Forecasting with a mismatch-enhanced labor market matching function

Christian Hutter
Enzo Weber

ISSN 2195-2663
Forecasting with a mismatch-enhanced labor market matching function

Christian Hutter (IAB)
Enzo Weber (IAB, Universität Regensburg)
## Contents

Abstract ................................................................................. 4  
Zusammenfassung ................................................................. 4  
1 Introduction ..................................................................... 5  
2 Theory and measurement .................................................... 6  
  2.1 Theoretical background ................................................. 6  
  2.2 Data ........................................................................... 7  
  2.3 Measuring mismatch ..................................................... 12  
3 Forecast evaluation ............................................................ 16  
  3.1 Forecast evaluation setting ............................................. 16  
  3.2 Performance of the standard mismatch indicator ................. 19  
  3.3 The importance of the different dimensions of mismatch .......... 22  
4 Conclusion ....................................................................... 24  
References ........................................................................ 26
Abstract

This paper investigates the role of mismatch between job seekers and job openings for the forecasting performance of a labor market matching function. In theory, higher mismatch lowers matching efficiency which increases the risk that the vacancies cannot be filled within the usual period of time. We investigate whether and to what extent forecasts of German job findings can be improved by a mismatch-enhanced labor market matching function. For this purpose, we construct so-called mismatch indicators that reflect regional, occupational and qualification-related mismatch on a monthly basis. In pseudo out-of-sample tests that account for the nested model environment, we find that forecasting models enhanced by the mismatch indicator significantly outperform their benchmark counterparts for all forecast horizons ranging between one month and a year. This is especially pronounced in the aftermath of the Great Recession where a low level of mismatch improved the possibility of unemployed to find a job again.

Zusammenfassung


JEL classification: C22, C52, C53, C78, E24, E27

Keywords: matching function, mismatch indicators, forecast evaluation

Acknowledgements: We are grateful to Anja Bauer as well as participants of the DIW Macroeconometric Workshop 2013, the IWH-CIREQ Macroeconometric Workshop 2013 and the UR-IAB seminar 2014 in Regensburg for helpful suggestions and valuable input.
1 Introduction

Many approaches for relating unemployment and vacancies to job findings rely on matching theory (see e.g. Mortensen and Pissarides (1994) and Petrongolo and Pissarides (2001)) and thus consider vacancies and unemployed as inputs into the production function of job findings. Most studies, however, assume the efficiency parameter to be constant over time. Only in recent years this strong assumption has been called into question. Barnichon and Figura (2011) or Sedlacek (2011) and Klinger and Weber (2014) for instance allow for time-varying matching efficiency.

We contribute to the literature by applying the concept of a time-varying efficiency parameter from a forecasting perspective. Particularly, we allow matching efficiency to depend on structural imbalance between the supply and demand side of the labor market. Employing a mismatch index in forecasting regressions is an economically attractive way of modeling structural change in the relationship between unemployment, vacancies and job findings.

The underlying data from the statistics department of the Federal Employment Agency (FEA) allow for disaggregation of unemployed and registered vacancies at the levels of 21 occupational segments, 50 labor market regions and 3 qualification groups. This way we get detailed empirical evidence on the size and development of mismatch in Germany in the past 13 years, including the years after the Hartz labor market reforms.

The focus is on discrepancies in all 21x50x3 markets due to occupational, regional and qualification-related incongruence between the attributes of the unemployed and the demands of the jobs. For this purpose, we construct a mismatch index that captures such imbalances in addition to the pure quantity of vacancies and unemployed and investigate whether models enhanced by a mismatch indicator can improve predictive ability.

After log-linearisation of the matching function, the mismatch indicator appears as an additional regressor in our forecasting equations. As a consequence, the usual out-of-sample test of Diebold and Mariano (1995) cannot be implemented. We thus employ the nested-model test described in Clark and West (2007), applying a one-sided test for equal predictive accuracy with the alternative hypothesis being worse forecast performance of the nesting model. We find that our enhanced matching function significantly outperforms its benchmark counterpart without mismatch indicator. Gains in forecast accuracy can be shown for all forecast horizons ranging between one month and one year. It is especially pronounced in the aftermath of the Great Recession where a low level of mismatch improved the job chances of the unemployed.

By gradually removing the occupational, regional or qualification-related dimension from our index calculations we generate alternative indices that allow quantifying the size and development of the different types of mismatch separately. This way we can identify which of the aforementioned dimensions has the biggest impact on forecast accuracy. The results emphasize the positive contribution of the occupational dimension for forecasts in the range of six to twelve months ahead. For the very short term, however, relying solely on qualificatory mismatch might be entirely sufficient with respect to forecasting job findings.
The remainder of the paper is structured as follows: The theoretical background of our forecasting equations, namely the enhanced stock-flow matching model, is introduced in the first part of section 2. The second part describes the data used for forecast evaluation and construction of the mismatch indices, and subsection 2.3 briefly describes various concepts of measuring mismatch and shows the development of the mismatch index following Jackman and Roper (1987). Section 3 compares the forecasting performance of a mismatch-enhanced labor market matching function to that of the corresponding benchmark models using nested model out-of-sample tests. Furthermore, it discusses the question which dimensions of mismatch provide the highest value added in forecasting German hiring figures. The last section concludes.

2 Theory and measurement

2.1 Theoretical background

The well-known search and matching theory (see e.g. Mortensen and Pissarides (1994), Petrongolo and Pissarides (2001), Shimer (2007) and Yashiv (2007)) states that vacancies ($V$) and unemployed ($U$) form matches ($H$ for hirings) through a Cobb-Douglas production function in the style of equation (1).

$$H_t = \Phi \cdot V_{t-1}^{\alpha} \cdot U_{t-1}^{\beta},$$  \hspace{1cm} (1)

where $\Phi$ denotes the efficiency parameter that carries the information about the location of the Beveridge curve. $\alpha$ and $\beta$ are the elasticities of new matches with respect to vacancies and unemployed, respectively.

