Monopsony Makes Firms not only Small but also Unproductive: Why East Germany has not Converged

Rüdiger Bachmann, Christian Bayer,
Heiko Stüber, and Felix Wellschmied*

December 21, 2021, Current version

Abstract

When employers face a trade-off between growing large and paying low wages—that is, when they have monopsony power—some productive employers will decide to acquire fewer customers, forgo sales, and remain small. We show that this mechanism has adverse consequences for aggregate labor productivity. Using high-quality administrative data from Germany, we document that East German plants (compared to West German ones) face a steeper size-wage curve, are smaller, and invest less into marketing. This mechanism, in a model with labor market monopsony, product market power, and customer acquisition produces a 10 percent lower aggregate labor productivity in East Germany.

Keywords: aggregate productivity, plant heterogeneity, unions, monopsony, size-wage curve, monopolistic competition, customer capital, size distortions
JEL: E20, E23, E24, J20, J42, J50

*Rüdiger Bachmann: University of Notre Dame, rbachman@nd.edu; Christian Bayer: University Bonn, christian.bayer@uni-bonn.de; Heiko Stüber: Friedrich-Alexander Universität Erlangen-Nürnberg, heiko.stueber@fau.de ; Felix Wellschmied: Universidad Carlos III de Madrid, fwellsch@eco.uc3m.es. We thank seminar participants at the 28*th* CEPR ESSIM, the Oxford Conference on Firm Heterogeneity and the Macroeconomy, 2021 NBER-SI, the 2021 Macroeconomics Section Meeting of the German Economic Association, SITE 2021, the CRC TRR-224 workshop, and the IV Workshop Spanish Macroeconomics Network, as well as at the universities of Erlangen-Nürnberg, Bayreuth, Kiel, and FU Berlin. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FTP/2007-2013) / ERC Grant agreement no. 282740. Christian Bayer was visiting professor at the German ministry of finance while part of this research was conducted. The usual disclaimer applies. Felix Wellschmied gratefully acknowledges support from the Agencia Estatal de Investigación del Gobierno de España through research grant *ECO2017-86261-P* and the Comunidad de Madrid through research grant *EPUC3M11 (V PRICIT)* and thanks the Department of Economics at ITAM for its hospitality. Heiko Stüber gratefully acknowledge support from the German Research Foundation (DFG) under priority program "The German Labor Market in a Globalized World" (SPP 1764).

1 Introduction

Union membership around the world has declined. A large literature has investigated and discussed the direct labor market consequences of this trend. In this paper, we shift the focus to the macroeconomic and misallocation consequences when employers strategically adjust to heterogeneous union retrenchment. The East and West German labor markets provide a good laboratory for this. Both regions share the same legal and cultural institutions. In the East, however, collective bargaining and union membership are underrepresented in small plants.[1] As a result, the size-wage curve for plants is steeper in the East than in the West. This creates disincentives in East Germany to choose business models that require a plant to grow large such that the most productive plants hire relatively few workers and create relatively small customer networks. The aggregate productivity effects of these disincentives are sizable. Thirty years after the German reunification, labor productivity and wages remain about 25 percent lower in East relative to West Germany and the disincentives from a steeper size wage curve explain at least ten percentage points thereof.

We arrive at this conclusion by employing high quality administrative wage data, combining them with a new heterogeneous-firm model in which plants have product market power and face an increasing size-wage curve when they decide about entry, their customer networks, and their size. Our model thus tractably combines elements from two recent, but separate, strands of the heterogeneous-firm literature: First, by letting plants face a trade-off between paying low wages and growing large, we capture the effects of monopsony power on heterogeneously productive plants and relate to, e.g., Berger, Herkenhoff, and Mongey (2019).[2] Second, by modeling plants' decisions about the number of customers they want to acquire, we marry the monopsony literature with that on customer capital accumulation (see, e.g., Sedláček and Sterk, 2017; Arkolakis, 2010). We show that, in such a model, a steeper size-wage trade-off not only generates sizable negative aggregate productivity effects but also explains, parsimoniously, the differences in the plant-size distributions between East and West Germany.

[1] In communist economies, trade unions did not have the role to represent worker interests. As a consequence, after reunification, union membership fell dramatically (see Schnabel, 2005).

[2] We simplify the setup by focusing on monopsonistic, as opposed to oligopsonistic, competition and restrict the analysis to allocative effects, abstracting from normative efficiency questions.

We first show that, in the data, aggregate and sectoral differences in labor productivity are systematically related to the absence of large plants in East Germany. The share of employment at large plants, that is plants with more than 249 employees, is almost twice as large in the West. Turning to the sectoral data, manufacturing has both a particularly large labor productivity difference (30% higher in the West) and employment is much more concentrated (25 percentage point higher share at large plants again in the West), while construction, for example, has virtually the same employment concentration and a much smaller labor productivity gap (15%).

What is more, the lack of large plants is systematically related to differences in size-wage curves. Overall, the relationship between plant size and wages is one fifth steeper in East Germany relative to West Germany. Exploiting differences across industries, we show that those industries with steeper size-wage curves in the East are also those industries with particularly many missing large plants. In turn, the steeper size-wage curves in the East can be traced back to the fact that workers at small plants in the East are more likely to have individually and not collectively bargained wages compared to their Western counterparts.

To quantify the effects of a steeper size-wage trade-off on the plant size distribution and labor productivity, we employ a heterogeneous-plant model. We model long-run optimal plant decisions in a static framework which allows us to characterize the solution in closed form. Within period, plants have the following three-stage decision problem. First, plants decide about market entry. After market entry, they choose how many customers to accumulate, trading off additional sales and marketing expenses. This customer-base choice also takes into account the additional labor needed to supply these customers and, thus, that a larger customer base drives up wages in line with the upward-sloping size-wage curve present in the data. Finally, plants, taking into account their product market power, decide about prices charged to each individual customer and, thereby, about the number of workers required to service this customer.

The trade-offs plants face in these decision problems lead them to accumulate fewer customers the steeper the size-wage schedule they encounter. Indeed, we show that, in the data, marketing expenses are particularly small in those industries in East Germany which have a particularly steeper size-wage curve than their

West German counterparts. This has two effects on aggregate productivity: First, as plants on average have an incentive to remain small and search little for customers, the average customer bundles from fewer plants. This makes the economy less efficient, an adverse love-of-variety effect from monopsony power in the labor market. Second, compared to a situation with lower monopsony power, the employment distribution across plants is compressed, and labor is reallocated from more to less productive plants. Again, the result is an aggregate productivity loss. We also show that this second effect is exacerbated by product market power.

We calibrate our model to the average plant size and the share of large plants in West Germany. Imposing the steeper size-wage trade-off from East Germany explains a 10 percentage points lower productivity in that region. In addition, untargeted, the model replicates the plant-size distribution in East Germany. That is, it matches the smaller average plant size and the relatively small number of large plants. For the manufacturing sector, where East-West differences in plant size, the size-wage trade-off, and aggregate productivity are particularly pronounced, the calibrated model explains 18 percentage points lower productivity in East Germany. In a decomposition, we finally show that largest part of the productivity loss stems from the compression of the plant size distribution.

The remainder of the paper is organized as follows: after reviewing the literature, Section 2 discusses the data sets we use. Section 3 provides the empirical analysis. Section 4 introduces our model, and Section 5 discusses its quantitative implications. Section 6 concludes. We relegate additional material to a number of appendices. In particular, we show in Appendix A that East-West differences in aggregate labor productivity are driven by aggregate total factor productivity, not by quality of labor inputs nor by capital intensity or quality.[3] What is more, we show that the differences in aggregate total factor productivity are unlikely the result of a higher degree of labor market flexibility in West Germany, nor of differences in industry composition. Furthermore, we show in Appendix B that differences in the size distribution between East and West Germany and thus differences in aggregate labor productivity are not driven by the fact that East Germany has fewer metropolitan areas.

[3]Hence, even within a country, we confirm the well-known finding from Hall and Jones (1999) that differences in total factor productivity explain a large fraction of dispersion in labor productivity across geographical units.

Literature First, our paper is related to the literature that explains aggregate productivity losses as a result of too little employment at the most productive plants. For example, Hsieh and Klenow (2014) and Braguinsky, Branstetter, and Regateiro (2011) take the relatively slow growth of plants/firms as evidence of high (implicit) taxes on growing large and quantify the resulting productivity loss. More recently the literature, like this paper, starts from existing institutions like firing protections and links them to aggregate productivity losses caused by their effects on the plant size distribution. Examples are Garicano, Lelarge, and Van Reenen (2016) and Cingano, Leonardi, Messina, and Pica (2016). This paper highlights a new force behind productivity losses from a compressed plant size distribution: monopsony power in the labor market, see also Berger, Herkenhoff, and Mongey (2019). To study this force, the German case is particularly interesting. Government policies (and their enforcement) are essentially constant across regions but there are East-West German differences in labor market power related to the historically given lack of collective bargaining in small plants in East Germany. This selectively increases the steepness of the size-wage curve there.[4]

Second, our paper relates to the large literature on productivity (non-)convergence between countries in general (see Johnson and Papageorgiou, 2020, for a recent survey), as well as former socialist countries in particular (see Svejnar, 2002, for a survey). We study non-convergence within a country and thus non-convergence within the same legal framework.[5] This is different from the early difficulties of some other former socialist countries with building good legal institutions. Studying non-convergence within a country has the additional advantage that we can use high-quality micro data with common measures of factor inputs across the regions.

The particular case of non-convergence within Germany has drawn previous attention in the literature. Regarding convergence in labor productivity, Burda (2006) emphasizes the role of capital accumulation frictions for the slow conver-

[4]It has also been documented for the U.S. and the U.K. that collective bargaining makes the size-wage curve flatter (see Stewart, 1987; Brown and Medoff, 1989; Blanchflower, Oswald, and Garrett, 1990; Green, Machin, and Manning, 1996).

[5]Non-convergence can also be found in other countries (Italy's "Mezzogiorno", the US' "Rustbelt", etc.). What makes the German case of regional non-convergence particularly interesting is that there is a well-defined starting date from which onward we should expect convergence (October 3, 1990), a point made by Uhlig (2006).

gence between the two regions. While capital accumulation has played an important role for convergence right after the reunification, it cannot explain the persistent differences between the regions. Snower and Merkl (2006) study unemployment differences between East and West Germany and relate them to government transfers. Uhlig (2006) shows that initial conditions, i.e., at reunification, may be self-perpetuating when agglomeration effects in production networks are important. In our model, differences in production networks also play a role. They arise, however, endogenously from differently steep size-wage curves. Using cross-boarder worker mobility, Fuchs-Schündeln and Izem (2012) find that job, in contrast to worker, characteristics explain lower wages in East Germany. Using matched employer-employee data, Heise and Porzio (2021) document a low mobility of German workers across the two parts of the country. What is more, they also find that plant productivity differences (as opposed to worker quality differences) drive the majority of wage differences between the two regions. While their paper takes these plant productivity differences as given and explains why *worker mobility* does not remove East-West German wage differences, our paper explains why firm productivity is lower in East Germany and *firm mobility* does not remove these wage differences either. We thus view both papers as complementary.

Lastly, in terms of model ingredients, our paper borrows from two literatures. We start by drawing from the large literature on monopsony power in the labor market where firms internalize upward-sloping labor supply curves (Lamadon, Mogstad, and Setzler, 2018; Card, Cardoso, Heining, and Kline, 2018; Manning, 2011, 2003; Burdett and Mortensen, 1998), which recently Berger, Herkenhoff, and Mongey (2019) picked up in a heterogeneous-firm model of oligopsony to discuss distortions arising form size-dependent wage mark-downs. We, by contrast, highlight that monopsony power, in the form of monopsonistic competition, distorts also investment, for example, into customer acquisition. Customer acquisition, in addition to differences in technical productivities, is another force the literature (see Einav, Klenow, Levin, and Murciano-Goroff, 2021; Sedláček and Sterk, 2017; Gourio and Rudanko, 2014; Drozd and Nosal, 2012; Arkolakis, 2010) has highlighted to explain the size distribution of plants. Combined with a love-of-variety argument (see, e.g. Bilbiie, Ghironi, and Melitz, 2012), less customer acquisition leads to lower aggregate labor productivity in a monopsony framework.

2 Data

For our analysis, we use publicly available aggregate, sectoral and regional data and two administrative micro data sets. We focus on the private, non-primary sector (industries 10 to 82 in the German WZ2008 industry classification system). Specifically, we use German national income and product accounts data, *Volkswirtschaftliche Gesamtrechnung (VGR)*, to compute labor productivity at the national sectoral and regional level. The micro data sets are, respectively, the Structure of Earnings Survey (SES), *Verdienststrukturerhebung*, and the Administrative Wage and Labor Market Flow Panel (AWFP).

