

Vollmer, Sebastian; Holzmann, Hajo; Ketterer, Florian; Klasen, Stephan

**Working Paper**

## Distribution dynamics of regional GDP per employee in unified Germany

Diskussionsbeitrag, No. 461

**Provided in Cooperation with:**

School of Economics and Management, University of Hannover

*Suggested Citation:* Vollmer, Sebastian; Holzmann, Hajo; Ketterer, Florian; Klasen, Stephan (2010) : Distribution dynamics of regional GDP per employee in unified Germany, Diskussionsbeitrag, No. 461, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät, Hannover

This Version is available at:

<https://hdl.handle.net/10419/66023>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Distribution Dynamics of Regional GDP per Employee in Unified Germany

Sebastian Vollmer\*

Hajo Holzmann<sup>†</sup>

Florian Ketterer<sup>†</sup>

Stephan Klasen<sup>‡</sup>

Leibniz Universität Hannover, Discussion Paper No. 461

ISSN 0949-9962

## Abstract

We investigate to what extent convergence in production levels per worker has been achieved in Germany since unification. To this end, we model the distribution of GDP per employee across German districts using two-component normal mixtures. While in the first year after unification, the two component distributions were clearly separated and bimodal, corresponding to the East and West German districts, respectively, in the following years they started to merge showing only one mode. Still, using the recently developed EM-Test for homogeneity in normal mixtures, the hypothesis of just a single normal component for the whole distribution is clearly rejected for all years. A Posterior analysis shows that about half of the East German districts were assigned to the richer component in 2006, thus catching up to levels of the West. The growth rate of a mover district is about one percentage point higher than the growth rate of a non-mover district which had the same initial level of GDP per employee.

**JEL classification:** O47, R11.

**Keywords:** Regional convergence, distribution dynamics, mixture models, Germany.

**Acknowledgements.** The authors gratefully acknowledge financial support from the the German Research Foundation (HO 3260/3-1, KL 1260/9-1 and VO 1592/3-1), the German Academic Exchange Service (Vollmer), the Claussen-Simon Stiftung (Holzmann) and the Landesstiftung Baden-Württemberg (Holzmann).

---

\*Harvard University & University of Hannover. Harvard Center for Population and Development Studies, 9 Bow Street, Cambridge MA 02138, USA. Email: svollmer@hsph.harvard.edu.

<sup>†</sup>University of Marburg.

<sup>‡</sup>University of Göttingen.

# 1 Introduction

About twenty years have passed since the unification of the two German states. Many predictions on the economic integration and convergence of the two German states had been made at the time of the unification. Sinn and Sinn (1991) and Akerlof et al. (1991) correctly predicted a massive output collapse after unification, linked to the exchange rate used for monetary union in 1990 which rendered East German moribund industry entirely uncompetitive; both papers suggested that unless corrective measures were taken (and they were not), it would take a long time for East German to recover from this output shock.

Comparable living standard across regions are a fundamental objective of both German and European Union regional policy. Thus, fiscal transfers to the East of Germany continue to be sizeable. While most of the transfers are effectively tied to higher unemployment and higher poverty in the East and thus are used to raise consumption levels, some of these transfers are used to promote production there by funding investments in infrastructure, industrial policies, and the like. These policies should promote spatial convergence of production levels. On the other hand, new economic geography models would suggest that agglomeration tendencies in advanced economies might make it quite difficult for East Germany to attract and retain advanced industries which would work against spatial convergence. Given these possibly opposing forces, it is important to empirically test whether or not convergence of GDP per employee can be observed.

We investigate to what extent convergence in production levels per worker has been achieved in Germany since unification. To this end, we apply a distribution dynamics approach to the distribution of GDP per employee across German districts. This approach to convergence analysis has been introduced to the literature by Quah (1993, 1996). He interpreted the emergence of a bimodal cross-country distribution of GDP per capita as polarization of distribution into a rich and a poor convergence club and coined the term "twin peaks". Bianchi (1997) was the first to empirically confirm the statistical significance of the second peak using a nonparametric procedure by Silverman (1981). Colavecchio et al. (2010) apply Silverman's test to the regional distribution of GDP per capita in Germany for the period 1992–2001 and conclude that the distribution is bimodal or even trimodal. Jüssen (2008) applies Silverman's test to the regional distribution of GDP per employee for the period 1992–2004 and finds that an initially bimodal distribution turns into a unimodal distribution around 2002. From this he concludes convergence and does not further investigate how distributional dynamics have developed within the East.

Recently, Vollmer et al. (2010) pointed out that it is misleading to look at the number of peaks of a distribution if convergence clubs or sub-distributions are the true purpose

of the analysis. They show that simple rescaling of the data (e.g. taking logs) produces a statistically significant third peak in the cross-country distribution of GDP per capita. Countries which were previously assigned to Quah’s poor convergence club are considered middle-income on the log-scale, which introduces an arbitrary element in these analyses. Vollmer et al. (2010) model the cross-country distribution of GDP per capita with mixture models instead, because this approach is invariant to strictly monotonic transformation of the data and is thus robust towards this shortcoming of the twin peaks approach. Paap and van Dijk (1998) have pioneered the modeling of the cross-country distribution of GDP per capita with mixture models. Recent developments in the methodology of likelihood ratio tests that were not available to Paap and van Dijk (1998) allow us to determine the number and type of components with rigorous statistical testing. We adopt this approach to study the regional distribution of GDP per employee in post-unification Germany for the period 1992–2006.

We find that the regional distribution of GDP per employee in Germany is best described by a mixture of two normal distributions. In 1992, the two component distributions were clearly separated, corresponding to the East and West German districts, respectively. In the following years the two components started to merge, leading to a single mode but continuing to consist of two separate component distributions. A posterior analysis shows that 46 East German districts (out of 102) were assigned to the richer component in 2006, thus catching up to levels of the West (while only four districts from the West fell back to the poorer component). Interestingly, whether the East German districts move to the richer component or stay in the poor component does not depend on their initial level of GDP per employee.