Based on the idea of Coles and Smith (1998), Ebrahimy and Shimer (2010) augmented the basic model by the respective flow counterparts of $V$ and $U$. The idea of the so-called stock-flow matching model is that job searchers first screen the stock of job openings and firms first look at the stock of unemployed. Those agents on the demand or supply side who could not successfully fill the vacancy or find a job are assumed to search among the newly incoming applicants or job openings only. It can be shown that in a steady state, stock-stock matches are less likely than stock-flow matches. The stock-flow matching function follows as:

$$H_t = \Phi \cdot V_{t-1}^{\alpha_{st}} \cdot U_{t-1}^{\beta_{st}} \cdot \dot{V}_t^{\alpha_{fl}} \cdot \dot{U}_t^{\beta_{fl}},$$  \hspace{1cm} (2)

where aggregate unemployment and vacancies enter the regression both as stock ($U$, $V$) and as flow ($\dot{U}$, $\dot{V}$) variables. The corresponding parameters are indexed by $st$ and $fl$, respectively.

However, even in the stock-flow matching model the efficiency parameter is assumed to be constant over time. This could be problematic since the efficiency parameter is degraded to
a sort of Solow residual that has to capture any dynamics that cannot be accounted for by
the stocks or flows of unemployment and vacancies. Sedlacek (2011), for instance, finds
that the efficiency parameter can explain about 25 percent of the fluctuations in the job
finding rate. Not accounting for these dynamics could not only lead to biased estimates of
the structural parameters of the matching function but also (negatively) affect the accuracy
in forecasts of hirings.

In the underlying paper, we model \( \Phi \) to depend on the extent of structural incongruence
between vacancies and unemployed. Structural imbalance occurs when a large number of
unemployed coincides with a small number of vacancies (and vice versa) in a given micro
market. It can lead to long-run unemployment because workers and vacancies would not
form a match even in absence of search frictions or imperfect information. A better con-
gruence of both groups with respect to relevant attributes such as qualification, occupation
and region is expected to influence matching efficiency as expressed in equation (3).

\[
\Phi_t = f(C, MM_{t-1}) = C \cdot MM_{t-1}^\gamma,
\]

(3)

where \( MM \) denotes an aggregate measure of structural imbalance capturing the relevant
dimensions of mismatch, and \( C \) other factors that are not modeled here. Hence, \( \gamma \) is
expected to be negative since a higher regional, occupational or qualificatory incongruence
should hamper matching efficiency.

Inserting (3) into (2) yields after log-linearisation:

\[
h_t = c + \gamma \cdot mmt_{t-1} + \alpha_{st} \cdot vt_{t-1} + \beta_{st} \cdot ut_{t-1} + \alpha_{fl} \cdot \dot{vt} + \beta_{fl} \cdot \dot{ut},
\]

(4)

where lower-case letters denote the natural logarithm. The measure of structural imbalance
\( MM \) now appears in additive form which has some implications when it comes to forecast
evaluation (see section 3).

2.2 Data

In order to construct the mismatch indices and to evaluate their value added in out-of-
sample forecasts, we use monthly data from the statistics department of the Federal Em-
ployment Agency (FEA). As target variable we take seasonally adjusted hiring figures on
the aggregate level, i.e the monthly outflow from unemployment into employment on the
primary labor market. Hence, transitions from unemployment into subsidized employment
or into labor-market measures are excluded.

Figure 1 shows the development of the aggregate, seasonally adjusted outflow from un-
employment into employment on the primary labor market, from January 2000 until March
2014. Among the variables on the right-hand-side of our forecast equation are the aggre-
gate number of both the stock and the inflow of unemployed and vacancies, respectively.
Figure 1: Aggregate number of hirings, seasonally adjusted

Figure 2: Unemployed and vacancies (stock variables), seasonally adjusted
Figure 2 shows the stock of aggregate unemployed and registered vacancies (both seasonally adjusted) as reported by the FEA, from January 2000 until March 2013. The latter variable comprises all job offers that employers report to the respective local agencies and that are approved for placement. Figure 3 shows the aggregate number of the respective monthly inflows.

Constructing the mismatch indices as described in subsection 2.3 requires detailed information about the regional, occupational and qualificatory distribution of the respective job openings and unemployed on a monthly frequency. For this purpose, we exploit a large data set from the FEA that allows us to account for the most relevant dimensions of mismatch. With regards to qualification-related incongruence, we distinguish between the following three groups: experts (people with academic training), skilled workers and specialists (people with completed educational or vocational training), and helpers (people without completed vocational training).

Furthermore, we account for regional mismatch by using the 50 labor market regions proposed by Kropp and Schwengler (2014) and shown in figure 4. The delineation of these functional labor market regions relies on commuter flows between all German municipalities since 1993. Its main advantages are the high stability over time and its excellent properties with respect to self-supply and commuter ratio. On average, 87.4 percent (standard deviation: 4.4 percentage points) of the jobs are taken by workers that live within the

---

1 Our last estimation period ends in March 2013. Our evaluation period for 12-months-ahead forecasts ends in March 2014. As a consequence, the observation periods shown in figures 2 and 3 differ from that in figure 1.

2 Consequently, the chosen variable does not cover the whole job market. However, there is no variable available at a monthly frequency that takes into account the potentially changing share of vacancies that are registered at the employment agencies.
Figure 4: Germany’s 50 labor market regions