2.1 Structure of Earnings Survey (SES)

The *SES* is a cross sectional matched employer-employee data set maintained by the German statistical agency (*Statistisches Bundesamt*). The SES is carried out every four years. The German statistical agency randomly samples plants and, by law, these plants are required to provide detailed information on their employees and their employees' monthly working hours and earnings. Hence, selection due to nonresponse does not arise. It contains the number of employees at a plant as well as industry classification and location information at the superregional level. In particular, the SES divides Germany up into 5 regions.[6] The sample is representative for the universe of all German plants with at least ten employees.[7]

For our analysis, we employ the 2006, 2010, and 2014 samples, which we pool for most empirical analysis. We drop all civil servants from our sample as well as all plants where at least 50% of employees are public servants. Moreover, we restrict the sample to full-time employees. The final sample contains 2,364,862 worker-plant observations. The 2006 sample uses a different industry classification than the later two samples. As a result, we have to merge some industries to have

[6] North: Schleswig Holstein, Hamburg, Bremen, Berlin, and Lower Saxony; West: Northrhine-Westphalia; South-West: Hesse, Rhineland Palatinate, and Saarland; South: Baden-Württemberg and Bavaria; East: Thuringia, Saxony, Saxony-Anhalt, Mecklenburg Western Pomerania, and Brandenburg. West Germany summarizes the North, West, South-West, and South.

[7] The restriction on ten or more employees is meant to reduce administrative burden on small enterprises.

a consistent classification. Table C1 in the Appendix C provides a crosswalk for this merger and shows how it relates to the sectors from the national accounts.

The SES provides the best-available data source for our analysis. First, the SES is a representative sample for non-farm German plants with at least 10 employees. We focus on the private non-primary sector. Self-employed workers are not covered. Second, data on regular earnings, overtime pay, bonuses, and hours paid, both regular and overtime, are extracted from the payroll accounting and personnel master data of plants and transmitted via software interface to the statistical office. Transmission error is, hence, negligible. That is, unlike German social security data, the SES reports the actual (virtually uncensored) pay and hours worked of employees. Third, it also provides detailed information on workers' education, occupation, age, tenure, and job levels. Fourth, the survey has information on about 3.2 million employees from roughly 28,700 establishments in 2006, 1.9 million employees from 32,200 establishments in 2010, and 0.9 million employees from 35,800 plants in 2014. The number of sampled employees decreased over time because the sampling probability of plants became smaller to reduce bureaucratic costs. In our analysis, we equalize observation weights across surveys so that all surveys receive equal weight.

2.2 Administrative Wage and Labor Market Flow Panel (AWFP)

For some analyses, principally for longer time series, we supplement the SES with the *AWFP* which is a quarterly plant-level data set based on German social security data which contains daily earnings, not wages, up to the social security cap. The data covers the universe of private German plants and is available for both West and East Germany from 1993 until 2014 (see Stüber and Seth, 2017; Bachmann, Bayer, Merkl, Seth, Stüber, and Wellschmied, 2021). The AWFP's data source is the Employment History (*Beschäftigten Historik*, BeH) of the German Institute for Employment Research (IAB). The BeH is an individual-level data set covering all workers in Germany subject to social security.[8] The information in

[8]Marginal part-time workers (*geringfügig Beschäftigte*) have been covered since 1999. The main types of employees not covered by the BeH are civil servants (*Beamte*), military personnel,

the BeH originates from the notification procedure for social security. Essentially, this procedure requires employers to keep the social security agencies informed about their employees by reporting any start and end date of employment and by annually confirming existing employment relationships. The *AWFP* aggregates this individual worker data to the plant level. We use the *AWFP* on occasion because it covers a longer time period than the *SES* and provides supplementary information about the plants, but its wage data are inferior to the *SES*.[9]

3 Size Distortions

We start this section by documenting that, at an aggregate level, East Germany has a lower aggregate labor productivity and compensation per worker, whether one includes the public and primary sector or not. The *SES* data allows us to establish that the lower labor productivity in East Germany is related to missing large plants in the East which itself is related to a steeper size-wage relationship there. We start by showing that, at the aggregate level, East Germany has fewer large plants than West Germany. Next, we show that productivity differences between East and West are particularly pronounced in sectors with more missing large plants in the East. We then document a novel size distortion: a relatively steeper size-wage relationship in East Germany. Using more detailed industry data, we show that a steeper size-wage relationship in an industry in the East predicts more missing large plants in that industry. This steeper size-wage relationship itself relates to the fact that workers at small plants in the East are less likely to be paid according to a collective bargaining agreement.

---

and the self-employed. East German employees were integrated with the West-German social security administration only after 1992.

[9]To ensure consistency over time, most variables in the AWFP—and all variables used in this paper—are calculated on a 'regular worker' basis. In the AWFP, a person is defined as a 'regular worker' when she is employed full-time and belongs to one of the following person groups: 'employees subject to social security without special features', 'seamen' or 'maritime pilots.' Therefore (marginal) part-time employees, employees in partial retirement, interns, etc., are not counted as regular workers.

3.1 Aggregate Productivity

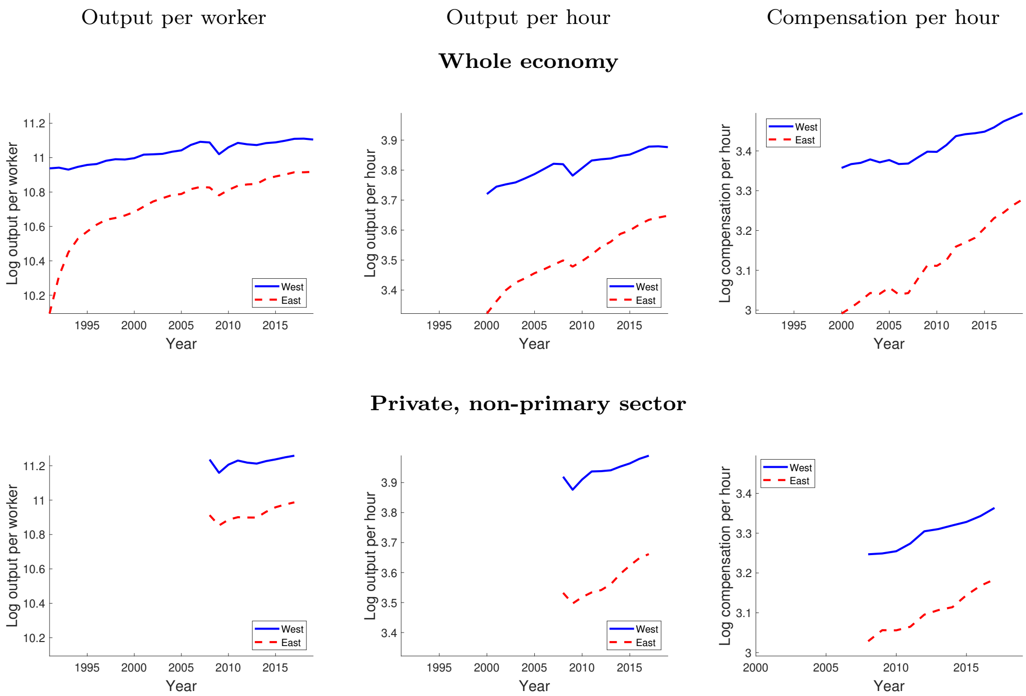
In 1991, when centrally planned East Germany reunited with West Germany and became a market economy, other factors depressing labor productivity played an important role. Capital was in short supply, machines were outdated, and political pressure had plants over-employ labor in the East. Consequently, labor productivity did not even reach 50% of the West German level in 1991 (see the first panel in Figure 1). During the first couple of years after reunification, labor productivity and wages grew quickly. However, this process ended soon, in about 1995. Since then, convergence in relative labor productivity and wages has almost come to a halt.[10] What is more, as the bottom panel of Figure 1 shows, the East-West productivity difference is with 25% even larger in the private (non-primary) sector. Finally, the rightmost panels show the same East-West differences in real wages. This makes an explanation unlikely that attributes the productivity differences to a pure statistical artifact from transfer pricing within firms. Since headquarters of most large firms are located in West Germany, income from unlocalized intangibile capital might be accounted for as headquarter income and thus in West Germany. If there were otherwise no underlying localized productivity differences, wages across the two regions should be the same.

3.2 Missing Large Plants in East Germany

That East Germany has fewer large plants than West Germany both in the private sector overall and in the large manufacturing sector in particular can be seen from Figure 2. The top panels show this in terms of the (employment-weighted) density of plants over log employment for the pooled samples. The bottom panels show this in terms of the CDF of employment over (log) plant sizes for each survey year. In all these years, employment at large plants is much more prevalent in the West. In this paper, we follow the definition of the German statistical agency and refer to large plants as those with more than 249 employees. In the West, close to 40% of employees are employed at such large plants in 2014, as the rightmost lower panel shows. The same number for East Germany is only around 25%. In Appendix B

[10] We use output per worker as our baseline measure of labor productivity. As the figure shows, differences in output per hour are even somewhat larger than those in output per worker.

Figure 1: Output and wages

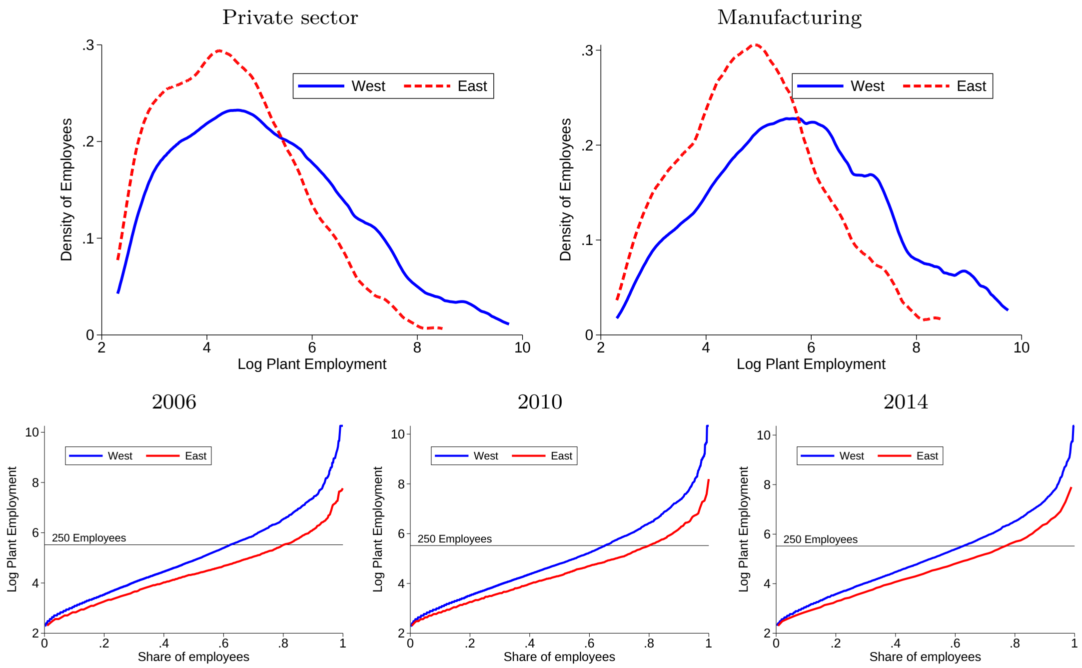


Notes: The figure displays yearly log output per worker, log output per hour, and log labor compensation per hour in East and West Germany. The top panel displays it for the whole economy, the bottom panel for the private, non-primary sector. Calculations are based on national accounts (VGR). The data is available by region and sector only since 2008, which is why the lower panel starts only in that year. Similarly, data on hours worked by region starts in 2000. Weinand and von Auer (2020) provide county-level consumer price indices for Germany in 2016 that we aggregate to the regional level using population weights. With 2016 as the base year, we then calculate a time series of regional prices using the regional GDP-deflator-based inflation rates from national accounts.

we show that the main result extends to earlier time periods and is not driven by differences in urbanization between East and West Germany.

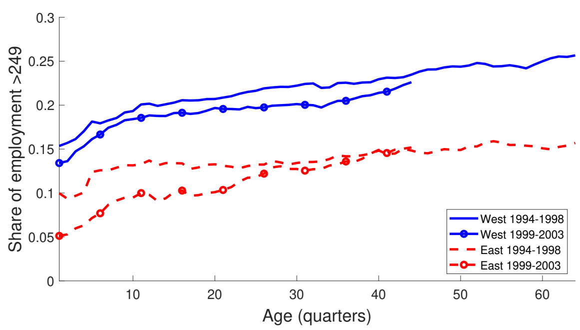
A potentially confounding factor for the East-West difference in the plant size distribution could be plant age. The restructuring of the East German economy led to the exit of many old, large plants. Figure 3 shows, however, that even conditional on plant age, East German plants are smaller, because they enter smaller and they remain smaller. Put differently, already at entry, plants in East Germany choose production technologies that imply a relatively small plant size.