## 2 Data & Methods

### 2.1 Data

Germany is structured into 16 states (Länder) and 429 districts (Kreise and kreisfreie Städte), 326 of these districts are located in former West Germany, 102 districts are located in former East Germany, Berlin is neither assigned to the East nor the West.<sup>1</sup> We use data from the German Federal Statistical Office (Statistisches Bundesamt) on nominal GDP per employee and prices to obtain our variable of interest, namely real GDP per employee for all districts. The data are available for the years 1992, 1994–2006. We exclude the 28 richest West German districts in 1992 from our analysis, since these stay way ahead of the rest of the Germany’s distribution of GDP per employee (forming their

---

<sup>1</sup> These figures predate a reform in Saxony which reduced the number of districts from 29 to 13. Berlin is just one district.

own group), and do not affect the main part of the distribution between east and west. Typically these districts include the central cities of major industrial centers in Western Germany where production is heavily concentrated and which pull in workers from a wide surrounding. They thus represent the extremes of concentrated production activities. However, these districts do not have any relevance for our research question, namely the catch-up process of East German districts to Western standards after unification.<sup>2</sup> Figure 1 shows the development of the mean and standard deviation for East and West Germany over time. Apparently we observe convergence, with GDP per employee in the East initially growing much faster than in the West and thus catching up. Since 1998 the gap between East and West has shrunk at a much lower rate than before. The overall standard deviation of GDP per employee decreased strongly in the first few years and stayed more or less constant since then. However, the standard deviation in the East increased since 1999 when the speed of convergence had already slowed down. Note that an assessment of income per capita would lead to quite different conclusions. On the one hand, due to much higher unemployment in the East and a lower share of working age people, the differential in income per capita between the East and West remains larger and has closed less than GDP per employee. On the other hand, the sizable transfer payments from West to East ensure that the incomes in the East are much higher than their output levels, sharply reducing the differential between East and West; as our focus is on convergence of labor productivity here, we focus on GDP/worker rather than GNI/capita.

## 2.2 Two-component normal mixtures

A natural way to model a heterogeneous population such as Germany's distribution of GDP per employee after unification is by finite mixture models. In a two-component normal mixture, the observations have density

$$f(x; \alpha, \mu_1, \mu_2, \sigma_1, \sigma_2) = (1 - \alpha)\phi(x; \mu_1, \sigma_1) + \alpha\phi(x; \mu_2, \sigma_2), \quad (1)$$

with  $0 \leq \alpha \leq 1$  and

$$\phi(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

---

<sup>2</sup> Specifically, we exclude Stuttgart, Landeshauptstadt, Kreisfreie Stadt; Böblingen, Landkreis; Heilbronn, Kreisfreie Stadt; Mannheim, Universitätsstadt, Kreisfreie Stadt; Ingolstadt, Kreisfreie Stadt; München, Landeshauptstadt, Kreisfreie Stadt; Altötting, Landkreis; Freising, Landkreis; München, Landkreis; Pfaffenhofen a.d.Ilm, Landkreis; Starnberg, Landkreis; Erlangen, Kreisfreie Stadt; Fürth, Kreisfreie Stadt; Hamburg; Frankfurt am Main, Kreisfreie Stadt; Offenbach am Main, Kreisfreie Stadt; Wiesbaden, Landeshauptstadt, Kreisfreie Stadt; Groß-Gerau, Landkreis; Hochtaunuskreis; Main-Taunus-Kreis; Wolfsburg, Kreisfreie Stadt; Wilhelmshaven, Kreisfreie Stadt; Düsseldorf, Kreisfreie Stadt; Rhein-Kreis Neuss, Kreis; Köln, Kreisfreie Stadt; Leverkusen, Kreisfreie Stadt; Rhein-Erft-Kreis; Ludwigshafen am Rhein, Kreisfreie Stadt from our analysis.

We assume without loss of generality that  $\mu_1 \leq \mu_2$ .  $\phi(x; \mu_1, \sigma_1)$  and  $\phi(x; \mu_2, \sigma_2)$  correspond to the distributions of the two sub-populations, and  $\alpha$  and  $1 - \alpha$  are interpreted their relative sizes.

We fit two-component normal mixtures to the log data. In order to check that the parametric components are well-specified, we investigate the log-data in 1992 for East and West separately, since in this year the distributions were clearly separated. We apply Shapiro-Wilk's (SW) and Anderson-Darling's (AD) tests to check whether normality can be rejected, yielding SW  $p$ -value East: 0.87; AD  $p$ -value East: 0.91; SW  $p$ -value West: 0.82; AD  $p$ -value West: 0.60. Hence we conclude that a mixture of two normal distributions fits the log-data well. Note that it is essential to set up a joint model for the two populations, since we want to investigate convergence within the complete distribution of GDP per employee in Germany.

The parameters  $\alpha$ ,  $\mu_1$ ,  $\mu_2$ ,  $\sigma_1$  and  $\sigma_2$  are estimated from the data by a likelihood based method. We allow for unequal variances  $\sigma_1^2$  and  $\sigma_2^2$ , because a likelihood ratio test shows that the simplifying assumption of equal variances does not hold for all years.

Let  $X_1, \dots, X_n$  denote independent, identically distributed observations with densities (1). The log-likelihood

$$\mathcal{L}_n(\alpha, \mu_1, \mu_2, \sigma_1, \sigma_2) = \sum_{i=1}^n \log((1 - \alpha)\phi(X_i; \mu_1, \sigma_1) + \alpha\phi(X_i; \mu_2, \sigma_2)).$$

in finite normal mixtures with different variances is unbounded, since for any given  $n$ ,  $\mathcal{L}_n(\alpha, \mu_1, \mu_2, \sigma_1, \sigma_2) \rightarrow \infty$ , if  $X_1 = \mu_1$  and  $\sigma_1 \rightarrow 0$ , holding the other parameters fixed. Thus, a global maximizer of the likelihood function does not exist. There are some formal ways around this problem, e.g. choose the largest local maximum or restrict the possible variances by restrictions of the form  $\sigma_2^2 \leq c\sigma_1^2$  and  $\sigma_1^2 \leq c\sigma_2^2$  for some  $c > 1$  (cf. Hathaway 1985), which again leads to the existence of a global maximum. We found that using reasonable starting values (which are easy to obtain in our problem by considering East and West German districts separately), maximization algorithms such as EM or quasi Newton found stable local maxima of the log-likelihood function.