Notes: The labor market regions (black borderlines) are taken from Kropp and Schwengler (2011) and slightly modified so that they are in accordance with the most recent county borders (grey borderlines).
<table>
<thead>
<tr>
<th>No.</th>
<th>Occupational Segment</th>
<th>Exemplary Occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Agricultural occupations</td>
<td>farmer, fisher, gardener, forester</td>
</tr>
<tr>
<td>2</td>
<td>Miner/chemical occupations</td>
<td>miner, mining engineer, chemical worker</td>
</tr>
<tr>
<td>3</td>
<td>Glass, ceramic, paper producer</td>
<td>ceramicist, glazier, paper producer, printer</td>
</tr>
<tr>
<td>4</td>
<td>Textile, leather producer</td>
<td>spinner, weaver, tailor, sewer, shoemaker</td>
</tr>
<tr>
<td>5</td>
<td>Metal producer</td>
<td>metal worker, plumber, mechanical engineer</td>
</tr>
<tr>
<td>6</td>
<td>Electricians</td>
<td>electrician, electrical engineer</td>
</tr>
<tr>
<td>7</td>
<td>Wood occupations</td>
<td>wood processor, woodworker, joiner</td>
</tr>
<tr>
<td>8</td>
<td>Construction occupations</td>
<td>bricklayer, carpenter, roofer, tiler</td>
</tr>
<tr>
<td>9</td>
<td>Hotel/restaurant occupations</td>
<td>baker, butcher, cook, barkeeper</td>
</tr>
<tr>
<td>10</td>
<td>Storage/transport occupations</td>
<td>conductor, motorist, driver, mail distributor</td>
</tr>
<tr>
<td>11</td>
<td>Merchandise occupations</td>
<td>merchant, cashier, accounting clerk, banker</td>
</tr>
<tr>
<td>12</td>
<td>White collar worker</td>
<td>accountant, clerk, member of parliament</td>
</tr>
<tr>
<td>13</td>
<td>Security occupations</td>
<td>gate keeper, firefighter, guard, chimney sweeper</td>
</tr>
<tr>
<td>14</td>
<td>Social/care occupations</td>
<td>child teacher, care taker, social worker</td>
</tr>
<tr>
<td>15</td>
<td>Medical occupations</td>
<td>nurse, helper in nursing, receptionist</td>
</tr>
<tr>
<td>16</td>
<td>Physicians</td>
<td>physician, dentist, veterinarian</td>
</tr>
<tr>
<td>17</td>
<td>Teaching professions</td>
<td>professor, teacher</td>
</tr>
<tr>
<td>18</td>
<td>Artists/Athletes</td>
<td>graphic designer, musician, professional sportsman</td>
</tr>
<tr>
<td>19</td>
<td>Natural scientists</td>
<td>chemist, physicist, mathematician</td>
</tr>
<tr>
<td>20</td>
<td>Humanists</td>
<td>publicist, translator, librarian, economist</td>
</tr>
<tr>
<td>21</td>
<td>Others</td>
<td>laborer without further specification of activity</td>
</tr>
</tbody>
</table>

Source: Matthes et al. 2008[1], p.22.
respective labor market region. This high degree of self-supply emphasizes that the chosen delineation is highly suitable for identifying distinct regional labor markets in the framework of the matching theory. In addition, the commuter ratio, i.e. the ratio of commuters into and out of the respective labor market region is 98 percent on average and varies only little (standard deviation: 5.4 percentage points), especially in case of big labor market regions. Hence, as a consequence of the chosen delineation method there is no single labor market region left that could be considered as being a typical in- or out-commuter-region.

The third dimension of incongruence between demand and supply on the labor market is the occupational mismatch. Similar to the regional delineation described in the previous paragraph, it is necessary to identify occupational segments that are characterized both by a high within-homogeneity and by a high across-segregation. We follow Matthes et al. (2008) and use the 21 occupational segments shown in table 1. A big advantage of using occupational segments instead of occupations is that the degree of homogeneity varies less across segments which implies that the segments are better comparable with respect to the number of job alternatives within a segment. Furthermore, the occupational segments are characterized by a reasonably high degree of discriminatory power with respect to realized job mobility. To be more precise, in 95 percent of the segments the number of occupational changes within the respective segment exceeds the number of occupational changes that involve other segments.

All necessary data are made available with a very small time lag. This enables us to immediately calculate the monthly value of the mismatch index and to use it in the respective forecasting equations (as shown in section 3).

2.3 Measuring mismatch

Economists have been interested in measuring mismatch since the advent of the Beveridge curve. However, there is no clear-cut definition of mismatch in the literature. In fact, the mismatch concept is rather loose and as a consequence, a variety of indices and interpretations coexist on identical observable facts. Another concept (applied by Franz (1991), for instance) considers mismatch as being not only a temporary phenomenon. For instance, the mismatch index developed by Lilien (1982) is based on the assumption that short-run shocks can lead to a change in the composition of sectoral demand in an economy. Since labor markets adjust only slowly, mismatch occurs when both unemployment (in the contracting sectors) and vacancies (in the expanding sectors) temporarily rise to elevated levels during the period of transition. However, this approach has not been embraced too much in studies on European countries where economists have rather been searching for the causes of a more permanent increase in unemployment.

Another concept (applied by Franz (1991), for instance) considers mismatch as being not only a temporary phenomenon. It is built upon a disequilibrium model in which the short

---

3 For an overview of various concepts for measuring mismatch, see e.g. Schioppa (1991) or Canon et al. (2013).
side of any distinct labor market determines its level of employment. Under the assumption that the short side is not the same in all micro markets, unemployment and vacancies coexist at the aggregate level, which leads to employment being below the minimum of the aggregate supply and demand. As a consequence, a higher level of mismatch is attributable to a higher variance between the markets at the micro level.

A main disadvantage of the second approach to mismatch is that it rules out the coexistence of vacancies and unemployed within each micro market (usually called frictional unemployment). In contrast, however, the approach to mismatch embraced by Jackman and Roper (1987) accepts that frictional unemployment is inevitable within each distinct labor market. Therefore, incongruence at the micro level is measured relative to a more realistic (since attainable) size of unemployment. Hence, the respective mismatch index measures by how much structural unemployment contributes to total unemployment, i.e. "the proportion of unemployment attributable to structural imbalance" (Jackman and Roper (1987), p.14). It is this useful interpretation why we will work with mismatch indices in the style of Jackman and Roper (1987) in section 3.

To be precise, the mismatch index as proposed by Jackman and Roper (1987) is given by:

\[
MM_t = 1 - \prod_{i=1}^{I} \left[ \frac{V_{it}}{V_t} \cdot \frac{U_{it}}{U_t} \right]^{0.5},
\]

where the indices denote the micro market \( i \) and the time period \( t \), respectively.

If the ratio of unemployed equals that of vacancies in all micro markets, \( MM_t \) becomes zero which indicates that there is no structural imbalance at all. If for each \( i \), \( V_{it} \cdot U_{it} \) equals zero, there is either no vacancy for the unemployed or no unemployed person for the vacancies in any micro market. As a consequence, \( MM_t \) becomes one which means that 100 percent of unemployment is due to structural mismatch.

Figure 5 shows the development of the seasonally adjusted mismatch index covering all three dimensions of mismatch \( I = 3150 \) since January 2000. Due to missing or unusable data there is a gap in 2006. Along with this break comes a change in the way the individuals are classified into the different qualificatory groups so that the level of mismatch before this break cannot be compared to that after the break. Throughout the forecast evaluation performed in section 3, we control for this issue by using impulse dummies for all twelve months of 2006 and a level shift dummy starting in January 2007.