Figure 2: Plant-size distributions in East and West Germany, 2014



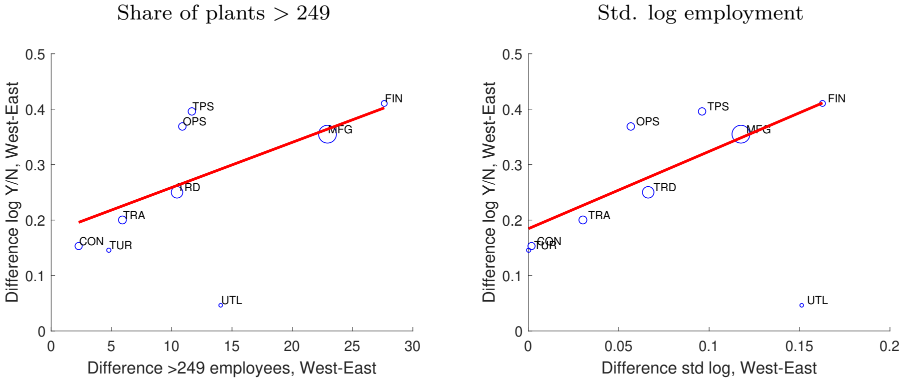
Notes: The figure displays the employment weighted plant size distribution in East and West Germany. The top panels display, respectively, an estimated density function (by a Gaussian kernel smoother) in the total private, non-primary sector and in the manufacturing sector. The bottom panels display, for different survey years, the empirical CDF of employment over plant log-employment for the total private non-primary sector. Data source: *SES*.

Figure 3: Employment share 250+ by cohort



Notes: The figure displays for different plant-entry cohorts the share of employment at plants with more than 249 employees over their life-cycles. Calculations are based on the *AWFP* data from the private, non-primary sector.

Figure 4: Productivity differences and large plants by sector



Notes: The figures relate 2014 log differences in output per worker between West and East Germany within major sectors to the share of employment at plants with more than 249 employees and the standard deviation of log plant employment. The lines show weighted-least squares regressions. *MFG*: Manufacturing, *UTL*: Utilities, *CON*: Construction, *TRD*: Wholesale and Retail, *TRA*: Transportation, *TUR*: Tourism, *FIN*: Finance, *TPS*: Technical professional services, *OPS*: Other professional services, see Appendix C. Data sources: *SES* and *VGR*.

What is more, the East-West difference in the employment share of large plants is essentially constant both in plant age and across entry cohorts.

Returning to Figure 2 and comparing its top two panels, one can also see that the East-West differences in the plant size distribution are not uniform across sectors. They are much stronger in the manufacturing sector, where in the West, 55% of all employees work at plants with more than 249 employees, while in the East it is only 31%. Figure 4 makes this comparison systematically across all sectors and relates it to sectoral productivity differences. The left panel uses the share of employment at plants with more than 249 employees to compare the plant size distribution. The right panel uses the standard deviation of log-employment, $\sigma^W_{\log e} - \sigma^E_{\log e}$, instead. The employment-weighted correlation between the two measures is 0.72 for the left panel and 0.70 for the right panel. Both scatter plots show that those sectors where productivity is particularly low in the East are also the sectors where particularly fewer workers are employed at large plants in East Germany relative to West Germany.

3.3 Size-Wage Relationships and Missing Large Plants

These differences in the plant-size distribution are in turn related to differences in the size-wage relationship that plants face. To show this, we use the *SES* data to estimate the following reduced-form relationship between individuals' log wages, $\ln w_{it}$, and the log employment at their plant, $\ln E_{it}$:

$$\ln w_{it} = \beta_0 + \beta_E \, East_i + \hat{\omega}_W \ln E_{it} + (\hat{\omega}_E - \hat{\omega}_W) \, East_i \ln E_{it} + \beta x_{it} + e_{it}, \quad (1)$$

where $East_i$ is a dummy that is one when the employer is located in East Germany and x_{it} are other observable plant or worker characteristics. The coefficient of interest is the difference in the size-wage slope $\hat{\omega}_E - \hat{\omega}_W$, the interaction term. In our baseline specification, we non-parametrically control for a workers' age and sex by a full set of interaction dummies and for time and industry fixed effects. For robustness, we consider a second (and a third) specification where we fully interact age, sex, education, and occupation (job-level) dummies (in addition to time and industry fixed effects) to allow for differences in occupational (job-level) patterns within industries between the two regions.

The top panel of Table 1 displays the results. It first shows that large plants pay higher average wages in both regions as $\hat{\omega}_{W,E} > 0$. Importantly, the size premium is larger in East Germany. In the West, increasing employment by 1% increases wages by 0.078%. The corresponding number for the East is 0.094%, one fifth higher. For example, if a plant wants to increase its employment from 50 to 100 (log difference 0.69) it has to pay 5.6% higher wages in the West, while in the East, the same size increase comes with a wage increase of 6.7%. The last two columns of Table 1 show that the difference between the two regions becomes yet slightly larger when we control additionally for age-, sex-, and education-specific occupational or job-level patterns.[11] For manufacturing, the difference in the size-

[11]This suggests that selection of higher paid workers (proxied for by education and occupations or job levels) into larger plants and regional differences therein are not driving our results in the sense that large plants in East Germany would attract a larger share of high ability workers. If anything, the effect of including more proxies for a worker's skills shows that this selection is stronger in West Germany. In Appendix D.2 we investigate the issue of selection further by using the social security data which allow us, with the caveat of that these are top-coded earnings as opposed to hourly wage data, to use estimates of plant-level fixed effects controlling for worker fixed effects. We find the same pattern of a steeper East German size wage curve. Furthermore,

Table 1: Size-wage relationships

	Non-primary private sector		
	Baseline	Occupation x Education	Job level x Education
Wage-size elasticity, West, $\hat{\omega}_W$	7.8 (0.1)	6.1 (0.1)	5.5 (0.1)
Difference in elasticities, $\hat{\omega}_E - \hat{\omega}_W$	1.6 (0.3)	2.0 (0.2)	2.5 (0.2)
Implied elasticity, East, $\hat{\omega}_E$	9.4	8.4	8.4
N (in thousands)	2365	2365	2228
	Manufacturing sector		
	Baseline	Occupation x Education	Job level x Education
Wage-size elasticity, West, $\hat{\omega}_W$	8.8 (0.2)	6.9 (0.1)	6.5 (0.1)
Difference in elasticities, $\hat{\omega}_E - \hat{\omega}_W$	4.3 (0.4)	4.9 (0.3)	5.4 (0.3)
Implied elasticity, East, $\hat{\omega}_E$	13.1	12.3	12.3
N (in thousands)	1025	1025	970
	Type of bargaining		
	Non-collective	Collective	
Wage-size elasticity, West, $\hat{\omega}_W$	7.7 (0.2)	5.8 (0.2)	
Difference in elasticities, $\hat{\omega}_E - \hat{\omega}_W$	-0.3 (0.4)	-0.3 (0.4)	
Implied elasticity, East, $\hat{\omega}_E$	7.4	5.5	
N (in thousands)	1378	986	

Notes: The table displays the estimated size-wage relationships for the non-primary private (manufacturing) sector in West and East Germany. Standard errors are in parentheses. The top panel is for all workers. The bottom panel splits the sample (non-primary private sector) by whether the worker is covered by a collective bargaining agreement or not. All coefficients are multiplied by 100 for better readability. *Baseline*: Controls for a workers' age and sex by a full set of dummy-interactions, plus time, and industry fixed effects. *Occupation x Education*: Controls for a workers' age, sex, education, and occupation by a full set of dummy-interactions, plus time and industry fixed effects. *Job level x Education*: Controls for a workers' age, sex, education, and job level (5 levels of job level, coding the level of autonomy, complexity, and responsibility a worker's job has, see Bayer and Kuhn, 2018) by a full set of dummy-interactions, plus time and industry fixed effects. Data source: *SES*.

Appendix D.1 considers two other robustness checks which both slightly increase the difference between West and East Germany. First, we include part-time workers into our sample. Second, we allow the size-wage relationship to be non-linear by including a common quadratic size effect into the regression.

wage relationship between East and West is even more pronounced.

In fact, that the East-West difference in the size-wage relationship is not uniform across industries generalizes. Importantly, it is also systematically related to industry variation in the prevalence of large plants, as Figure 5 shows. Industries with many missing large plants in the East have also a steeper size-wage relationship in the East.

Concretely, we estimate Equation (1) for 21 individual industries. We plot the the difference $\hat{\omega}_E - \hat{\omega}_W$ against (a) the difference in the share of employment at large plants and (b) the difference in the standard deviation of log employment for each industry. Here we can go beyond sectoral disaggregation as we do not need to rely on regional *VGR* data (as we needed to calculate productivity in Figure 4).[12] We find a positive relationship between steeper size-wage relationships in the East and a larger difference in the share of employment at large plants (industry-size weighted correlation of 0.3). The industry-size weighted correlation for the standard deviation of log plant employment is 0.34. In Appendix E we repeat everything in Figure 5 (as well as Figure 4) splitting up West German industries by four regions. The resulting correlations are similar but come with a higher degree of statistical confidence.

What lies behind these differences in the steepness of the size-wage relationship? It could be that East Germans have more specific workplace preferences, leading to lower degrees of substitutability between employers. By contrast, we highlight the role of collective wage bargaining and the differences in the role of unions rooted in the different historical developments before 1990. We find that, once we condition on whether individual employment contracts are subject to collective bargaining, the size-wage relationship in West and East Germany is basically identical (see the bottom panel of Table 1). Since collectively bargained wages are in general higher,[13] the fact that the size-wage relationship is flatter for collectively bargained wages (e.g. 5.8 vs. 7.7 in the West) means that collective

[12] Table C1 in Appendix C shows the mapping between the two classifications.

[13] For all plant sizes collective bargained wages are higher in our sample. Incidentally, this explains why the economy-wide size-wage curve is steeper than those conditional on the bargaining arrangement. Wages at larger plants are more often subject to collective bargaining, so that the higher collectively bargained wages play an ever larger role in the composite regression when moving from smaller to larger plants.

bargaining in particular raises wages at small plants.[14] One way to interpret this flatter size-wage relationship is that unions reduce a plant's monopsony power. Putting together the lack of a difference in the size-wage relationship conditional on collective bargaining and the overall higher collectively bargained wages means that the overall steeper size-wage relationship in the East is driven by composition differences between small and large plants regarding the prevalence of collective bargaining. The aforementioned preference-based explanation appears to be difficult to reconcile with this pattern in Germany.

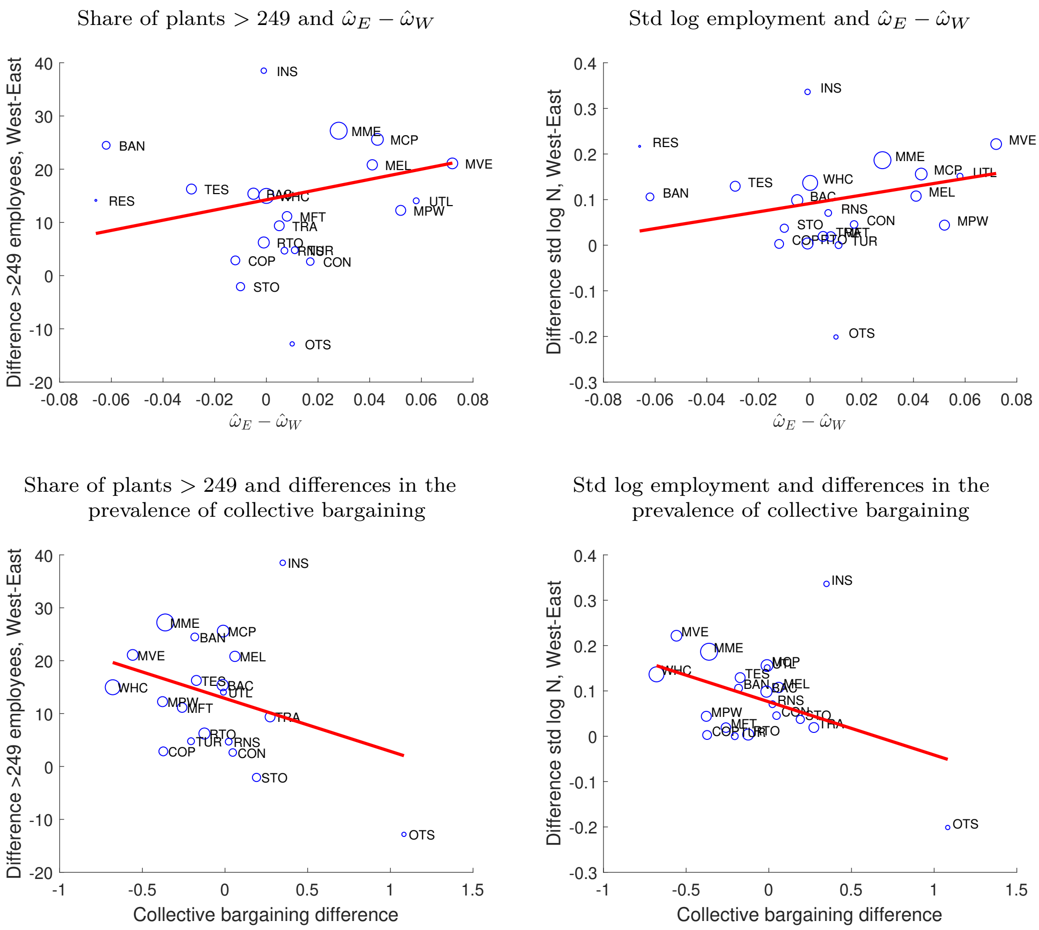
In fact, workers at small plants in the East have a particularly low probability to be covered by a collective bargaining agreement relative to their West German counterparts, as we will show next.[15] The bottom panels of Figure 5 show on the x-axes, for each industry, a double difference in the prevalence of collectively bargained wage contracts between large and small plants and between East and West. For the majority of industries this double difference is negative. This means that the fraction of collectively bargained wage contracts increases indeed more in plant size in the East than it does in the West. This double difference is plotted against our two measures of East-West differences in the plant size distribution: the share of employment at large plants (left panel) and the standard deviation of log plant-level employment (right panel). The relationship between collective-bargaining prevalence differences and plant-size differences is negative with an employment-weighted correlation of -.3 and -.4, respectively. Industries in which the prevalence of collectively bargained wages increases relatively more in plant size in the East are also those industries where, compared to the West, large plants are particularly missing in the East.