In order to formally investigate whether the two components in Germany's distribution of GDP per employee finally merged, one can test in model (1) whether it effectively consists of just a single component. This amounts to testing the hypothesis

$$H_0 : \alpha(1 - \alpha) = 0 \text{ or } (\mu_1, \sigma_1^2) = (\mu_2, \sigma_2^2)$$

against the full model (1). This turns out to be a quite difficult parametric testing problem, see Chen and Chen (2003) for some history. In the following we present a novel

approach, the EM-test by Chen and Li (2009) for normal mixtures in mean and variance parameters, which overcomes many drawbacks of the simple likelihood ratio test for the same problem. The test by Chen and Li (2009) is based on a penalized log-likelihood function

$$pl_n(\alpha, \mu_1, \mu_2, \sigma_1, \sigma_2) = \mathcal{L}_n(\alpha, \mu_1, \mu_2, \sigma_1, \sigma_2) + p(\alpha) + p_n(\sigma_1) + p_n(\sigma_2),$$

Here,  $p : [0, 1] \rightarrow \mathbb{R}$  is a continuous function that is maximized at  $\alpha = 0.5$  and tends to negative infinity as  $\alpha$  goes to 0 or 1 and  $p_n : [0, \infty) \rightarrow \mathbb{R}$  is bounded, when  $\sigma$  is large, but tends to negative infinity when  $\sigma$  goes to 0. The test statistic is computed as follows:

**Step 0** Choose a set of initial  $\alpha$  values, say  $\alpha_1, \alpha_2, \dots, \alpha_J$  and a positive integer  $K$ . Compute

$$(\hat{\mu}_0, \hat{\sigma}_0) = \arg \max_{\mu, \sigma} pl_n(0.5, \mu, \mu, \sigma, \sigma).$$

Let  $j = 1, k = 0$ .

**Step 1** Let  $\alpha_j^{(k)} = \alpha_j$ .

**Step 2** Compute

$$(\mu_{j1}^{(k)}, \mu_{j2}^{(k)}, \sigma_{j1}^{(k)}, \sigma_{j2}^{(k)}) = \arg \max_{\mu_1, \mu_2, \sigma_1, \sigma_2} pl_n(\alpha_j^{(k)}, \mu_1, \mu_2, \sigma_1, \sigma_2)$$

**Step 3** For  $i = 1, 2, \dots, n$ , compute the weights

$$w_{ij}^{(k)} = \frac{\alpha_j^{(k)} \phi(X_i; \mu_{j2}^{(k)}, \sigma_{j2}^{(k)})}{(1 - \alpha_j^{(k)}) \phi(X_i; \mu_{j1}^{(k)}, \sigma_{j1}^{(k)}) + \alpha_j^{(k)} \phi(X_i; \mu_{j2}^{(k)}, \sigma_{j2}^{(k)})}.$$

and then use the  $M$ -step to update the parameters

$$\alpha_j^{(k)} = \arg \max_{\alpha} \left( (n - \sum_{i=1}^n w_{ij}^{(k)}) \log(1 - \alpha) + \sum_{i=1}^n w_{ij}^{(k)} \log(\alpha) + p(\alpha) \right)$$

and

$$(\mu_{j1}^{(k+1)}, \mu_{j2}^{(k+1)}, \sigma_{j1}^{(k+1)}, \sigma_{j2}^{(k+1)}) = \arg \max_{\mu_1, \mu_2, \sigma_1, \sigma_2} \sum_{l=1}^2 \left( \sum_{i=1}^n w_{ij}^{(k)} \log\{\phi(X_i; \mu_l, \sigma_l)\} + p_n(\sigma_l) \right)$$

Let  $k = k + 1$  and repeat Step 3, until  $k = K$ .

**Step 4** Let  $j = j + 1, k = 0$  and go to Step 1, until  $j = J$ .

**Step 5** Calculate

$$EM_n^{(K)} = \max\{M_n^{(K)}(\alpha_j), j = 1, 2, \dots, J\}$$

where

$$M_n^{(K)}(\alpha_j) = 2\{pl_n(\alpha_j^{(K)}, \mu_{j1}^{(K)}, \mu_{j2}^{(K)}, \sigma_{j1}^{(K)}, \sigma_{j2}^{(K)}) - pl_n(0.5, \hat{\mu}_0, \hat{\mu}_0, \hat{\sigma}_0, \hat{\sigma}_0)\}$$

Chen and Li (2009) show that under the null hypothesis  $H_0$ , if one of the  $\alpha_j$ 's is equal to 0.5, then as  $n \rightarrow \infty$ ,

$$EM_n^{(K)} \xrightarrow{d} \chi_2^2.$$

As parameters of the EM-test, following the recommendations in Chen and Li (2009) we choose  $p(\alpha) = \log(1 - |1 - 2\alpha|)$  and  $p_n(\sigma) = -0.25\{s_n/\sigma^2 + \log(\sigma^2/s_n)\}$  where  $s_n = \sum_{i=1}^n (X_i - \bar{X})^2/n$ . Further we choose  $\{\alpha_1, \alpha_2, \alpha_3\} = \{0.1, 0.3, 0.5\}$  and  $K = 3$ .

One advantage of modeling (log) GDP per employee is that we can relate the estimated density to any sub-populations. That means that we can use mixture models for discriminant analysis, see e.g. Fraley and Raftery (2002). Once we have fitted the two-component normal mixture

$$f(x; \bar{\alpha}, \bar{\mu}_1, \bar{\mu}_2, \bar{\sigma}_1, \bar{\sigma}_2) = (1 - \bar{\alpha})\phi(x; \bar{\mu}_1, \bar{\sigma}_1) + \bar{\alpha}\phi(x; \bar{\mu}_2, \bar{\sigma}_2)$$

to the data (here  $\bar{\alpha}, \bar{\mu}_1, \bar{\mu}_2, \bar{\sigma}_1$  and  $\bar{\sigma}_2$  denote the EM-estimators, maximizing  $EM_n^{(K)}$ ), each observation  $X_i$  can be assigned the posterior probabilities

$$p(1; X_i) = \frac{(1 - \bar{\alpha})\phi(X_i; \bar{\mu}_1, \bar{\sigma}_1)}{f(X_i; \bar{\alpha}, \bar{\mu}_1, \bar{\mu}_2, \bar{\sigma}_1, \bar{\sigma}_2)}, \quad p(2; X_i) = 1 - p(1; X_i),$$

which give the probability that  $X_i$  belongs to the corresponding component in the mixture model. One may then assign  $X_i$  to one of the components by using the maximum a-posterior estimate (MPE), which assigns the  $j \in \{1, 2\}$  to district  $X_i$  for which  $p(j; X_i)$  is maximal.