The development of the mismatch indicator is characterized by a relatively strong decline from about 0.17 in 2000 to roughly 0.13 in 2003, followed by a moderate increase in the two following years. The second half of our observation period is marked by fairly smooth up and down movements in the range between 0.10 and 0.14. Hence, approximately 10

---

4 To be more precise: It is their mismatch index \( I_3 \) (p.13) we refer to in this paragraph.

5 This interpretation holds under the assumption that the matching technology follows a Cobb-Douglas production function with constant returns to scale and equal elasticity of 0.5.
Figure 5: Development of the mismatch index, seasonally adjusted

Notes: The mismatch indicator is calculated according to equation (5), with $I = 3150$. Grey areas denote periods of recessions.

to 14 percent of unemployment in Germany can be attributed to regional, occupational and qualification-related mismatch.

It is striking that the development of the mismatch indicator seems to depend – at least to some extent – on the business cycle. The grey areas in figure 5 denote periods of recessions (identified by negative values of the year-on-year growth rate of the real GDP). We find that mismatch tends to decline in late stages of expansions and during recessions and to increase again after recessions. This pattern is especially pronounced in the years immediately before, during, and immediately after the Great Recession of 2008/2009. Given that structural imbalances on the labor market as measured by the mismatch index negatively affect matching efficiency, our results are in accordance with the findings of Barnichon and Figura (2011). In their study on the regression residual of the matching function, the authors find that matching efficiency tends to grow in the later stages of expansions and during recessions and to decrease in the aftermaths of recessions. In contrast, Sedlacek (2011) finds that matching efficiency tends to be procyclical which would not conform to what happened during our observation period. We argue that the way a recession influences the extent of mismatch is not unambiguous a priori. In principle, any downturn can lead to an increase in cyclical or structural unemployment, or to a combination of both. If unemployment increases are solely cyclical it is quite possible that mismatch declines which raises matching efficiency.

Another advantage of our detailed data set is that we can observe how the different types of mismatch develop over time and which dimension drives the development of the original mismatch indicator covering all three dimensions. For this purpose, we calculate the
Figure 6: The various dimensions of mismatch

Notes: The graph shows the seasonally adjusted development of mismatch accounting for occupational (O), qualificatory (Q), and regional (R) imbalance between unemployed and vacancies, or accounting for any combination of the three dimensions.

mismatch indicator of equation (5) using two or only one instead of three dimensions of mismatch.

Table 2: The dimensions of mismatch

<table>
<thead>
<tr>
<th>combination of dimensions</th>
<th>no. of micro markets</th>
</tr>
</thead>
<tbody>
<tr>
<td>region x occupation x qualification</td>
<td>3150</td>
</tr>
<tr>
<td>region x occupation</td>
<td>1050</td>
</tr>
<tr>
<td>region x qualification</td>
<td>150</td>
</tr>
<tr>
<td>occupation x qualification</td>
<td>63</td>
</tr>
<tr>
<td>region</td>
<td>50</td>
</tr>
<tr>
<td>occupation</td>
<td>21</td>
</tr>
<tr>
<td>qualification</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2 shows all seven possible versions of the mismatch indicator together with the size of the underlying matrices. One can see that the level of disaggregation, i.e. the number of micro markets taken into account in equation (5), varies from $I = 3$ to $I = 3150$. The development of the resulting mismatch indices is shown in figure 6. It depicts the seasonally adjusted development of mismatch accounting for occupational (O), qualificatory (Q), and regional (R) imbalance between unemployed and vacancies or for any combination of the
three dimensions. It demonstrates that taking into account an additional factor increases the value of the respective mismatch index at any point in time.\footnote{However, the dynamics do not necessarily add up exactly due to covariance between the three dimensions.}

The pronounced decline in mismatch in the early part of our observation period is solely attributable to a decreasing regional mismatch. Even since 2007, regional mismatch has declined whereas occupational and qualificatory mismatch have increased in tendency. This development is likely to be the consequence of a longstanding adjustment process through domestic migration out of structurally disadvantaged regions into prospering regions. Indeed, since German re-unification, and hence during all of our observation period, the net figures of internal migration have been negative for the eastern part of Germany, with the lion’s share of internal migrants moving to the southern federal states. Although this internal migration can help to approximate the shares of vacancies and unemployed across German labor market regions it does not automatically mitigate the problems of occupational or qualification-related incongruence.

\section{Forecast evaluation}

\subsection{Forecast evaluation setting}

This subsection tests from an out-of-sample perspective the hypothesis of a time varying efficiency parameter that negatively depends on structural imbalance on the micro labor markets. For this purpose, it compares the forecasting performance of the stock-flow matching model with constant efficiency parameter (introduced in subsection 2.1) to that of a model enhanced by $\mathcal{M} \mathcal{M}$, the mismatch indicator described in subsection 2.3.

As a consequence, the parsimonious benchmark model is nested in the larger model, which is of crucial importance in tests of equal predictive accuracy. Clark and West (2007) argue that the mean squared prediction error (MSPE) of the larger model is upward-biased due to additional noise stemming from the need to estimate a parameter which – under the null hypothesis of equal predictive performance – is zero in population\footnote{For a discussion of the difference between a null hypothesis of equal accuracy in the \textit{population vs. finite sample}, see e.g. Clark and McCracken (2009, 2012).} and which is correctly set to zero in the parsimonious model. In a sense, the smaller benchmark model is more efficient and hence benefits from not carrying the burden of estimating the parameter of a redundant variable to zero. Consequently, usual tests in the style of Diebold and Mariano (1995) are undersized and have poor power in a nested model environment. Therefore, we implement the nested-model test described in Clark and West (2007), applying a one-sided test for equal predictive accuracy with the alternative hypothesis being worse forecast performance of the nesting model. Since multiperiod-ahead forecast errors are usually autocorrelated, we use the heteroskedasticity and autocorrelation robust covariance estimator proposed in Newey and West (1987) in case of multiple-step forecasts. Inference based on asymptotic critical values – as proposed in McCracken (2004) or Clark and McCracken (2001) – might not be appropriate in case of small sample sizes. Therefore, the fixed regressor bootstrap method proposed in Clark and McCracken (2012a, b) is
implemented. We argue that bootstrapping considerably strengthens the validity of our test results. Furthermore, horizon-specific sets of critical values are implemented.