In summary, the data suggests that plants in East Germany face a stronger trade-off between growing large and paying low wages. This stronger trade-off appears to originate from the relative concentration of collective bargaining in East Germany at large plants. Most importantly, across industries/sectors the

[14] The size-wage relationship for collectively bargained wages is not completely flat for at least two reasons. First, firms can negotiate firm-specific agreements that are collective in the sense that they hold for their entire workforce. Second, the typical industry-wide collective bargaining agreement in Germany establishes a wage floor for all plants bound by the agreement, but allows to pay an individual worker better, e.g., through bonuses.

[15] See also Table 2 in Schnabel (2005).

Figure 5: The share of large plants, the size-wage relationship, and collective bargaining



Notes: The top panel relates differences between West and East Germany in the share of employment at large plants and the standard deviation of log plant employment to differences in size-wage relationships. The bottom panel relates differences between West and East Germany in the share of employment at large plants and the standard deviation of log plant employment to the following double difference: $\log P(C|L,W) - \log P(C|S,W) - [\log P(C|L,E) - \log P(C|S,E)]$, where $P(C|\cdot)$ is the conditional probability of a worker being subject to collective bargaining in our sample in (L)arge (>249 employees) or (S)mall ($\leq$ 249 employees) plants in the (E)ast and (W)est. The lines show weighted-least square regressions. MFT: Food and textile manufacturing, MPW: Paper and wood manufacturing, MCP: Chemical and plastic manufacturing, MME: Metal manufacturing, MEL: Electronics manufacturing, MVE: Vehicle manufacturing, UTL: Utilities, CON: Construction, COP: Construction preparations, WHC: Wholesale and car retail, RTO: Other retail, TRA: Transportation, STO: Storage, TUR: Tourism, BAN: Banking, INS: Insurance, RNS: Research services, TES: Technical services, RES: Rental services, BAC: Building and area care, OTS: Other services, see Appendix C. Data sources: Calculations using the *SES*.

stronger size-wage trade-off in the East correlates with missing large plants and lower productivity.

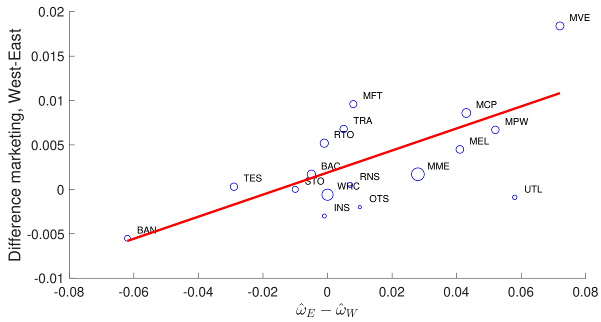
4 A Model of Missing Large Plants

To understand why a stronger size-wage trade-off leads to missing large plants and lower productivity in East Germany, we introduce labor market power into a heterogeneous plant-size model with endogenous demand (customer) accumulation. This accords with the literature which has emphasized two forces to explain heterogeneous plant sizes: productivity and demand differences. For the former, we follow much of the literature that treats productivity differences as exogenous. For the latter, we draw on a recent and growing literature that puts a form of customer accumulation at the center stage in addition to productivity differences (see Arkolakis, 2010; Drozd and Nosal, 2012; Gourio and Rudanko, 2014; Sedláček and Sterk, 2017). In these models, in order to grow, plants have to make potential customers aware of their products through marketing. We highlight that this decision naturally interacts with labor-market power: firms that face a steeper size-wage trade-off will tend to accumulate fewer customers. Indeed, we see this prediction to be borne out by the data: Figure 6 shows that East-West differences in marketing expenditures are systematically related to the size-wage trade-off at the industry level. Industries with particularly steep size-wage trade-offs in the East spend, relative to West Germany, little on marketing (a size-weighted correlation of 0.71).[16]

Concretely, we introduce a size-wage trade-off into the following framework: There are intermediate good producers with heterogeneous productivities who can use labor to produce a differentiated good. First, these potential producers decide on market entry; second, conditional on entry, they learn their productivity and decide on marketing expenditures that are used to form production networks with final goods producers (bundlers). Third, intermediate good producers hire labor and produce, facing both a size-wage and an output-price trade-off. Fourth,

[16]Data for the ratio of marketing expenditures relative to sales at the industry level comes from the *Mannheimer Innovationspanel*. We are extremely grateful to the team at the ZEW, in particular Christian Rammer, who shared this data with us.

Figure 6: Marketing expenditures and $\hat{\omega}_E - \hat{\omega}_W$



Notes: The figure relates differences between West and East Germany in the ratio of marketing expenditures relative to sales (West minus East in 2014) to differences in the size-wage relationship for those industries where marketing expenses are available. The line shows a weighted-least square regression. MFT: Food and textile manufacturing, MPW: Paper and wood manufacturing, MCP: Chemical and plastic manufacturing, MME: Metal manufacturing, MEL: Electronics manufacturing, MVE: Vehicle manufacturing, UTL: Utilities, CON: Construction, COP: Construction preparations, WHC: Wholesale and car retail, RTO: Other retail, TRA: Transportation, STO: Storage, TUR: Tourism, BAN: Banking, INS: Insurance, RNS: Research services, TES: Technical services, RES: Rental services, BAC: Building and area care, OTS: Other services, see Appendix C. Data sources: Calculations using the *SES* and the *Mannheimer Innovationspanel*.

bundlers produce a perfectly substitutable consumption good under perfect competition.[17] Finally, given that East-West differences in plant size are relatively stable in plant age and across cohorts (see Figure 3) we abstract from plant dynamics to maintain tractability.

4.1 Bundlers

There is a unit mass of bundlers who are indexed by j. Each produces a final consumption good, Y_j, using a Dixit-Stiglitz aggregator.

$$Y_j = \left(\int \gamma_i \theta_{ij} y_{ij}^{\frac{\eta-1}{\eta}} \, di \right)^{\frac{\eta}{\eta-1}}. \tag{2}$$

[17]We emphasize the interaction of customer accumulation and labor market power in shaping plant size and productivity, and, therefore, we abstract, for tractability reasons, from how interregional trade additionally influences this nexus. We thus model each East and West Germany as closed economies which is tantamount to assuming that the bundlers in both regions produce perfect substitutes. In addition, since plants in both regions, because of free entry, make zero expected profits in equilibrium, there is no incentive for plants to start up in another region.

They bundle differentiated goods, y_{ij}, of a continuum of potential intermediate good producers i (again of mass one).

A potential intermediate good producer may enter and be active, $\gamma_i = 1$, or not, $\gamma_i = 0$. Not all active intermediate good producers are known to each bundler, and producer i is known to bundler j only if $\theta_{ij} = 1$. A bundler can only buy an intermediate good from a producer that is both active and known to the bundler. This implies that the demand for producer i's product by bundler j is given by

$$y_{ij} = \gamma_i^\eta \theta_{ij}^\eta \left(\frac{p_{ij}}{\bar{P}_j} \right)^{-\eta} Y_j, \tag{3}$$

where $\bar{P}_j$ is the cost minimizing price at which bundler j sells its bundle, and p_{ij} is the price of the intermediate good charged by producer i to bundler j.

The cost minimizing price of bundler j is given by

$$\bar{P}_j = \left(\int (\gamma_i \theta_{ij})^\eta p_{ij}^{1-\eta} di \right)^{1/(1-\eta)}, \tag{4}$$

which can be written as

$$\bar{P}_j = \left(\int (\gamma_i \theta_{ij})^\eta di \right)^{1/(1-\eta)} \left(\int p_{ij}^{1-\eta} di \right)^{1/(1-\eta)} \tag{5}$$

because we assume that prices and γ and θ are independent. The latter reflects random matching, the former is tantamount to assuming, without loss of generality, that inactive producers set a price as if they were active and could sell (a weakly dominant strategy). What is more, random matching implies that the integral

$$\left(\int (\gamma_i \theta_{ij})^\eta di \right)^{1/(1-\eta)} \tag{6}$$

does not depend on the specific bundler j, and in turn all bundlers charge the same price:

$$\bar{P}_j = (\Gamma \bar{\Theta})^{1/(1-\eta)} \hat{P}_j, \tag{7}$$

where Γ is the mass of all active producers, $\bar{\Theta}$ is the average fraction of active producers known to a bundler, which by symmetry is also the average fraction of bundlers that an active producer sells to (and therefore has no j index), and

$$\hat{P}_j = \left(\int p_{ij}^{1-\eta} di \right)^{1/(1-\eta)} \tag{8}$$

is the average price charged by intermediate good producers. Because all bundlers j charge the same price, we focus on the symmetric equilibrium in which $Y_j = Y$, $\hat{P}_j = \hat{P}$, and $\bar{P}_j = \bar{P}$.

4.2 Intermediate Good Producers

Intermediate good producers operate a constant returns to scale production function that transforms l_i unit of labor into $y_i = z_i l_i$ units of the intermediate good, where z_i denotes idiosyncratic productivity. Because in the symmetric equilibrium $Y_j = Y$ and $\bar{P}_j = \bar{P}$ the intermediate goods producer supplies the same amount of goods to each bundler she knows, we can drop the subscript j and let y_i denote the representative quantity that an active producer supplies to each bundler she knows and l_i the number of workers that are needed to produce this representative quantity. The total number of employees of an intermediate good producer is $l_i \Theta_i$, where Θ_i is the number of bundlers known to that producer.

Moreover, an intermediate good producer faces monopsonistic competition in the labor market, i.e., the wage is a function of its total number of employees. As in our empirical specification, Equation (1), we assume a constant elasticity:

$$w_i = \left(\frac{l_i \Theta_i}{\bar{l} \bar{\Theta}} \right)^{\hat{\omega}} W, \tag{9}$$

where we express size relative to the average producer size in the economy, $\bar{l}\bar{\Theta}$, and W is a wage index, which we set to 1, making labor the numeraire. While a wage curve like (9) could be derived from preferences for specific workplaces (see e.g. Berger, Herkenhoff, and Mongey, 2019), we do not need to take a stance on its precise micro-foundation because our research question is not of a normative nature. Nonetheless, our results from the previous section that East-West differences in the size-wage curve vanish after controlling for bargaining arrangements suggest an institutional rather than a preference-based micro-foundation for these differences. In turn, these institutional differences are arguably history driven so that we can view them as exogenous to our question.

Given this environment, we solve the decision problem of the intermediate good producers backward, starting with the optimal price-setting to one bundler. Thereafter, we solve for the optimal marketing policy given the downstream price-setting decisions.

4.2.1 Price-Setting and Profits within a Single Market

Since intermediate good producers in each single (bundler/product) market face monopolistic competition for any bundler they are known to, they set prices as a mark-up over marginal costs, given by wages w_i relative to productivity z_i:[18]

$$p_i = \frac{\eta}{\eta - 1} \frac{w_i}{z_i}. \tag{10}$$

Hence, a producer who knows Θ_i bundlers has a total gross profit of:

$$\pi_i(\Theta_i) = \Theta_i \left(p_i y_i - y_i \frac{w_i}{z_i} \right) = \Theta_i \left(y_i \frac{1}{\eta - 1} \frac{w_i}{z_i} \right), \tag{11}$$

where the terms in brackets are the gross profits earned from commerce with an individual bundler.