## 3 Results

### 3.1 The Regional Distribution of GDP per Employee

Figure 2 shows the fitted normal mixture and a kernel density estimate of the regional distribution of GDP per employee for the first and last year of our analysis, 1992 and 2006, for which we chose the bandwidth according to Silverman's rule of thumb (Silverman, 1986). Figures for all other years can be found in the appendix (c.f. Figures 5 and 6). It is apparent that we observe two quite distinct distributions in 1992, which have moved considerably together by 2006. The pictures show two modes and two distributions in 1992, in 2006 there is only a single mode and it is not obvious whether the population



consists of two sub-populations or just a single one. Therefore, we test both for the number of modes and for the number of components for all years. Silverman (1981) introduced a nonparametric test for the hypothesis that a density function has  $k$  modes against the alternative that it has more than  $k$  modes, Bianchi (1997) was the first to apply this test to income distributions. We refer to these two papers for a detailed description of the test. Arguably, the more important feature of the distribution is the number of components that generate this distribution rather than the number of modes. To test for the number of components here, we will apply the EM test as discussed above.

Silverman's test rejects unimodality in favor of bimodality from 1992 to 2000. In 2001 the hypothesis of unimodality can only be rejected at the 10 percent level, from 2002 on it cannot be rejected anymore. However, the EM test clearly finds two different components for all years (no matter which level of confidence we apply, the first three digits of the p-values are always zero).

The interpretation of this finding is, that levels of GDP per employee in East and West have moved close enough together so that the two underlying distributions do no longer result in separate modes. However, the complete distribution of GDP per employee continues to be best described by a mixture of two separate underlying distributions. Jüssen (2008) exclusively focuses on the number of modes in the distribution and therefore concludes convergence at this point. His study on the number of modes misses the point that the distribution is still generated by two components which cannot be revealed with Silverman's test. In addition, looking at the dynamics within these two components turns out to be fruitful, to which we turn now.

## 3.2 Posterior Analysis

The posterior analysis confirms the visual first impression that in 1992 all East German districts belonged to the first (poorer) component and all West German districts belonged to the second (richer) component (based on the MPE). Over the years, 46 East German districts moved up from the first to the second component, and 4 West German districts, namely Aurich, Friesland, Wittmund (all in Lower Saxony), and Bottrop (in Northrhine Westphalia) fell back from the second to the first component.

In the remainder of the paper we will focus on the East districts which moved up from the first to the second component and on those which failed to do so. Figure 4 shows a map of Germany where all Eastern districts which moved up to the second component are colored black and the four Western districts which moved down to the first component are colored gray. Table 2 lists all Eastern districts that moved up to the second component with growth rates and ranks in 1992 and 2006 (within the East). We would expect that initial conditions, i.e. the GDP per employee in 1992, should at least to some extent

determine whether or not a district catches up. In particular, if an East German district is initially closer to the second component than other East German districts, and we assume that all districts uniformly converge to a common steady state, then it should be able to close the gap to the second component prior to the districts which initially are farther away.

The empirical evidence contradicts our expectations. Table 3 shows the results of a probit regression with a mover dummy (coded equal to 1 if a district moves from the first to the second component between 1992 and 2005) as the dependent variable. Districts that have a common border with the Czech Republic or Poland were 25 percent less likely to move to the second component than other East German districts. The coefficient for a common border with West Germany turned out to be positive but insignificant. But surprisingly, the experience of catching-up to the second component cannot be explained by initial levels of GDP per employee as this variable is insignificant in the regression. Considering the two most extreme examples, Märkisch-Oderland was the second richest district in the East in 1992, but it did not move to the second component by 2006. This implies that all 46 districts, which moved to the second component, (and also some others) are ranked higher than Märkisch-Oderland in 2006. On the other extreme, Wartburgkreis was the sixth poorest district in 1992, but it nevertheless managed to catch-up to the West, in fact it had the highest average annual growth rate of all districts, both East and West.

One also can approach the question of convergence using another technique. Following the classical Barro and Sala-I-Martin (1991, 1992) framework, we find  $\beta$ -convergence among all districts, but also within East and West respectively (c.f. Figure 3 (left)). However, the growth rate of a mover district in the East is about one percentage point higher than the growth rate of a non-mover district which had the same GDP per employee in 1992 (c.f. Figure 3 (right)). Figure 1 shows that the standard deviation of GDP per employee across East German districts increases; in other words we do not find  $\sigma$ -convergence in the East. However, we observe  $\sigma$ -convergence within the first component. As reported in Table 1, the  $\sigma$ -parameter of the first component decreases over time (which is mostly due to the fact that the more successful districts move up to the second component). These findings suggest that there are two distinct convergence clubs for GDP per employee in the East of Germany.

This divergence pattern is in contrast to the predictions of the basic neoclassical growth model which, assuming free mobility of capital and labor, would lead to convergence of GDP per employee. Instead it appears that agglomeration tendencies outweigh these processes and prevent greater convergence for the majority of districts. Schäfer and Steger (2010) set up a dynamic macroeconomic model that is motivated by the process of eco-

conomic development in Eastern Germany. In their model they describe a small open economy with factor mobility and aggregate increasing returns to scale. The model includes multiple equilibria as well as indeterminacy. Expectations matter for the equilibrium dynamics. Their model can explain our findings as well as some other stylized facts.

## 4 Conclusion

We find that the regional distribution of GDP per employee in Germany is best described by a mixture of two normal distributions that is twin peaked in 1992. The two components move closer together to a single peaked distribution but continue to consist of two separate component distributions over the entire observation period. Our analysis is based on a parametric mixture model, which is obtained through rigorous testing of all assumptions involved.