For computing multi-step forecasts we use direct, lead time-dependent forecasts. At least in theory, direct forecasts are more immune to model misspecification than iterated forecasts since they use the chosen model only once. In applying direct forecasts we avoid forecasting the mismatch indicator itself and modeling feedback effects to our target variable. Furthermore, the asymptotic theory of the nested model test we use in our application requires the forecasts to be linear functions of parameters which applies in direct forecasts but not in iterated approaches. Based on equation (4), the general lead time-dependent estimation specification for the natural logarithm of aggregate hires follows

\[
(1 - \rho_1 \cdot L + \rho_2 \cdot L^{f+1}) h_{t+f} = c + \alpha_{st} \cdot v_t + \beta_{st} \cdot u_t + \alpha_{fl} \cdot \dot{v}_t + \beta_{fl} \cdot \dot{u}_t + \gamma \cdot mm_t + \text{dummies} + \epsilon_{t+f},
\]

with \( f \) denoting the forecast horizon, \( L \) the lag-operator, and \( \epsilon \) the error term. Note that – in contrast to equation (4) – all variables on the right hand side are observed in \( t \), the time when the forecast is conducted. We do not index the coefficients by \( f \) for simplicity.

We include several shift dummies to capture changes in institutional settings or structural breaks in our target variable \( h \). In January 2005, for instance, the last stage of the Hartz labor market reforms came into effect. Along with this step there was a change in the way the unemployed are counted which lead to a sudden jump in the official unemployment figures (see figure 2) and, shortly afterwards, in the hiring figures (see figure 1). We find that a level-shift dummy that is zero before and 1 as from April 2005 is well suited to control for this issue. Furthermore, we account for the sudden decline in the hiring figures two years later (see figure 1) by including another level-shift dummy that takes on the value 1 as from January 2007. Since autoregressive lags are included in equation (6), the corresponding lags of the two dummy variables are employed, too.

As described in section 2 there is a short period of missing or unusable data with respect to the mismatch indicator (see figure 5). In case \( MM \) is included, we thus employ impulse dummies for all twelve months of 2006. The level-shift dummy of 2007 described above not only captures the break in our target variable but also in the mismatch indicator. The last dummy employed in equation (6) is an impulse dummy for January 2002, a month where it was necessary to impute the value of \( MM \) due to unusable data.

Since we are interested in forecasting the number of hires and not the natural logarithm of \( h \), we use the exponential to undo the log. In order to avoid a systematic underestimation

8 On the other hand, parameter estimates are more efficient in the iterated approach because it usually allows eliminating residual autocorrelation. Literature on this topic is ambiguous, ranging from emphasizing the advantages of direct forecasts (e.g. Klein (1968)) over mixed results (e.g Kang (2003)) to the finding of an empiric study on 170 U.S. macroeconomic variables that iterated forecasts typically outperform direct forecasts (Marcellino et al. (2006)).

9 Throughout the evaluation process we target forecasts of aggregate hiring figures (and hence not forecasts of the job finding rate).
of the expected value of $h$ due to the nonlinear transformation, we use the adjustment proposed by Wooldridge (2009):

$$\hat{h} = e^{\hat{\sigma}^2} \cdot e^{\log(h)},$$

(7)

where $\hat{\sigma}$ is the standard error of our regression stemming from equation (6). Since we cannot reject the null of $h$ having a log-normal distribution, we can be confident that the adjustment of equation (7) produces consistent predictions of the number of hirings.

Now we discuss the choice of the underlying parsimonious benchmark model. One could think of models relying solely on the own past such as AR(p)-models or the random walk (RW). In their GDP growth application, Clark and West (2007) use an AR(1) with constant as benchmark model, Clark and McCracken (2009) use models with just a constant in order to predict stock returns. Sometimes AR models of higher order, determined by in-sample information criteria, are employed. The Bayesian Information Criterion (BIC) with monthly hiring data from 2000 to 2007 (and hence excluding data from our evaluation period) gives evidence for using AR models with order not higher than 2.\textsuperscript{10} As expected from the very low persistence in the hirings variable (see figure 1), neither the information criteria nor the out-of-sample performance improve by forcing $\rho(L)$ to have a unit root (i.e. by modeling the first difference of hirings).

Our benchmark model does not include the mismatch indicator ($\gamma = 0$) and the dummy variables that are necessary to account for the data limitations inherent to $MM$ (see subsection 2.3).\textsuperscript{11} Since we use the direct approach, the model type changes with forecast horizon. For instance, the first model becomes an $AR(1)$ for 1-step-ahead forecasts and an $AR(6)$ without the first five lags for 6-step-ahead forecasts. As any direct f-step-ahead forecasting equation implies a $MA(f-1)$ error structure, we also considered the respective ARMA models. However, out-of-sample performance of these models turned out to be worse than that of their AR counterparts in most cases such that we do not report ARMA results.

It is not clear a priori how long it takes until changes in the incongruence between the profiles of supply and demand are fully incorporated in the matching efficiency. This is why we investigate all forecast horizons ranging from one to twelve months. We divide the sample into an estimation period which is updated for each iteration, and an evaluation period. The initial estimation period for 1-step-ahead forecasts ranges from March 2000 to February 2008 (96 observations) using data from January and February 2000 as initial observations. Our evaluation period ranges from March 2008 to April 2013 in case of 1-step-ahead forecasts and from February 2009 to March 2014 in case of 12-step-ahead forecasts. As a consequence, the evaluation period consists of 62 forecasts for all forecast horizons. The estimation period is regularly updated by adding the month that has

\textsuperscript{10} This result remains valid in case data after 2007 are included.

\textsuperscript{11} We also checked the out-of-sample performance of a benchmark model not including $MM$ itself, but including the corresponding dummy variables. However, forecast accuracy turned out to be worse in most cases.
become recently available (recursive scheme). Hence, the last estimation period ends in March 2013 for all forecast horizons. Each time the forecasts are calculated, the respective forecasting model is re-estimated first. Since we use lead time-dependent forecasts, the necessary number of initial observations differs across forecast horizons.