Substituting into the gross profits the demand curve from an active market,

$$y_i = \left(\frac{p_i}{\bar{P}} \right)^{-\eta} Y, \tag{12}$$

as well as the optimal price, Equation (10), allows us to express gross profits as a function of known bundlers, the wage, and idiosyncratic productivity:

$$\pi(\Theta_i) \quad = \quad \Theta_i \left(\frac{w_i}{z_i} \right)^{1-\eta} \bar{P}^{\eta} Y \frac{(\eta - 1)^{\eta - 1}}{\eta^{\eta}}. \tag{13}$$

[18]The intermediate good producers' price-setting can ignore the fact that they are in monopsonistic competition in the labor market, as each bundler is infinitesimally small and, hence, a marginal increase in the quantity sold to a single bundler has only a second-order impact on the producer's total labor demand and is thus irrelevant for the producer's first-order condition.

4.2.2 Optimal Marketing

The intermediate good producer maximizes gross profits net of marketing cost but takes into account wages as a function of the total number of employees. Therefore, we first need to express wages in (13) as a function of the number of bundlers the are known to the producer. To this end, we plug the number of workers, $\frac{y_i}{z_i}$, required to fulfill a producer's demand, (12), into the size-wage trade-off, (9):

$$w_i = \left(\frac{\left(\frac{p_i}{\bar{P}}\right)^{-\eta} Y \Theta_i}{z_i \bar{l} \bar{\Theta}} \right)^{\hat{\omega}}. \tag{14}$$

Next, substituting p_i with the optimal pricing decision (10), solving for the wage w_i, and redefining terms, we obtain wages as a function of the number of known bundlers as well as productivity and aggregates:

$$w_i = z_i^{\frac{(\eta-1)\hat{\omega}}{1+\eta\hat{\omega}}} \bar{w} \left(\frac{\Theta_i}{\bar{\Theta}} \right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}}, \tag{15}$$

where $\bar{w} = \left(\bar{P}^{\eta} Y \left(\frac{\eta}{\eta-1} \right)^{-\eta} / \bar{l} \right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}}$ summarizes the aggregate terms that affect wages.

Given this reformulation of the size-wage trade-off, we are now ready to solve for the optimal marketing policy. To get to know one additional bundler, the intermediate good producer has to pay marketing expenditures, $\mu\bar{P}$ (μ measures costs in terms of the output good). The resulting operating profits are

$$\Pi_i = \pi(\Theta_i) - \mu \bar{P} \Theta_i. \tag{16}$$

Substituting in gross profits, (13), and the wage-size trade-off, (15), yields

$$\Pi_i = \Theta_i \left(\frac{\Theta_i}{\bar{\Theta}} \right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}(1-\eta)} z_i^{-(1-\eta)\frac{\hat{\omega}+1}{1+\eta\hat{\omega}}} \bar{P}^{\eta} Y \frac{(\eta-1)^{\eta-1}}{\eta^{\eta}} \bar{w}^{1-\eta} - \mu \bar{P} \Theta_i. \tag{17}$$

The optimal scope of producer i follows from the first order condition, $\frac{\partial \Pi_i}{\partial \Theta_i} = 0$, ignoring, for simplicity, that $\Theta_i \leq 1$:

$$\bar{P}^{\eta-1}\frac{Y}{\mu}\frac{(\eta-1)^{\eta-1}}{\eta^{\eta}}\bar{w}^{1-\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\left(\frac{\Theta_i}{\bar{\Theta}}\right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}(1-\eta)} = z_i^{(1-\eta)\frac{\hat{\omega}+1}{1+\eta\hat{\omega}}}, \tag{18}$$

which, solving for Θ_i, simplifies to

$$\frac{\Theta_i}{\bar{\Theta}} = z_i^{\frac{1+\hat{\omega}}{\hat{\omega}}}\left[\frac{Y}{\mu}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\frac{1}{\eta}\left(\frac{\bar{P}}{\bar{w}}/\frac{\eta}{\eta-1}\right)^{\eta-1}\right]^{\frac{1+\eta\hat{\omega}}{\hat{\omega}(\eta-1)}}. \tag{19}$$

This equation relates the optimal amount of known bundlers to a producer's idiosyncratic productivity, z_i. More productive producers find it optimal to accumulate more customers. A yet different way to think about the producers' optimal marketing decision is to use (15) and express (19) in terms of the real wage targeted by a producer:

$$\frac{w_i}{\bar{P}} = \frac{\eta-1}{\eta}\left[\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\frac{Y}{\mu}\frac{1}{\eta}\right]^{\frac{1}{\eta-1}} z_i. \tag{20}$$

The real wage is proportional to idiosyncratic technical productivity, z_i. Resulting from producers' product market power, workers receive only the inverse mark-up in the product market, $\frac{\eta-1}{\eta}$, of the producer's technical productivity. Additionally, producers pay a mark-down stemming from their monopsonistic power in the labor market, $\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}$. The mark-down is applied to the additional amount of goods sold per producer for one unit of marketing expenses, Y/μ, multiplied by the profit margin per goods sold (in terms of goods), $1/\eta$.

To solve the model further, we need to make a functional form assumption about the distribution of idiosyncratic productivity, z_i. Assuming that z_i is log-normally distributed, $z_i \sim LN(\ln \bar{z}, \Sigma^2)$, (and learned after entry) allows us to solve the model analytically.[19]

Recall that for any log-normally distributed random variable $z \sim LN(\ln \bar{z}, \Sigma^2)$

[19] Strictly speaking, we solve an approximation that ignores the upper bound on Θ_i. The support of the log-normal distribution of z_i has no upper bound and, hence, there are always some firms for which (18) produces a $\Theta_i > 1$. However, as we later show in our calibration, that fraction is small and can be made arbitrarily small by choosing the marketing cost μ appropriately without changing the model results of interest.

and real number x:

$$E(z^x) = \bar{z}^x \phi^{x^2}, \quad \text{with} \quad \phi = \exp(0.5\Sigma^2).$$

Further, observe that, since $\bar{\Theta}$ is the expected value of Θ_i, we obtain from (19) after taking expectations:

$$\left[\frac{Y}{\mu}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\frac{1}{\eta}\left(\frac{\bar{P}}{\bar{w}}/\frac{\eta}{\eta-1}\right)^{\eta-1}\right]^{-\frac{1+\eta\hat{\omega}}{\hat{\omega}(\eta-1)}} = E\left(z_i^{\frac{1+\hat{\omega}}{\hat{\omega}}}\right) = \bar{z}^{\frac{1+\hat{\omega}}{\hat{\omega}}}\phi^{\left(\frac{1+\hat{\omega}}{\hat{\omega}}\right)^2} \tag{21}$$

Dividing Equation (19) by Equation (21) allows us to express individual marketing choices in a more compact form:

$$\frac{\Theta_i}{\bar{\Theta}} = \left(\frac{z_i}{\bar{z}\phi}\phi^{-\frac{1}{\hat{\omega}}}\right)^{\frac{1+\hat{\omega}}{\hat{\omega}}}. \tag{22}$$

This equation highlights that the more a producer's productivity exceeds average productivity ($z_i > \bar{z}\phi$) the more customers it accumulates relative to the average. Furthermore, we see from (22) that $\log \Theta_i$ is normally distributed and has a larger variance than $\log z_i$ because $\frac{1+\hat{\omega}}{\hat{\omega}} > 1$. This means that the distribution of networks, i.e., the distribution of Θ_i, which is log-normal, is more right skewed than the productivity distribution. In particular, a producer with average productivity knows fewer customers than the average producer by a factor of $\phi^{-\left(\frac{1+\hat{\omega}}{\hat{\omega}^2}\right)} < 1$. This effect is the stronger the larger the variance of idiosyncratic productivity, ϕ. This means that the endogenous customer acquisition decision amplifies productivity heterogeneity. Furthermore, the difference between the productivity distribution and the distribution of networks becomes smaller as $\hat{\omega}$ increases. In the limit, $\hat{\omega} \to \infty$, productivity and customer accumulation are proportional. A stronger size-wage trade-off renders the acquisition of additional customers less attractive because wages rise too fast.

Plugging Equation (22) into the size-wage trade-off, (15), yields that producer-level wages are proportional to idiosyncratic productivity:

$$w_i = z_i \bar{w} \bar{z}^{-\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}} \phi^{-\frac{(1+\hat{\omega})^2}{\hat{\omega}(1+\eta\hat{\omega})}}, \tag{23}$$

so that marginal costs are constant across producers. Consequently, using (10), all producers charge the same price:

$$p_i = \frac{\eta}{\eta - 1}\bar{w}\bar{z}^{-\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}}\phi^{-\frac{(1+\hat{\omega})^2}{\hat{\omega}(1+\eta\hat{\omega})}}. \tag{24}$$

This, in turn, implies that each producer sells the same quantity of the intermediate good to each bundler it knows. Given that idiosyncratic prices are constant, we have, using (7), that this quantity is:

$$l_i z_i = y_i = \left(\frac{p_i}{\bar{P}}\right)^{-\eta} Y = Y(\Gamma\bar{\Theta})^{\eta/(1-\eta)}. \tag{25}$$

In particular, this also means that the real price charged to a bundler is given by:

$$\frac{p_i}{\bar{P}} = \frac{\hat{P}}{\bar{P}} = (\Gamma\bar{\Theta})^{-1/(1-\eta)}, \tag{26}$$

and, using (10), the real wage is given by:

$$\frac{w_i}{\bar{P}} = (\Gamma\bar{\Theta})^{-1/(1-\eta)}\frac{\eta - 1}{\eta}z_i. \tag{27}$$

Finally, plugging

$$\frac{\bar{P}}{\bar{w}} = (\Gamma\bar{\Theta})^{1/(1-\eta)}\frac{\eta}{\eta - 1}\bar{z}^{-\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}}\phi^{-\frac{(1+\hat{\omega})^2}{\hat{\omega}(1+\eta\hat{\omega})}}, \tag{28}$$

into Equation (21), we can derive the average network size:

$$\bar{\Theta} = \frac{Y/\Gamma}{\mu}\frac{1}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}. \tag{29}$$

Importantly, the average network size depends negatively on the size-wage elasticity, $\hat{\omega}$, as higher monopsony power discourages customer accumulation. It depends positively on market size per producer gained by a unit of marketing costs, $\frac{Y/\Gamma}{\mu}$.

This concludes our discussion of the optimal marketing choice. Before we turn to the final producer decision, namely, market entry, we point out two properties of optimal producer size. Combining (22) and (25), producer size is given by:

$$l_i\Theta_i = z_i^{1/\hat{\omega}} Y (\Gamma\bar{\Theta})^{\eta/(1-\eta)} \bar{\Theta} \left(\frac{1}{\bar{z}\phi} \phi^{-\frac{1}{\hat{\omega}}} \right)^{\frac{1+\hat{\omega}}{\hat{\omega}}} \tag{30}$$

From this equation follows immediately that producer size is increasing in idiosyncratic productivity. We note that this holds despite the fact that workers per known bundler, l_i, decrease in idiosyncratic productivity given the independence of y_i from idiosyncratic productivity. This is because more productive producers choose to know more bundlers.

Secondly, from (30), we obtain an explicit solution for the standard deviation of log employment at the producer level:

$$std\left(\log(l_i\Theta_i)\right) = std\left(\frac{1}{\hat{\omega}} \log z_i\right) = \frac{1}{\hat{\omega}}\Sigma. \tag{31}$$

That is, the distribution of log producer employment is, similar to the distribution of networks, normally distributed. Its dispersion depends positively on the dispersion of idiosyncratic productivities, Σ. Importantly, and consistent with the data in Figure 5, it depends negatively on the size-wage elasticity.[20]

4.2.3 Producer Entry

We assume free producer entry which implies that competition drives average producer profits to zero. We note that producers learn their idiosyncratic productivity level only after entry. Let $\lambda\bar{P}$ (λ is measured again in terms of the output good) be the costs to establish a producer. Given the marketing and downstream price-setting behavior, we obtain that producers enter until average operating profits, (16), equal entry costs:

$$\int \Theta_i y_i \left(p_i - \frac{w_i}{z_i}\right) - \int \mu\bar{P}\Theta_i = \lambda\bar{P} \tag{32}$$

[20]Specifically, we refer to the right-upper panel in Figure 5 which shows a positive relationship but only because the y-axis displays West-East differences, instead of East-West differences, in the standard deviation of log employment at the industry level.

which implies, using Equations (25)–(27):

$$\lambda = \int \left[\Theta_i Y \left(\Gamma\bar{\Theta}\right)^{\frac{\eta}{1-\eta}} \left(\Gamma\bar{\Theta}\right)^{-\frac{1}{1-\eta}} \left(1 - \frac{\eta-1}{\eta}\right)\right] di - \mu \int \Theta_i di \tag{33}$$

$$\lambda = \frac{Y}{\Gamma}\frac{1}{\eta} - \mu\bar{\Theta}. \tag{34}$$

This equation has an intuitive interpretation. The market entry costs in goods λ has to be equal to the goods sold per producer, Y/Γ, multiplied by the profit margin per goods sold (in terms of goods), $1/\eta$, net of expected marketing costs, $\mu\bar{\Theta}$. The steepness of the size-wage trade-off determines on which of the two margins, entry versus marketing, the profits from goods sold per producer are spent. The flatter the trade-off, the more this decision is tilted towards marketing and, thus, the larger the average and, in particular, the most productive producers become.