In a posterior analysis we have also identified the East German districts which converged to West German levels of GDP per employee since unification. Movements to the second component cannot be explained by initial levels of GDP per employee. The annual growth rate of a mover district is about one percentage point higher than the growth rate of a non-mover district which had the same level of GDP per employee in 1992. We thus conclude that districts in the East converge to two different steady states that are independent of their initial levels of GDP per employee.

Districts that share a border with the Czech Republic or Poland do systematically worse than other East German districts. We are thus quite optimistic that the inclusion of the Czech Republic and Poland to the European Union and a continuing European Integration will be beneficial for the East German districts which are currently lagging behind. Their geographic location might, in due time, turn into an advantage in an integrated Europe.

## References

- Akerlof, G., A. Rose, J. Yellen, and H. Hesselius (1991) East Germany in from the Cold: The Economic Aftermath of Currency Union. *Brookings Papers on Economic Activity* **22**, 1–106.
- Barro, R. and X. Sala-i-Martin (1991) Convergence across States and Regions. *Brookings Papers on Economic Activity* **22**, 107–182.
- Barro, R. and X. Sala-i-Martin (1992) Convergence. *Journal of Political Economy* **100**, 223–251.
- Bianchi, M. (1997) Testing for Convergence: Evidence from Non-Parametric Multimodality Tests. *Journal of Applied Econometrics* **12**, 393–409.
- Chen, H. and J. Chen (2003). Tests for homogeneity in normal mixtures with presence of a structural parameter. *Statistica Sinica* **13**, 351-365.
- Chen, J. and P. Li. (2009) Hypothesis test for Normal Mixture Models: the EM Approach. *Annals of Statistics* **37**, 2523-2542.
- Colavecchio, R., D. Curran and M. Funke (2010) Drifting Together or Falling Apart? The Empirics of Regional Economic Growth in Post-Unification Germany. *Applied Economics* forthcoming.
- Fraley, C. and A. Raftery (2002) Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association* **97**, 611–631.
- Hathaway, R. J. (1985) A constrained formulation of maximum-likelihood estimation for normal mixture distributions. *Annals of Statistics* **13**, 795-800.
- Jüssen, F. (2008) A distribution dynamics approach to regional GDP convergence in unified Germany. *Empirical Economics* **37**, 627–652.
- Paap, R. and H. K. van Dijk (1998) Distribution and mobility of wealth of nations. *European Economic Review* **42**, 1269-1293.
- Quah, D.T. (1993) Galton’s fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics* **95**, 427-443.
- Quah, D.T. (1996) Empirics for economic growth and convergence. *European Economic Review* **40**, 1353-1375.
- Schäfer, A. and T. Steger (2010) History, Expectations, and Public Policy: Economic Development in Eastern Germany. Discussion paper. University of Leipzig.
- Silverman, B. (1981) Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society Series B* **43**, 97–99.
- Silverman B. (1986) *Density estimation for statistics and data analysis*. Chapman and Hall, New York.

Sinn, G. and H.-W. Sinn (1991) *Kaltstart. Volkswirtschaftliche Aspekte der Deutschen Vereinigung*. Mohr-Siebeck, Tübingen.

Vollmer, S., H. Holzmann and J. Weisbrod (2010) Peaks vs. Components. Discussion Paper. University of Hannover.

Table 1: Estimated parameters, p-values of the EM test and p-values of Silverman's test.

Year	$\hat{\alpha}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	p EM	p 1vs2modes	p 2vs3modes
1992	0.743	4.42	4.67	0.0455	0.0385	< 0.01	< 0.01	0.79
1994	0.743	4.52	4.68	0.0372	0.0358	< 0.01	< 0.01	0.47
1995	0.738	4.54	4.68	0.0354	0.0356	< 0.01	< 0.01	0.42
1996	0.731	4.56	4.69	0.0308	0.0333	< 0.01	< 0.01	0.89
1997	0.726	4.57	4.69	0.0318	0.0323	< 0.01	< 0.01	0.97
1998	0.731	4.57	4.70	0.0329	0.0333	< 0.01	< 0.01	0.89
1999	0.720	4.59	4.70	0.0328	0.0310	< 0.01	< 0.01	0.22
2000	0.737	4.59	4.71	0.0322	0.0321	< 0.01	< 0.01	0.55
2001	0.716	4.61	4.71	0.0359	0.0336	< 0.01	0.05	0.70
2002	0.707	4.62	4.71	0.0343	0.0320	< 0.01	0.96	0.68
2003	0.700	4.63	4.72	0.0318	0.0334	< 0.01	0.76	0.72
2004	0.711	4.63	4.72	0.0328	0.0343	< 0.01	0.40	0.40
2005	0.875	4.61	4.71	0.0204	0.0454	< 0.01	0.15	0.72
2006	0.868	4.62	4.72	0.0219	0.0449	< 0.01	0.15	0.87

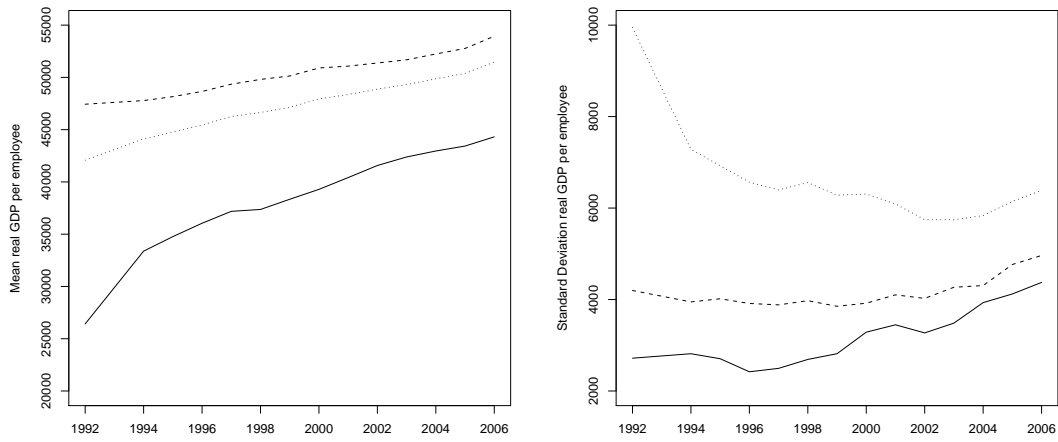


Figure 1: Mean and Standard Deviation of GDP per employee for all districts (dotted line), districts from the East (solid line) and districts from the West (dashed line).

