr>
<td>SR1</td>
<td>66960.13</td>
</tr>
<tr>
<td>SR2</td>
<td>33563.39</td>
</tr>
<tr>
<td>EG</td>
<td>211318.2</td>
</tr>
<tr>
<td>WG</td>
<td>227010.8</td>
</tr>
<tr>
<td>HU</td>
<td>219431.2</td>
</tr>
<tr>
<td>PL</td>
<td>1072322.0</td>
</tr>
</tbody>
</table>