4.3 Equilibrium

In equilibrium, the total amount of employment needs to equal aggregate labor supply. We abstract from agglomeration effects whereby a larger economy enjoys more product varieties and is, therefore, more productive. For this reason, we fix the aggregate labor supply at one unit.[21] Hence, labor demand of all active producers, (30), integrated over all producers needs to be one:

$$\Gamma \int \Theta_i l_i di = \Gamma \int z_i^{\frac{1}{\hat{\omega}}} Y (\Gamma\bar{\Theta})^{\eta/(1-\eta)} \bar{\Theta} \left(\frac{1}{\bar{z}\phi}\phi^{-\frac{1}{\hat{\omega}}}\right)^{\frac{1+\hat{\omega}}{\hat{\omega}}} di = 1 \tag{35}$$

which solving for Y yields:

$$Y = \bar{z}\phi \underbrace{(\Gamma\bar{\Theta})^{\frac{1}{\eta-1}}}_{=\hat{P}/\bar{P}} \phi^{\frac{2}{\hat{\omega}}}. \tag{36}$$

This equation highlights key properties of the model: First, aggregate output increases not only with expected technical productivity, $\bar{z}\phi$, but also in the number

[21] If we analyzed only one geographical unit, for instance, West Germany, this would be an innocuous normalization. However, when we calibrate the model separately for East and West Germany, we make this abstraction for both regions, and, thus, disregard the possibility that East Germany is less productive simply because it is smaller.

of intermediate good producers known to the representative bundler, $\Gamma\bar{\Theta}$, through a love-of-variety effect. The effect can alternatively be expressed as the ratio of the average price charged to a bundler, $\hat{P}$, and the aggregate price index, $\bar{P}$. It reflects the fact that a larger variety of intermediate inputs used by the final goods producer increases its efficiency and, thus, lowers aggregate prices. Second, the last term, $\phi^{\frac{2}{\hat{\omega}}}$, is an effect similar to an Oi-Hartman-Abel effect (see Oi, 1961; Hartman, 1972; Abel, 1983) discussed in the investment literature. It arises through the complementarity of labor and technical productivity, z_i, which can be exploited better, by building larger networks, if a low $\hat{\omega}$ allows for a higher concentration of labor at the most productive producers.

Ultimately, Equation (36) together with the average network size, (29), and producer entry, (34), determine the aggregate equilibrium in the economy. Normalizing average producer productivity $\bar{z}\phi$ to one and solving these equations for aggregate output, the average number of known bundlers, and the share of active producers yields:

$$Y = \left(\frac{1}{\mu\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)^{\frac{1}{\eta-2}} \left(\phi^{\frac{2}{\hat{\omega}}}\right)^{\frac{1}{\eta-2}} \phi^{\frac{2}{\hat{\omega}}}, \tag{37}$$

$$\bar{\Theta} = \frac{\lambda}{\mu}\left[\frac{1}{\eta-1}\left(\frac{1+\hat{\omega}}{\hat{\omega}}\right)\right], \tag{38}$$

$$\Gamma = \frac{1}{\lambda}\frac{\eta-1}{\eta}\frac{\hat{\omega}}{1+\eta\hat{\omega}}Y. \tag{39}$$

Equation (37) shows that output is the product of three terms that are all negatively affected by the size-wage trade off. The last term, $\phi^{\frac{2}{\hat{\omega}}}$, is the aforementioned Oi-Hartman-Abel effect on output that would also be present in a monopsony model with heterogeneous producers but without product market power, as we show in Appendix F. The first two terms represent the aforementioned love-for-variety effect because a steeper size-wage trade off restricts the varieties available to bundlers. The first of the two, $\left(\frac{1}{\mu\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)^{\frac{1}{\eta-2}}$, would also be present in a model without producer heterogeneity, $\phi = 1$. The second of the two, $\left(\phi^{\frac{2}{\hat{\omega}}}\right)^{\frac{1}{\eta-2}}$, captures the interaction of product market power, monopsony power, and producer heterogeneity.

Whether one interprets the impact of $\hat{\omega}$ on the allocation of labor across differently productive producers—through $\phi^{\frac{2}{\hat{\omega}}}$ and $\left(\phi^{\frac{2}{\hat{\omega}}}\right)^{\frac{1}{\eta-2}}$—as an inefficiency depends on the ultimate source of $\hat{\omega}$. For instance, if $\hat{\omega}$ were to reflect purely preferences of workers across employers, then the output costs associated with a lower concentration of workers at productive producers cannot be interpreted as arising from misallocation (c.p. Berger, Herkenhoff, and Mongey, 2019). If, by contrast, $\hat{\omega}$ reflects institutions of collective bargaining, the output reduction arises from misallocation and therefore reflects an inefficiency. Given the positive focus of this paper, we do not need to take a strong stance on this question. We note, however, that the evidence from Table 2 and the lower panel of Figure 5, suggests that at least the difference in the level of $\hat{\omega}$ between East and West Germany is stemming from collective bargaining institutions.

From these equations also follows that aggregate labor compensation, which equals aggregate output minus entry and marketing costs, is proportional to aggregate output, where the proportionality factor is the inverse markup:

$$LC = Y - \Gamma(\lambda + \mu\bar{\Theta}) = Y\left[1 - \left(\frac{\eta-1}{\eta}\frac{\hat{\omega}}{1+\eta\hat{\omega}} + \frac{1}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)\right] = Y\frac{\eta-1}{\eta}. \quad (40)$$

This means that it is irrelevant whether we compare Y or LC differences across geographical units in what follows.

5 Implications

5.1 Quantitative Results

Given this structure, we need to determine only five parameters to evaluate the quantitative implications of our model: the standard deviation of productivity, Σ, the degree of product market power, η, the unit marketing costs, μ, the entry costs, λ, and the elasticity of wages with respect to employment, $\hat{\omega}$. Our strategy is to calibrate the model to the West German economy given the estimate of $\hat{\omega}_W = 0.078$ from Section 3.3.

Note from Equation (37) that, once we fix $\hat{\omega}$, the key parameters to understand the relative output between two regions, the main statistic of interest, are product market power, η, and the dispersion of idiosyncratic productivities, ϕ. Bundesbank (2017) finds an average price-cost margin of 1.4 in Germany, and, therefore, we set $\eta = 3.5$. We want to calibrate the dispersion for idiosyncratic productivities to match the share of employment at large plants. In the data, we identify a large plant as one that has more than 249 employees. In addition, given that the data is truncated at plants with at least 10 employees, we have to impose the same truncation in the calibration. For both reasons, we require a notion of plant size in the model. Therefore, we effectively calibrate Σ (0.16) and the entry costs, λ (0.05), jointly to match the average plant size (62 employees) and the share of employment at large plants (39%) in the data (for West Germany).

Given this calibration, marketing costs μ do not affect relative productivities between East and West Germany, and, thus, different choices would only lead to a recalibration of λ. However, the marketing cost can be pinned down by the data we use in Figure 6, and we therefore set μ to match an average West German ratio of marketing costs to sales of about one percent.

To isolate and quantify the effect of a steeper size-wage trade-off in East Germany, we start from the parameters calibrated to West Germany and change exclusively $\hat{\omega}$, setting it to the value estimated in Section 3.3 for East Germany. Table 2 displays the results of this exercise, in the column entitled "*Model East*". For the private, non-primary sector (top panel), $\hat{\omega}_W$ is 0.078 and $\hat{\omega}_E$ is 0.094.

First and importantly, by varying only $\hat{\omega}$, the model matches the moments of the plant-size distribution that were targeted in West Germany extremely well in East Germany, where they were not targeted. That is, the average plant size decreases from 62 to 45 employees compared to 46 in the data, and the share of workers employed at large plants decreases from 39 to 22 percent compared to 21 percent in the data. Second, the model, through these effects of $\hat{\omega}$ on the plant size distribution, implies a substantial drop in productivity by ten percentage points. In other words, the model explains roughly 40 percent of the observed output differences per worker between the two regions. From Equation (40) it follows that the model also rationalizes a ten percentage points lower labor compensation in the East relative to the West.

Section 3.3 shows that differences in the size-wage trade-off are particularly large in manufacturing. To investigate the implications for productivity differences,

Table 2: Size distortions and output losses: model vs. data

Variable	Model West	Model East	Data West	Data East
	Private non-primary sector $\hat{\omega}_W = 0.078$ and $\hat{\omega}_E = 0.094$			
$1/\Gamma$	61.4	44.6	61.4	46.4
$std(\log(\theta_i l_i))$	0.96	0.84	0.91	0.83
Share E > 249	0.39	0.22	0.39	0.21
Y_{east}/Y_{west}	0.90		0.74	
	Manufacturing sector $\hat{\omega}_W = 0.088$ and $\hat{\omega}_E = 0.131$			
$1/\Gamma$	98.5	57.1	98.5	64.2
$std(ln(\theta_i l_i))$	1.11	0.90	1.05	0.94
Share E > 249	0.55	0.24	0.55	0.31
Y_{east}/Y_{west}	0.84		0.70	

Notes: The table compares model simulated moments to data moments from the *SES* and German national accounts for the private, non-primary sector (top panel) and manufacturing (bottom panel). $1/\Gamma$: Average plant size, $std(\log(\theta_i l_i))$: Standard deviation of log plant size. *Share* $E > 249$: Share of employment at plants with more than 249 employees. Y_{east}/Y_{west}: Output per worker in East relative to West Germany.

we next, keeping the general calibration strategy constant, recalibrate our economy to the manufacturing sector in West Germany. The bottom panel of Table 2 shows that the average plant size in manufacturing is larger than in the total private, non-primary sector and that a larger share of workers is employed at large plants. Accordingly, we adjust the dispersion of idiosyncratic productivity, Σ (0.17), and entry costs, λ (0.82). Bundesbank (2017) finds that average price-cost margins in manufacturing are lower than in the private sector as a whole, implying $\eta = 6$.

Table 3: Decomposition of output loss

	Private non-primary	Manufacturing
Total productivity difference	10.3%	15.5%
OHA-allocation effect, $\phi^{\frac{1}{\hat{\omega}}}$	5.2%	10.6%
Love of variety (LoV), $\left(\frac{1}{\mu\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)^{\frac{1}{\eta-2}}$	1.9%	2.9%
Allocation LoV interaction, $\phi^{\frac{1}{\hat{\omega}}\frac{1}{\eta-2}}$	3.5%	2.7%

Notes: The table displays the relative output loss per worker in East relative to West Germany, $1 - Y_{east}/Y_{west}$ decomposed, to the first order, into the three channels highlighted in the discussion of Equation (37).

The panel shows that also for the manufacturing sector the difference in the size-wage trade off alone is able to explain the smaller average plant size and the lower share of employment at large plants in East Germany. Importantly, and consistent with the data, the model produces output differences in manufacturing that are larger than in the private sector as a whole. The model predicts that output is 16 percent lower in East relative to West Germany, in the data the difference is 30 percent.

Table 3 now decomposes the predicted output losses into the three channels we have highlighted in Equation (37). In the private, non-primary sector, the Oi-Hartman-Abel allocation effect that would also be present in a model without customer accumulation makes up roughly half of the total productivity loss. In manufacturing, the share of productivity losses attributed to this effect is somewhat larger.

Of the two terms representing the distortionary effect of monopsony power on network size, the interaction term is quantitatively larger than the pure love-of-variety effect that would also be present in a homogeneous producer model. In other words, the model implies that monopsony power is particularly costly, when it hinders the most productive producers to accumulate customers and grow, rendering the entire production network in the economy less efficient.

5.2 Policy Discussion

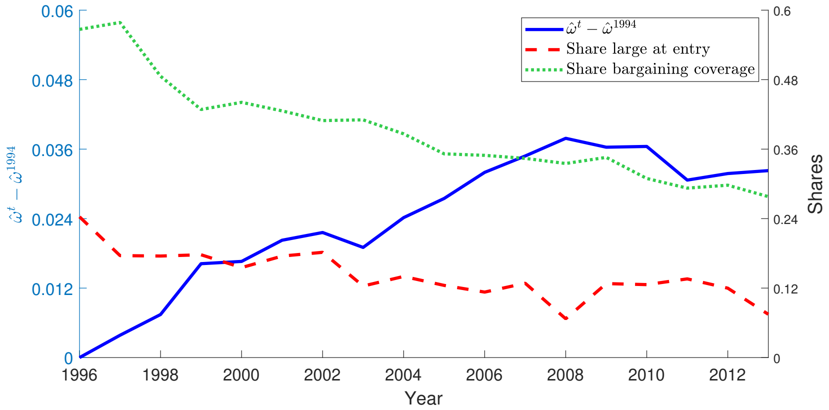
5.2.1 Subsidies

The standard output loss associated with monopsony power is due to underemployment. We deliberately abstract from this effect by assuming an inelastic labor supply. Instead, we highlight two additional sources of output loss: first labor is allocated away from large, productive plants towards small, less productive plants and, second, producers establish too small production networks by underinvesting in marketing. These two sources are not affected by untargeted wage subsidies, the standard policy instrument recommended to eliminate the distortions resulting from monopsony power, as we show in Appendix G. Intuitively, such subsidies raise the labor demand of all producers but neither change the relative distortions of labor demands nor create incentives to invest in larger networks.