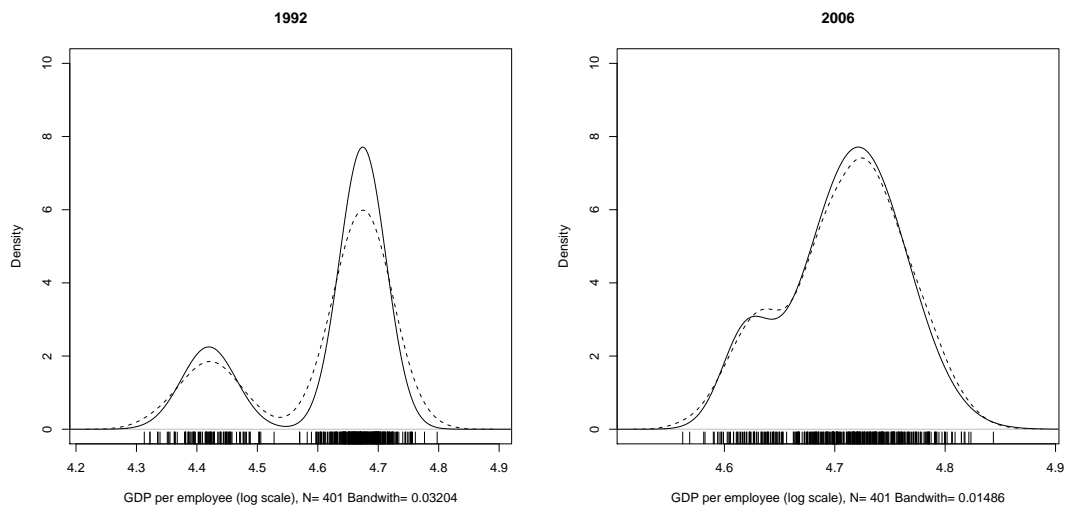


Figure 2: Distribution of GDP per employee in 1992 and 2006. Our model (solid line) compared to a kernel density estimator (dashed line).

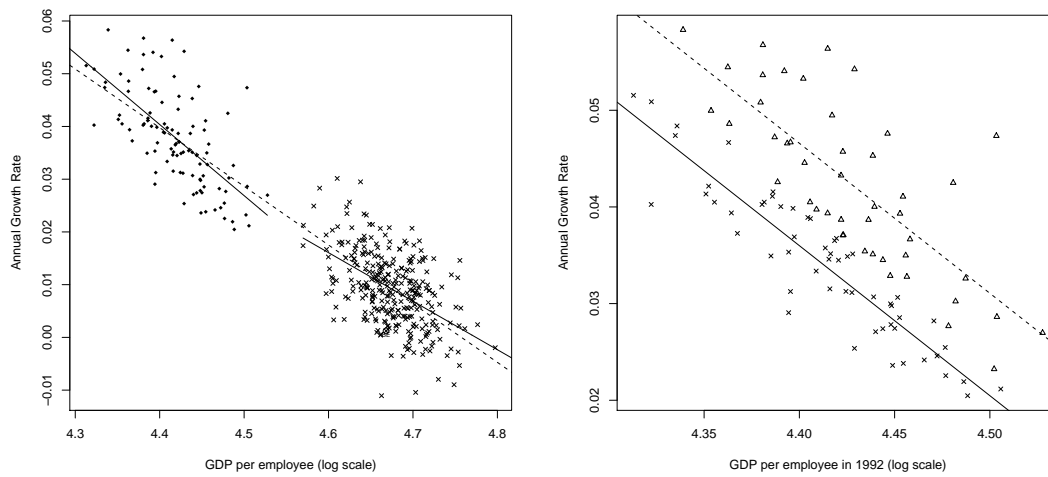


Figure 3: Left:  $\beta$ -convergence in Germany (dashed line) and East and West respectively (solid lines). Right: Two convergence clubs in the East, movers (triangles) and non-movers (crosses).

Table 2: East German districts that caught up to the West

County Name	Rank 1992	Rank 2006	Year(s) moved	Growth Rate
Brandenburg (out of 18)				
Dahme-Spreewald, Landkreis	4	1	1996-2005	4.74
Spree-Neie, Landkreis	10	4	1995	4.25
Oberspreewald-Lausitz, Landkreis	21	11	2003-2005	4.11
Oder-Spree, Landkreis	48	13	1999	4.57
Ostprignitz-Ruppin, Landkreis	50	31	2006	3.87
Uckermark, Landkreis	55	9	2000	4.95
Teltow-Fläming, Landkreis	75	25	1999	4.66
Oberhavel, Landkreis	84	7	2001	5.68
Prignitz, Landkreis	90	33	2005	4.86
Mecklenburg-Vorpommern (out of 18)				
Rostock, Kreisfreie Stadt	7	18	1998-2004	3.26
Nordwestmecklenburg, Landkreis	11	43	2005	2.77
Ludwigslust, Landkreis	30	42	2005	3.29
Wismar, Kreisfreie Stadt	31	5	2002-2004	4.76
Güstrow, Landkreis	32	34	2005	3.45
Demmin, Landkreis	68	6	2003	5.33
Saxony (out of 29)				
Niederschlesischer Oberlausitzkreis	1	14	1994-2004	2.7
Kamenz, Landkreis	3	23	2005	2.86
Delitzsch, Landkreis	5	46	2005	2.32
Chemnitzer Land, Landkreis	9	30	2004	3.02
Muldentalkreis	19	29	1994-2005	3.5
Dresden, Kreisfreie Stadt	35	21	2003	4.00
Meien, Landkreis	38	37	2005	3.51
Mittweida, Landkreis	39	28	2005	3.87
Plauen, Kreisfreie Stadt	47	41	2006	3.71
Leipziger Land, Landkreis	51	20	2005	4.33
Döbeln, Landkreis	71	24	2005	4.67
Freiberg, Landkreis	85	12	2005	5.36
Riesa-Groenhain, Landkreis	92	16	2005	5.45
Saxony-Anhalt (out of 14)				
Burgenlandkreis	17	22	2002-2005	3.67
Jerichower Land	22	15	2002-2005	3.93
Börde	37	10	2001	4.53
Stendal	40	40	2005	3.54
Saalekreis	59	3	2000	5.64
Altmarkkreis Salzwedel	60	36	2005	3.93
Salzlandkreis	62	39	2005	3.97
Wittenberg	65	38	2005	4.05
Harz	67	26	2005	4.46
Anhalt-Bitterfeld	87	19	2003-2005	5.08
Thuringia (out of 23)				
Weimarer-Land, Kreis	18	32	2005	3.28
Sömmerda, Kreis	41	2	2000	5.42
Saale-Orla-Kreis	49	44	2005	3.71
Jena, Kreisfreie Stadt	76	8	2003	5.41
Altenburger Land, Kreis	78	45	2006	4.26
Saalfeld-Rudolstadt, Kreis	79	27	2005	4.72
Ilm-Kreis	94	35	2005	5.00
Wartburgkreis	97	17	2005	5.83