### 3.2 Performance of the standard mismatch indicator

This subsection treats the performance of the standard mismatch indicator \((I = 3150)\) in forecasting hirings. Table 3 shows the test results for the 1-, 2-, 6- and 12-step-ahead forecasts, respectively. The first column displays the forecast horizon. The second column shows the mean squared prediction errors (MSPEs) of the benchmark stock-flow matching model, whereas the third column displays the MSPEs of the alternative larger model enhanced by the mismatch indicator. In all cases, the MSPE of the benchmark model exceeds that of the larger model. Adjusted for the upward bias (fourth column), all reported test statistics are significantly positive at least at the 5 percent level. Hence, the test results show that the null hypothesis of equal predictive accuracy can be rejected and that models enhanced by the mismatch indicator outperform their benchmark counterparts. This supports our hypothesis that – from an out-of-sample perspective – models based on stock-flow matching can be improved even further by allowing for a time varying efficiency parameter and by enhancing them by an appropriate measure of structural imbalance on the micro labor markets. This potential for an improvement in forecast accuracy is verified for all investigated forecast horizons between 1 and 12 months.

<table>
<thead>
<tr>
<th>forecast horizon</th>
<th>( MSPE_1 )</th>
<th>( MSPE_2 )</th>
<th>adj. term</th>
<th>( \Delta MSPE_{adj} ) (test statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>204.6</td>
<td>174.7</td>
<td>14.1</td>
<td>43.9 (3.28)**</td>
</tr>
<tr>
<td>2 months</td>
<td>221.1</td>
<td>196.6</td>
<td>17.3</td>
<td>41.8 (1.65)**</td>
</tr>
<tr>
<td>6 months</td>
<td>254.2</td>
<td>141.9</td>
<td>126.0</td>
<td>238.4 (2.58)**</td>
</tr>
<tr>
<td>12 months</td>
<td>545.6</td>
<td>516.0</td>
<td>182.0</td>
<td>211.7 (1.80)**</td>
</tr>
</tbody>
</table>

**Notes:** \( MSPE_1 \) is the out-of-sample MSPE of the parsimonious benchmark model. \( MSPE_2 \) is the out-of-sample MSPE of the alternative larger model including the lagged mismatch indicator. adj. term is the adjustment term according to Clark and West (2007). \( \Delta MSPE_{adj} \) presents a point estimate of the adjusted difference in MSPEs with the respective test statistic in parentheses. All figures (except test statistics) are to be multiplied by \(10^6\). *, **, *** denote significance at the 10, 5, 1 percent level, respectively. Critical values are calculated following the fixed regressor bootstrap procedure proposed in Clark and McCracken (2012b) using 9,999 replications. The heteroskedasticity and autocorrelation robust covariance estimator proposed in Newey and West (1987) was used in case of multiple-step forecasts.

These findings are also supported by the respective in-sample results. For all forecasting horizons, we find \( \gamma \) to be negative and highly significant. The total effect of mismatch
Figure 7: Comparison of 6-months-ahead forecasts over time

Notes: The graph shows the number of actual hirings together with forecasts of hirings based on a benchmark stock-flow matching model and forecasts of hirings based on the same model enhanced by the mismatch indicator calculated according to equation (5) with $I = 3150$. Forecast horizon: 6 months. Grey areas denote periods of recessions.

The results do not differ substantially for other forecast horizons.
the Great Recession was an inflow into the pool of job searchers due to cyclical, not structural, reasons. The additional unemployment was not due to mismatch problems and thus laid the foundation for a rather quick and strong upswing. Figure 6 clarifies that it was mainly the occupational dimension that was responsible for the strong decline in structural imbalance at that time. This is in accordance with the finding that the enormous financial and economic crisis hit some sectors (and hence occupations), e.g. the export-dependent sectors in the southern and western parts of Germany, more than others.\footnote{And since the sectors and occupations are not equally distributed across the German regions, the regional component of mismatch was also affected to a certain extent, which can be seen in figure 6 as well.} As a consequence, the additional unemployed were exactly those needed to match the vacancies for the subsequent upswing.

To gain further intuition for the mismatch indicator, we shed light on the roles of the supply and the demand side. While mismatch could be driven by migrations on the supply side, also the composition of vacancies could have changed and hence contributed to the rather low level of mismatch at the end of the Great Recession. One approach to investigate which side of the labor market drives the dynamics of the mismatch indicator is to alternately hold constant one side and allow the other side to change in time.

Therefore, we define fictitious mismatch indices as follows:

\[
MM_t^U = 1 - \sum_{i=1}^{I} \left[ \frac{\bar{V}_i}{\bar{V}_t} \cdot \bar{U}_i \right]^{0.5}, \tag{8}
\]

\[
MM_t^V = 1 - \sum_{i=1}^{I} \left[ \frac{\bar{U}_i}{\bar{U}_t} \cdot \bar{V}_i \right]^{0.5}, \tag{9}
\]

where \(\bar{U}_i\) and \(\bar{V}_i\) are the average numbers of unemployed and vacancies in micro market \(i\), respectively. As a consequence, \(MM_t^U\) is the index that holds constant the supply side while \(MM_t^V\) holds constant the demand side and allows unemployment to change in time.

Figure 8 shows the development of the original mismatch indicator \(MM\) (blue line) since 2007\footnote{This period covers all values after the phase of missing data in 2006.} together with the two fictitious mismatch indicators. The red line shows the mismatch indicator freezing the distribution of unemployed at its average level according to equation (8). In contrast, the green line shows the analogous development holding constant the distribution of vacancies and hence allowing the distribution of unemployed to change (equation (9)).

Although both sides of the labor market contribute to the dynamics of structural imbalance, the changing distribution of the unemployed obviously dominated the development of mismatch in the last couple of years. This finding is supported by a remarkably high correlation between \(MM_t^V\) and \(MM\) of 0.910 whereas the correlation between \(MM_t^U\) and \(MM\) is only 0.719. This finding supports our reasoning that during the relevant evaluation period
The mismatch indicator was primarily driven by changes in the distribution of the unemployed, especially with respect to occupational imbalance, and that mismatch fell due to cyclical unemployment.

Furthermore, the results shed some light on the German job miracle: Not only did the firms make use of short-time work, hoarding and flexible working time. In addition, the fit of the stock of unemployed to the demand side of the labor market was improved. This might explain why the German labor market could regain its positive pre-crisis dynamics in the years 2010 and 2011, a period where a lowering speed due to the phasing-out of the Hartz-reform effects was already expected.