The output losses we highlight are also not affected by entry subsidies, another policy tool often suggested by policy makers to help distressed regions catch up. This can be seen in Equation (37), where the entry costs, λ, do not show up. In the model, increasing the number of active producers through entry subsidies crowds out network investments one-for-one so that the equilibrium production network size, $\Gamma\bar{\Theta}$, remains unaffected, (see Equation 38). While our model is admittedly a special case with full crowding out, it highlights a general force, where existing producers adversely react to entry with their choice of production networks and thereby reduce the number of varieties available to bundlers.

Such larger production networks can be created, however, by subsidizing marketing expenses. In fact, output (net of entry and marketing costs) is not maximal when producers privately pay the marketing costs, as we show in Appendix H. The intuition for this result is the positive externality created by larger networks that render all producers more productive (in terms of the final good) by making more varieties available to bundlers.

Figure 7: Large plants, steepness of the size-wage curve, and collective bargaining over time



Notes: The figure displays for all Germany over time the share of plants covered by a collective bargaining agreement (Share bargaining coverage), the East West difference in the steepness of the size-wage curve minus its steepness difference in 1996 ($\hat{\omega}^t - \hat{\omega}^{1996}$), and the share of plants of an entering cohorts having at least 250 employees 4 quarters after entry (Share large at entry). Data sources: *AWFP* and *IAB Betriebspanel*.

The output-net-of-cost maximizing subsidy is 37% in West Germany and would increase output net of costs by 9%. Owing to the steeper size-wage curve, the optimal subsidy is slightly larger in East Germany (38%) and the output gain (again net of costs) would be 10%. Note, this subsidy works exclusively through increasing networks, i.e., the variety of products known to different bundlers, which increases their productivity. It does not alter the allocation of workers across producers, and, thus, leaves the associated output losses that explain most of the East-West productivity difference unaffected.

5.2.2 Collective bargaining coverage

Given that simple subsidies cannot fully eliminate the output losses from monopsony, another way to cast the policy discussion is to ask how to directly affect the size-wage trade-off of employers. The empirical evidence from Table 1 suggests that collective bargaining coverage is one such avenue. Employers subject to collective bargaining face flatter size-wage curves, and differences between East and West Germany even disappear once we condition on bargaining arrangements.

In turn, this raises the question whether changes in the prevalence of collective bargaining over time are also reflected in changes of the steepness of the size-wage curve and optimal plant size. Figure 7 provides suggestive time series evidence in favor. First, it shows that, in Germany, collective bargaining has substantially declined over time. In 1996, more than 55% of plants were covered by a collective bargaining agreement. This number decreases to less than 30% by 2013. Second, and in line with what one would expect from our cross-sectional evidence, this decline in collective bargaining goes along with a steepening of the size-wage curve. Finally, and in line with the cross-sectional data (industry differences across East-West) and our theory, there is a parallel trend towards smaller plant sizes (at entry). Figure 7 shows that about 24% of all employment of an entry cohort used to be at large plants in 1996. This share has declined to around 12% by 2013.

6 Conclusion

Large aggregate labor productivity differences persist across regions where governmental policies (and legal institutions enforcing these) are almost identical. We consider the case of Germany where, 23 years after reunification, the East German private, non-primary sector remains about 25% less productive today than its West German counterpart. In this context, we show that differences in collective bargaining coverage lead to higher labor market monopsony power in East relative to West Germany. We show that this monopsony power leads to aggregate labor productivity losses by reducing investment into production networks and by distorting the distribution of employment over plants. The difference in monopsony power that we estimate explains about ten percentage points of the lower labor productivity in the East German private, non-primary sector.

References

ABEL, A. B. (1983): "Optimal Investment under Uncertainty," *American Economic Review*, 73, 228–233.

ARKOLAKIS, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118, 1151–1199.

BACHMANN, R., C. BAYER, C. MERKL, S. SETH, H. STÜBER, AND F. WELLSCHMIED (2021): "Worker Churn in the Cross Section and Over Time: New Evidence from Germany," *Journal of Monetary Economics*, 117, 781–797.

BAYER, C. AND M. KUHN (2018): "Which Ladder to Climb? Wages of Workers by Job, Plant, and Education," *CEPR Discussion Paper No. DP13158.*

BELLMANN, L., B. LOCHNER, S. SETH, S. WOLTER, ET AL. (2020): "AKM Effects for German Labour Market Data," *Institut für Arbeitsmarkt-und Berufsforschung (IAB).*

BERGER, D. W., K. F. HERKENHOFF, AND S. MONGEY (2019): "Labor Market Power," *National Bureau of Economic Research No. 25719.*

BILBIIE, F. O., F. GHIRONI, AND M. J. MELITZ (2012): "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy*, 120, 304–345.

BLANCHFLOWER, D. G., A. J. OSWALD, AND M. D. GARRETT (1990): "Insider Power in Wage Determination," *Economica*, 143–170.

BLOSSFELD, H.-P. (1987): "Labor-Market Entry and the Sexual Segregation of Careers in the Federal Republic of Germany," *American Journal of Sociology*, 93, 89–118.

BOERI, T. AND K. TERRELL (2002): "Institutional Determinants of Labor Reallocation in Transition," *Journal of Economic Perspectives*, 16, 51–76.

BRAGUINSKY, S., L. G. BRANSTETTER, AND A. REGATEIRO (2011): "The Incredible Shrinking Portuguese Firm," *National Bureau of Economic Research No. 17265.*

Brown, C. and J. Medoff (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1059.

Bundesbank, D. (2017): "Mark-ups of Firms in Selected European Countries," *Monthly Report*, 53.

Burda, M. (2006): "Factor Reallocation in Eastern Germany after Reunification," *American Economic Review (Papers and Proceedings)*, 96, 368–374.

Burdett, K. and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 257–273.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and Labor Market Inequality: Evidence and some Theory," *Journal of Labor Economics*, 36, S13–S70.

Cingano, F., M. Leonardi, J. Messina, and G. Pica (2016): "Employment Protection Legislation, Capital Investment and Access to Credit: Evidence from Italy," *The Economic Journal*, 126, 1798–1822.

Dijkstra, L., H. Poelman, and P. Veneri (2019): "The EU-OECD Definition of a Functional Urban Area," *OECD Regional Development Working Papers*.

Drozd, L. A. and J. B. Nosal (2012): "Understanding International Prices: Customers as Capital," *American Economic Review*, 102, 364–95.

Einav, L., P. J. Klenow, J. D. Levin, and R. Murciano-Goroff (2021): "Customers and Retail Growth," *National Bureau of Economic Research WP No. 29561.*

Foster, L., J. C. Haltiwanger, and C. Krizan (2001): "Aggregate Productivity Growth, Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, University of Chicago Press.

Fuchs-Schündeln, N. and R. Izem (2012): "Explaining the Low Labor Productivity in East Germany - A Spatial Analysis," *Journal of Comparative Economics*, 40, 1–21.

Garicano, L., C. Lelarge, and J. Van Reenen (2016): "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 106, 3439–79.

Gourio, F. and L. Rudanko (2014): "Customer Capital," *Review of Economic Studies*, 81, 1102–1136.

Green, F., S. Machin, and A. Manning (1996): "The Employer Size-Wage Effect: Can Dynamic Monopsony Provide an Explanation?" *Oxford Economic Papers*, 48, 433–455.

Hall, R. E. and C. I. Jones (1999): "Why do some Countries Produce so much more Output per Worker than Other?" *Quarterly Journal of Economics*, 114, 83–116.

Hartman, R. (1972): "The Effects of Price and Cost Uncertainty on Investment," *Journal of Economic Theory*, 5, 258–266.

Heise, S. and T. Porzio (2021): "The Aggregate and Distributional Effects of Spatial Frictions," *NBER Working Paper No. 28792*.

Hsieh, C.-T. and P. J. Klenow (2014): "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 129, 1035–1084.

Johnson, P. and C. Papageorgiou (2020): "What Remains of Cross-Country Convergence?" *Journal of Economic Literature*, 58, 129–75.

Lamadon, T., M. Mogstad, and B. Setzler (2018): "Imperfect Competition and Rent Sharing in the US Labor Market," *University of Chicago mimeo.*

Lochner, B., S. Seth, and S. Wolter (2020): "Decomposing the Large Firm Wage Premium in Germany," *Economics Letters*, 194, 109368.

Manning, A. (2003): *Monopsony in Motion*, Princeton University Press.

——— (2011): "Imperfect Competition in the Labor Market," in *Handbook of Labor Economics*, Elsevier.

OI, W. Y. (1961): "The Desirability of Price Instability under Perfect Competition," *Econometrica*, 58–64.

SCHNABEL, C. (2005): "Gewerkschaften und Arbeitgeberverbände: Organisationsgrade, Tarifbindung und Einflüsse auf Löhne und Beschäftigung," *Zeitschrift für Arbeitsmarkt Forschung-Journal for Labour Market Research*, 38, 181–196.

SEDLÁČEK, P. AND V. STERK (2017): "The Growth Potential of Startups over the Business Cycle," *American Economic Review*, 107, 3182–3210.

SNOWER, D. J. AND C. MERKL (2006): "The Caring Hand that Cripples: The East German Labor Market After Reunification," *American Economic Review*, 96, 375–382.

STEWART, M. B. (1987): "Collective Bargaining Arrangements, Closed Shops and Relative Pay," *The Economic Journal*, 97, 140–156.

STÜBER, H. AND S. SETH (2017): "The Administrative Wages and Labor Market Flow Panel," *FAU Discussion Papers in Economics*.

SVEJNAR, J. (2002): "Transition Economies: Performance and Challenges," *Journal of Economic Perspectives*, 16, 3–28.

UHLIG, H. (2006): "Regional Labor Markets, Network Externalities and Migration: The Case of German Reunification," *American Economic Review (Papers and Proceedings)*, 96, 368–374.

WEINAND, S. AND L. VON AUER (2020): "Anatomy of Regional Price Differentials: Evidence from Micro-Price Data," *Spatial Economic Analysis*, 15, 413–440.

A Differences in Input Factors and Reallocation

In principle, lower output per worker in East Germany could be the result of differences in the quality and quantity of factor inputs and total factor productivity (TFP). TFP differences, in turn, could result from differences in access to technology or institutions (which is unlikely in the German context), differences in the capability of the labor market to reallocate workers to firms that become more productive—a sclerotic labor market in the East—or a persistent misallocation of workers to relatively unproductive plants (as in our model, where we attribute this misallocation to the disincentives of the most productive plants to acquire a large customer network).

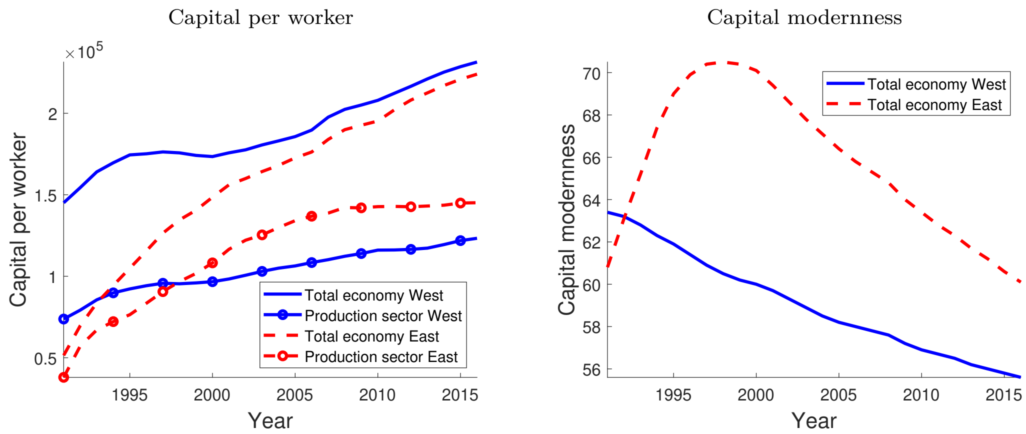
In this section, we establish that, first, differences in factor inputs are unlikely the reason behind the observed differences in output per worker. In other words, it must be TFP. Second, we show that a more sclerotic labor market in East Germany is not to blame either.