Table 3: Probit regression predicting movement from the first to the second component

Variable	Marginal effect (standard error)
GDP per employee in 1992 (log)	1.873 (1.161)
Border East	-0.249** (0.114)
Border West	0.071 (0.137)

\*\* denotes statistical significance at the 5 percent level.

\* denotes statistical significance at the 10 percent level.

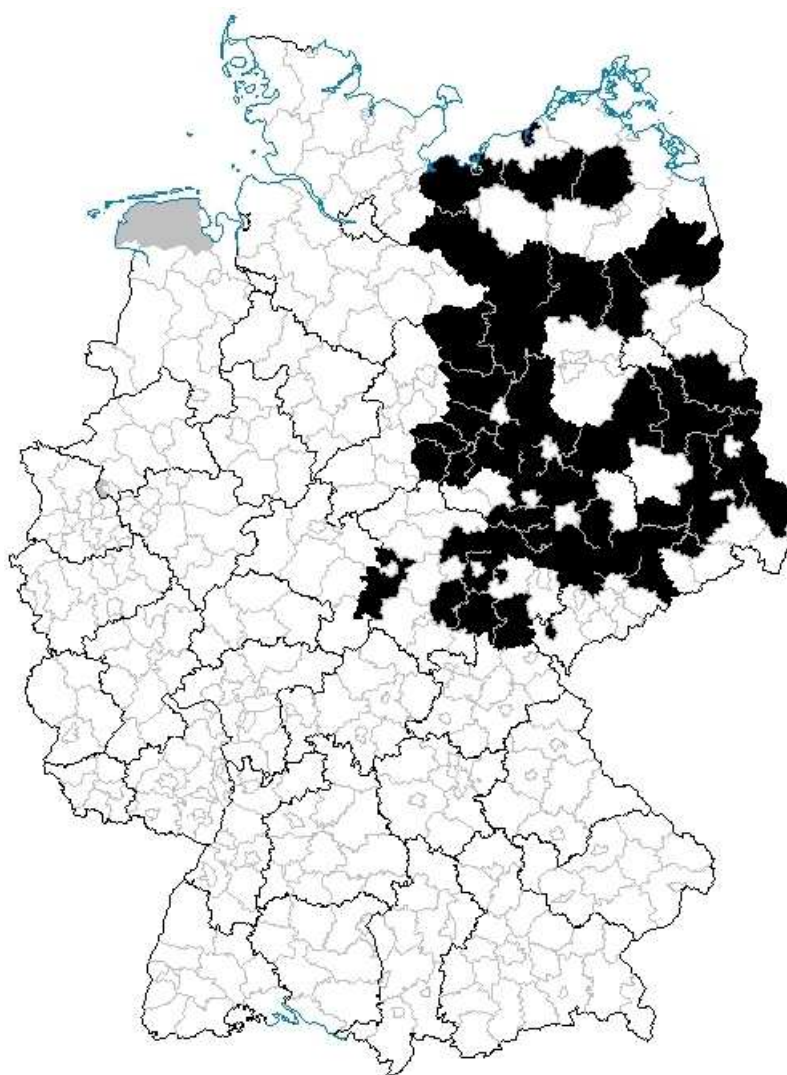


Figure 4: Map of Germany. East German districts that catch up to the West are colored black, and the four West German districts that fell back are colored gray.