### 3.3 The importance of the different dimensions of mismatch

The mismatch indicator accounting for all three dimensions of mismatch ($I = 3150$) obviously provides the most complete picture of structural imbalance given the available data set. However, this does not necessarily mean that it performs best in out-of-sample forecasts. In theory, including an additional dimension can add valuable information or improve the lead-time properties of the respective measures of mismatch. However, it can also add nonessential information and hence lead to worse predictions of hirings. Therefore, this subsection investigates whether ignoring certain dimension harms or improves predictability of hirings, and if so at which forecast horizons this is the case.
Table 4: Forecast evaluation using different dimensions of mismatch

<table>
<thead>
<tr>
<th>dimensions</th>
<th>forecast horizon</th>
<th>f=1</th>
<th>f=2</th>
<th>f=6</th>
<th>f=12</th>
</tr>
</thead>
<tbody>
<tr>
<td>O x Q</td>
<td>MSPE₂</td>
<td>178.9</td>
<td>199.4</td>
<td>176.3</td>
<td>385.3</td>
</tr>
<tr>
<td></td>
<td>adj. term</td>
<td>22.7</td>
<td>26.6</td>
<td>44.3</td>
<td>85.8</td>
</tr>
<tr>
<td></td>
<td>∆MSPEadj</td>
<td>48.4</td>
<td>48.3</td>
<td>122.3</td>
<td>246.1</td>
</tr>
<tr>
<td></td>
<td>test statistic</td>
<td>2.94</td>
<td>2.26</td>
<td>3.53</td>
<td>3.47</td>
</tr>
<tr>
<td>O x R</td>
<td>MSPE₂</td>
<td>185.1</td>
<td>209.6</td>
<td>183.7</td>
<td>580.7</td>
</tr>
<tr>
<td></td>
<td>adj. term</td>
<td>7.7</td>
<td>14.2</td>
<td>91.1</td>
<td>263.0</td>
</tr>
<tr>
<td></td>
<td>∆MSPEadj</td>
<td>27.2</td>
<td>25.7</td>
<td>161.6</td>
<td>227.9</td>
</tr>
<tr>
<td></td>
<td>test statistic</td>
<td>2.62</td>
<td>1.22</td>
<td>2.17</td>
<td>1.33</td>
</tr>
<tr>
<td>Q x R</td>
<td>MSPE₂</td>
<td>173.1</td>
<td>197.6</td>
<td>203.5</td>
<td>702.6</td>
</tr>
<tr>
<td></td>
<td>adj. term</td>
<td>16.0</td>
<td>28.6</td>
<td>150.2</td>
<td>169.9</td>
</tr>
<tr>
<td></td>
<td>∆MSPEadj</td>
<td>47.5</td>
<td>52.2</td>
<td>200.9</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>test statistic</td>
<td>2.75</td>
<td>1.33</td>
<td>2.10</td>
<td>0.14</td>
</tr>
<tr>
<td>O</td>
<td>MSPE₂</td>
<td>188.6</td>
<td>215.2</td>
<td>198.2</td>
<td>395.9</td>
</tr>
<tr>
<td></td>
<td>adj. term</td>
<td>10.9</td>
<td>15.7</td>
<td>30.0</td>
<td>73.1</td>
</tr>
<tr>
<td></td>
<td>∆MSPEadj</td>
<td>26.9</td>
<td>21.6</td>
<td>82.0</td>
<td>222.8</td>
</tr>
<tr>
<td></td>
<td>test statistic</td>
<td>2.68</td>
<td>1.38</td>
<td>2.41</td>
<td>2.59</td>
</tr>
<tr>
<td>Q</td>
<td>MSPE₂</td>
<td>168.1</td>
<td>185.4</td>
<td>235.5</td>
<td>560.0</td>
</tr>
<tr>
<td></td>
<td>adj. term</td>
<td>17.1</td>
<td>24.1</td>
<td>9.3</td>
<td>27.8</td>
</tr>
<tr>
<td></td>
<td>∆MSPEadj</td>
<td>53.6</td>
<td>59.8</td>
<td>28.0</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td>test statistic</td>
<td>3.40</td>
<td>2.45</td>
<td>2.31</td>
<td>0.25</td>
</tr>
<tr>
<td>R</td>
<td>MSPE₂</td>
<td>192.2</td>
<td>216.6</td>
<td>243.3</td>
<td>784.5</td>
</tr>
<tr>
<td></td>
<td>adj. term</td>
<td>6.9</td>
<td>16.4</td>
<td>145.2</td>
<td>327.1</td>
</tr>
<tr>
<td></td>
<td>∆MSPEadj</td>
<td>19.2</td>
<td>20.9</td>
<td>156.1</td>
<td>88.2</td>
</tr>
<tr>
<td></td>
<td>test statistic</td>
<td>1.78</td>
<td>0.91</td>
<td>1.86</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Notes: The table shows the results of the forecast evaluation test following Clark and West (2007). The implemented mismatch indicators account for occupational (O), qualificatory (Q) or regional (R) imbalance between unemployed and vacancies, or for any combination of the three dimensions. MSPE₂ is the out-of-sample MSPE of the enhanced model including the mismatch indicator. adj. term is the adjustment term according to Clark and West (2007). ∆MSPEadj presents a point estimate of the adjusted difference in MSPEs. All figures (except test statistics) are to be multiplied by 10^6. *, **, *** denote significance at the 10, 5, 1 percent level, respectively. Critical values are calculated following the fixed regressor bootstrap procedure proposed in Clark and McCracken (2012b) using 9,999 replications. The heteroskedasticity and autocorrelation robust covariance estimator proposed in Newey and West (1987) was used in case of multiple-step forecasts.
Similar to table 3, table 4 shows the results of the forecast evaluation test following Clark and West (2007) for all investigated forecast horizons. Instead of using the mismatch indicator that accounts for all three dimensions of structural imbalance, two or even only one dimension is accounted for. We find that the mismatch indicator solely considering regional incongruence ($I = 50$) performs worst for all forecast horizons. This could be a consequence of the longtime loss of importance of the regional dimension.