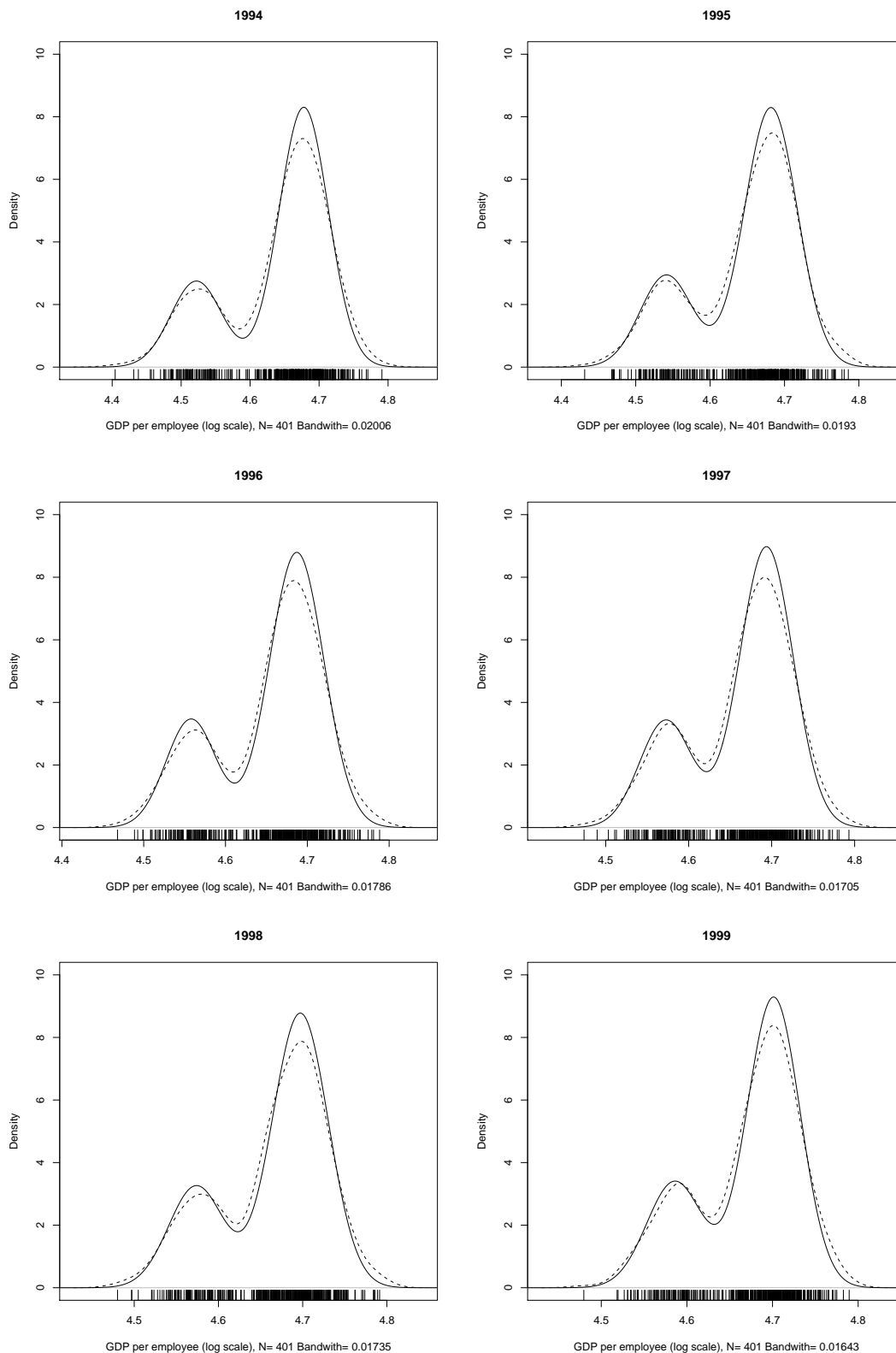


Figure 5: Distribution of GDP per employee between 1994 and 1999. Our model (solid line) compared to a kernel density estimator (dashed line).

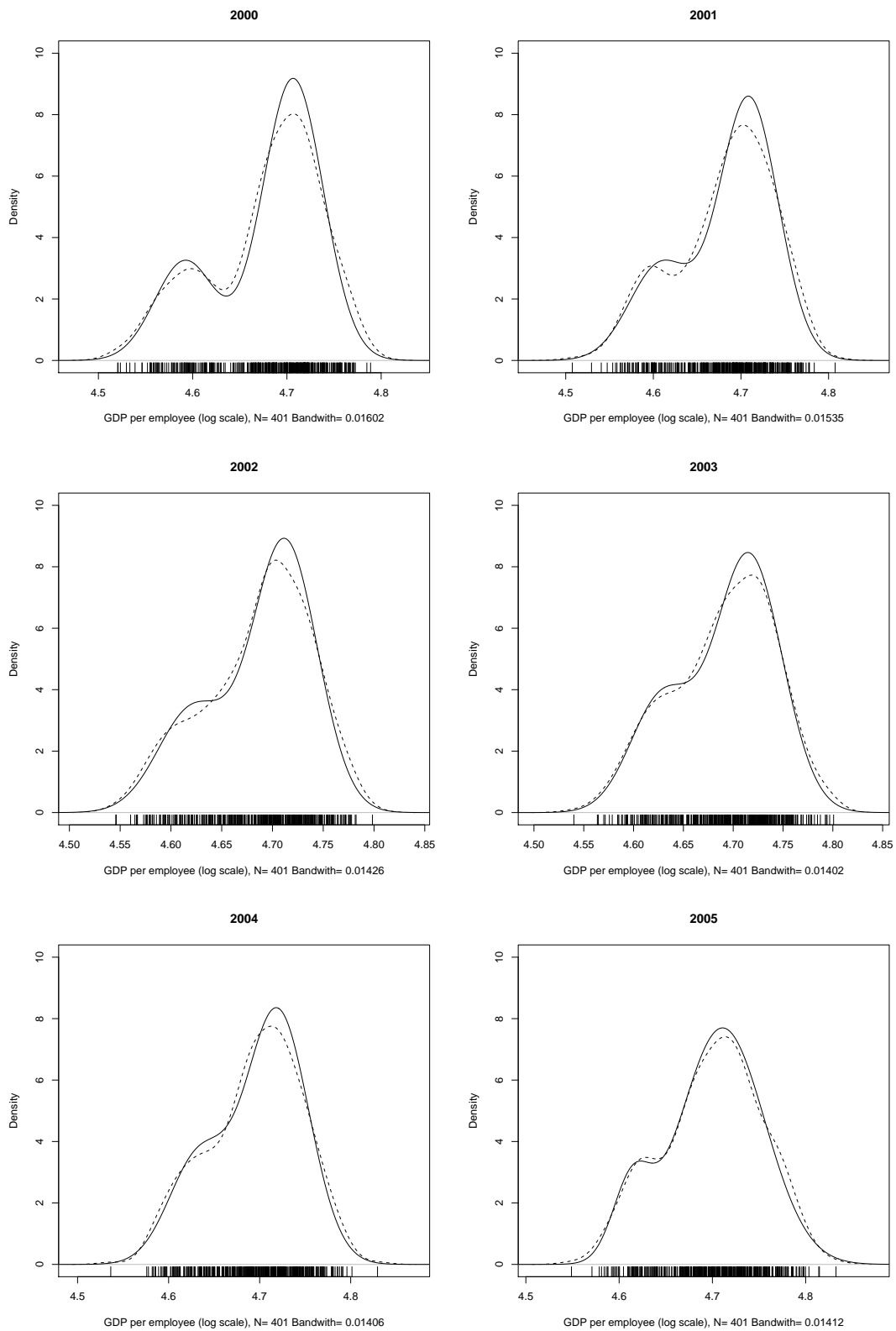


Figure 6: Distribution of GDP per employee between 2000 and 2005. Our model (solid line) compared to a kernel density estimator (dashed line).