For the very near future, i.e. in case of 1- or 2-step-ahead forecasts, the indicator covering qualificatory mismatch performs best as it leads to the lowest $MSPE$s and the highest test statistics. This result is rather astounding since the qualificatory approach separates the German labor market into three distinct micro markets only (see table 2).

For the more distant future, i.e. in case of forecasts 6 or even 12 months ahead, incorporating occupational mismatch as an additional factor seems to pay off as the combination of occupational and qualification-related mismatch produces the lowest $MSPE$s and the highest test statistics. Only in case of $f = 6$, the mismatch indicator covering all three dimensions (see table 3) produces a lower $MSPE$ of $141.9 \times 10^6$ although the respective test statistic is lower than in the case of the OxQ-combination.

To conclude, the results in this subsection emphasize the positive contribution of the occupational dimension for forecasts in the range of six to twelve months ahead. However, they also show that for the very short term, relying solely on qualificatory mismatch might be best with respect to forecasting.

4 Conclusion

This paper aimed at enhancing a basic stock-flow matching function by an appropriate measure of structural imbalance between the demand and supply side of micro labor markets. It investigated whether and to what extent the enhanced model performs better in out-of-sample forecasts of hirings.

For this purpose, we loosen the assumption of a constant efficiency parameter and allow matching efficiency to depend on the level of regional, qualificatory and occupational mismatch between unemployed and vacancies. In a sense, we go beyond the purely quantitative view of considering the plain number of vacancies and unemployed in our forecasting equations and add a qualitative perspective of how well the two groups match with respect to relevant categories.

Our data set reveals detailed empirical evidence on the size and development of mismatch in Germany in the past 13 years, including the years after the Hartz labor market reforms. We find a pronounced decline in regional mismatch, especially in the early part of our observation period. In contrast, both occupational and qualificatory mismatch have shown an increasing tendency.

\textsuperscript{15} We also checked whether including a time trend in equation (6) improves forecast accuracy. We find that this is only the case if the regional dimension is considered in $MM$, and then only for forecasts of the more distant future, i.e. in case of $f = 6$ and $f = 12$. 

IAB-Discussion Paper 16/2014
In pseudo out-of-sample tests that account for the nested model environment, we find that forecasting models enhanced by the mismatch indicator outperform their benchmark counterparts without indicator. This supports our hypothesis that – from an out-of-sample perspective – models based on stock-flow matching can be improved by allowing for a time varying efficiency parameter and based on a measure of structural imbalance on the micro labor markets. This potential for an improvement in forecast accuracy is verified for all invested forecast horizons between 1 and 12 months. It is especially pronounced in the aftermath of the Great Recession where a low level of mismatch improved the possibility of unemployed to find a job again.

Furthermore, we find that – although both sides of the labor market contribute to the dynamics of structural imbalance – the changing distribution of the unemployed dominated the development of mismatch in the last couple of years. This supports our reasoning that mismatch fell due to cyclical unemployment, i.e. that the persons becoming unemployed during the Great Recession were mostly those needed to match the demand side for the subsequent upswing. This might explain why unemployment could regain its strong pre-crisis dynamics in the years 2010 and 2011, a period where a lowering speed due to the phasing-out of the Hartz-reform effects was already expected.

Prospective research could benefit from extending the concept of measuring mismatch to the newly incoming vacancies and unemployed. Such an approach would take into account two measures of structural incongruence: One for the stock, and the other for the flow variables, which would probably complement the stock-flow matching model even better. Furthermore, one could investigate whether and to what extent further potential driving factors of the efficiency parameter (such as intensity of job search) can improve forecasts of hirings.
References


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2014</td>
<td>Hawranek, F.</td>
<td>Your very private job agency</td>
<td>1/14</td>
</tr>
<tr>
<td>2/2014</td>
<td>Kiesl, H., Drechsler, J.</td>
<td>Beat the heap</td>
<td>2/14</td>
</tr>
<tr>
<td>3/2014</td>
<td>Schäffler, J., Hecht, V., Moritz, M.</td>
<td>Regional determinants of German FDI in the Czech Republic</td>
<td>2/14</td>
</tr>
<tr>
<td>4/2014</td>
<td>Prantl, S., Spitz-Oener, A.</td>
<td>Interacting product and labor market regulation and the impact of immigration on native wages</td>
<td>2/14</td>
</tr>
<tr>
<td>5/2014</td>
<td>Kohlbrecher, B., Merkli, C., Nordmeier, D.</td>
<td>Revisiting the matching function</td>
<td>2/14</td>
</tr>
<tr>
<td>6/2014</td>
<td>Kopf, E., Zabel, C.</td>
<td>Active labour market programmes for women with a partner</td>
<td>2/14</td>
</tr>
<tr>
<td>7/2014</td>
<td>Rebien, M., Kubis, A., Müller, A.</td>
<td>Success and failure in the operational recruitment process</td>
<td>3/14</td>
</tr>
<tr>
<td>8/2014</td>
<td>Mendolicchio, C., Pietra, T.</td>
<td>On the efficiency properties of the Roy's model under asymmetric information</td>
<td>3/14</td>
</tr>
<tr>
<td>10/2014</td>
<td>Bossler, M.</td>
<td>Sorting within and across establishments</td>
<td>4/14</td>
</tr>
<tr>
<td>11/2014</td>
<td>Gillet, H., Pauser, J.</td>
<td>Efficiency in public input provision in two asymmetric jurisdictions with imperfect labour markets</td>
<td>4/14</td>
</tr>
<tr>
<td>12/2014</td>
<td>Antoni, M., Janser, M., Lehmer, F.</td>
<td>The hidden winners of renewable energy promotion</td>
<td>5/14</td>
</tr>
<tr>
<td>13/2014</td>
<td>Müller, S., Stegmaier, J.</td>
<td>Economic failure and the role of plant age and size</td>
<td>5/14</td>
</tr>
<tr>
<td>14/2014</td>
<td>Gärtnner, D., Grimm, V., Lang, J., Stephan, G.</td>
<td>Kollektive Lohnverhandlungen und der Gender Wage Gap</td>
<td>5/14</td>
</tr>
<tr>
<td>15/2014</td>
<td>Kruppe, T., Matthes, B., Unger, S.</td>
<td>Effectiveness of data correction rules in process-produced data</td>
<td>6/14</td>
</tr>
</tbody>
</table>

As per: 2014-06-18

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