A.1 Capital and Labor Inputs

Burda (2006) puts forwards an explanation where capital accumulation is subject to frictions. The East had a lower capital stock in 1992, implying a low initial labor productivity, and if it takes time for the East to accumulate capital, this would explain a persistent productivity gap. Figure B1 (left panel) compares the (net, i.e., after depreciation) capital stock per worker in East to West Germany, and it shows that the capital stock per worker, differently from output per worker, had almost converged by 2005. In 2019, the difference in the capital stock per worker is only 3%. Thus, with a constant returns to scale Cobb-Douglas production function and a standard capital share of 30% this difference in capital intensity would explain 0.9 percentage points of labor productivity differences.

We are particularly interested in differences in the private, non-primary sector. Unfortunately, the German national accounts do not provide the capital stock by detailed industry and region. It does provide data on the production sector (manufacturing, mining, utilities, and construction), and Figure B1 (left panel) shows that, in that sector, East Germany has even overtaken the West German

Figure B1: Capital stock



Notes: The left panel displays the capital stock (after depreciation) per worker in East and West Germany. It shows is for the total economy and the production sector (manufacturing, mining, utilities, and construction). The right panel displays the modernness of the current capital stock (the ratio of the net and gross capital stock). Calculations are based on national account data.

economy in terms capital intensity by 1998.

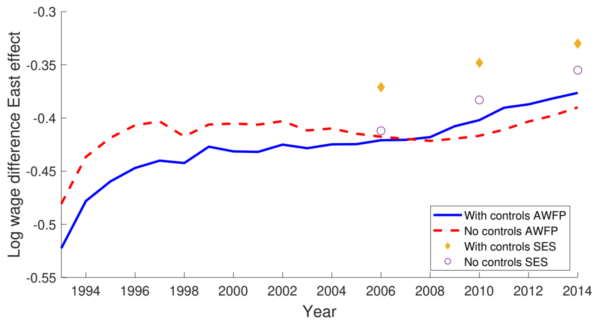
In this comparison, a confounding factor could be capital quality. Specifically, one might expect that East German plants still produce with outdated capital from before the reunification. Figure B1 in the right panel displays the modernness of the capital stock, i.e., the ratio of the net and gross capital stock. Consistent with the large catch-up in capital accumulation shown in the left panel, the capital stock is of a younger vintage in East Germany suggesting that, if anything, it is of relatively higher quality.

Another potential explanation for the lower labor productivity in East Germany could be lower quality of its labor input. If this was the case, then wage differences between East and West Germany should be explainable by measures of worker quality, such as age, sex, education, and occupation. We observe these in the *SES* and, thus, estimate, at the worker-level, the following regression for the years 2006, 2010, and 2014:

$$\ln w_{it} = \alpha_0 + East_i + F(age_{it}, sex_{it}) + educ_{it} + occ_{it} + \epsilon_{it}, \qquad \text{(B.1)}$$

where $East_i$ is a dummy that is one when when a worker works at a plant that

Figure B2: Worker quality

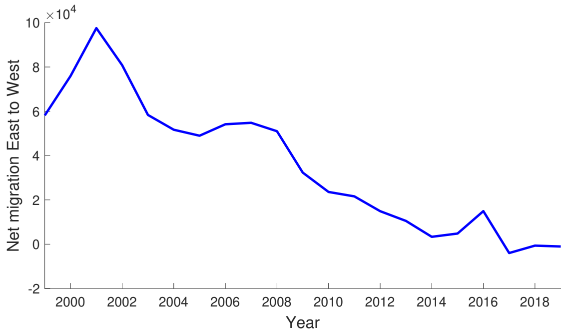


Notes: The figure displays the predicted log wage effect of a plant being located in East Germany (*No controls*) and the predicted effect of a plant being located in East Germany when controlling for worker observables (*With controls*). Estimation is based on the non-primary, private sector from either the *SES* or the *AWFP*. Worker observables in the *SES* are age and sex fully interacted, education, and occupation. Worker observables in the *AWFP* are the share of employment of workers across different age, sex, education, and task categories at the plant level.

is located in the East, and *age*, *sex*, *educ*, and *occ* are sets of dummy variables for workers' age, sex, education, and occupation, respectively. We estimate two versions of this regression, one with worker observables, age and sex fully interacted, and one without it. The restricted regression simply estimates the mean log-wage differences between East and West Germany for each year. The baseline regression provides the same but controlling for different worker skill distributions in East and West Germany. Figure B2 compares the two regressions. It shows that the mean difference in log wages and the mean difference in log wages after controlling for observable worker characteristics are very similar. Controlling for worker observables can explain some of the lower wages in East Germany but, even among similar workers, differences are about 0.35 log points between the two regions.

The *AWFP* data allows us to extend this analysis back in time. However, the *AWFP* being a plant-level data set, we can only do so at the plant level, using plant-level averaged earnings and plant-level shares of worker observables. In addition, the *AWFP* summarizes occupations in four broad groups called work

Figure B3: Net migration



Notes: The figure displays the net migration from East to West Germany. The data is from the German statistical agency (Bevölkerung und Erwerbstätigkeit Fachserie 1 Reihe 1.2).

tasks.[22] This leads to the following plant-level regression for each year:

$$\ln w_{jt} = \alpha_0 + East_j + age_{jt} + sex_{jt} + educ_{jt} + task_{jt} + \epsilon_{jt}, \tag{B.2}$$

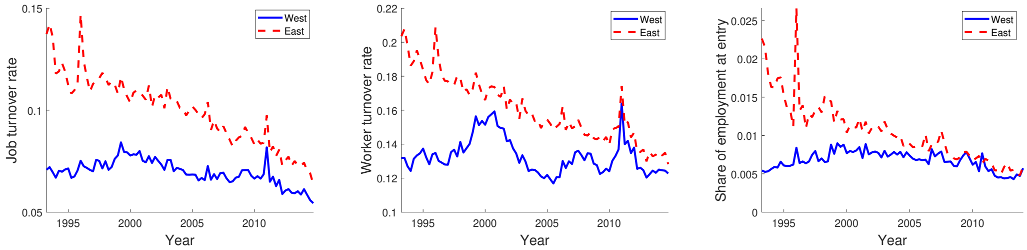
where $\ln w_{jt}$ is the log average wage at plant j in year t, $East_j$ is a dummy that is one when plant j is located in the East, and age_{jt} is the share of employment of workers across different age categories, sex_{jt} the share of employment of workers across different sex categories, $educ_{jt}$ the share of employment across different education categories, and $task_{jt}$ the share of employment across different task categories at the plant. We demean all covariates by their West-German mean and estimate again two versions of the regression, one with the covariates of worker observables and one without it.

Again, worker observables explain little of the wage differences. In fact, during the early years, worker characteristics have been somewhat better in East relative to West Germany. The relative improvement of the West German worker skill distribution has in part resulted from an outflow of workers from East Germany, see Uhlig (2006). However, as just argued, the overall distributions of qualities remain very similar in the two regions. Moreover, Figure B3 shows that net-outflows from the East to the West have converged to zero by 2013.

[22] *Low:* Agricultural occupations, elementary manual occupations, elementary personal services occupations, elementary administrative occupations, *Medium:* Skilled manual occupations, skilled services occupations, skilled administrative occupations, *Semi-high:* Technicians, associate professionals, *High:* Professional occupations, managers (Blossfeld, 1987, see).

Figure B4: Job and worker turnover rates

(a) Job turnover rate (b) Worker turnover rate (c) Entry rate



Notes: The first panel displays the job turnover rate (the sum of job creation and job destruction). The second panel displays the worker turnover rate (the sum of accessions and separations). The third panel displays the share of employment at plants entering in a quarter. Calculations are based on the *AWFP* data from the private, non-primary sector.

In line with the above, Fuchs-Schündeln and Izem (2012) and Heise and Porzio (2021) find that plant or job characteristics, rather than worker characteristics, explain the bulk of wage differences between East and West Germany even when unobserved worker heterogeneity is controlled for.

A.2 Missing Reallocation

Given that it is neither capital nor the quality of labor that explains productivity differences, any explanation must rest on TFP. In the German context, reunification has been a major shock, and one possibility might be that, even after 30 years, East Germany has failed to reallocate labor from the former state-run, unproductive plants towards more productive plants.[23] Using the *AWFP*, we show that common measures of labor market reallocation are not lower in East Germany.

To this end, we study quarterly job and worker reallocation rates as defined and explained in detail in Bachmann, Bayer, Merkl, Seth, Stüber, and Wellschmied (2021). Figure B4a displays the job turnover rate for East and West Germany. Job reallocation in the East has been relatively high following the years after reunification, likely contributing to the rapid productivity growth during these

[23] Boeri and Terrell (2002) finds that such job reallocation has indeed been important in understanding productivity growth in former Soviet Republic countries. Even for the U.S., the evidence suggests that much of long-run productivity growth is driven by the reallocation of jobs from less to more productive plants (see Foster, Haltiwanger, and Krizan (2001)).

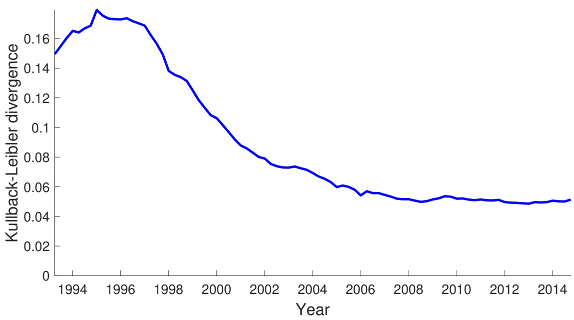
years, yet, missing reallocation does not appear to be the reason for the missing productivity convergence afterward. That is, job reallocation has remained higher in East than in West Germany throughout the sample period. In fact, the amount of job turnover in East Germany was sufficient to destroy and create every job 2.8 times since 1993.

An economy may reallocate workers across plants also without reallocating jobs, for example, to increase match quality between existing jobs and workers. Figure B4b shows that the East also does not fall short in terms of worker reallocation relative to the West. In particular, worker reallocation has been particularly high after reunification in East Germany and has nearly converged to the West level afterward.

Finally, the third panel considers one particular form of job reallocation, namely, that arising from new plant entry. It displays the share of total employment in a quarter that is due to employment at plant start-ups. Again, if anything, the East is the economy with more reallocation.

A particular form of reallocation is the growing and shrinking of industries. Therefore, since the industry composition has been significantly different in the East at the time of reunification, it could be that East Germany failed to reallocate jobs to more promising industries. To better understand the role of different industry structures between the two regions, Figure B5 plots the Kullback-Leibler divergence as a measure of the distance between the West and East German employment distribution over 21 industries. Initially, the industry distributions have been different but this difference has decreased between 1995 and 2008. Neither does the period of high productivity growth in East Germany, that is the years before 1995, coincide with a convergence in industry structure, nor does the period of convergence in industry structure, that is 1995 to 2008, show any particular pattern of aggregate productivity convergence. Most importantly, when looking at productivity differences within industries, as already seen in Figure 4, differences in output per worker are as large within sectors as in the economy as a whole: The East is less productive in each sector, and differences range from 0.44 log differences in finance to 0.08 in electricity and water supply. Hence, any yet missing convergence of the industry structure is unlikely to explain the persistent differences in output per worker between the two regions.

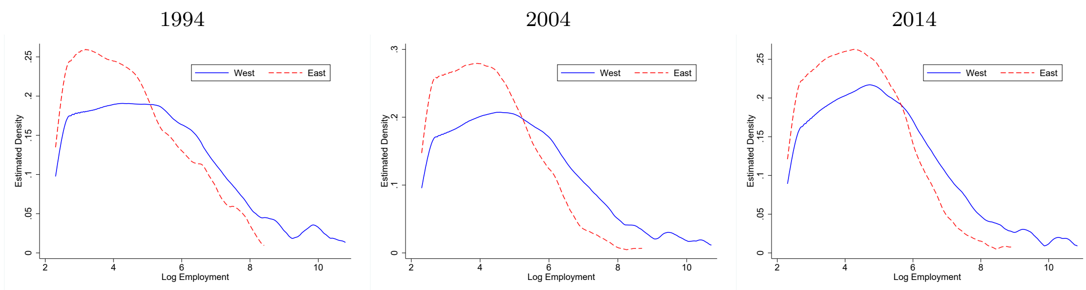
Figure B5: Industry convergence



Notes: The figure displays the Kullback-Leibler divergence index between the West and East German employment distribution over 21 industries: $KL = \sum_{i=1}^{2} 1 P(x_i) log \frac{P(x_i)}{Q(x_i)}$, where $P(x_i)$ is the employment share of industry i in the West and $Q(x_i)$ is the corresponding share in the East. Calculations are based on the *AWFP* data from the private, non-primary sector.

B Further Data on Plant Size Distributions and Wages in East and West Germany

Figure C1: Size distribution *AWFP*



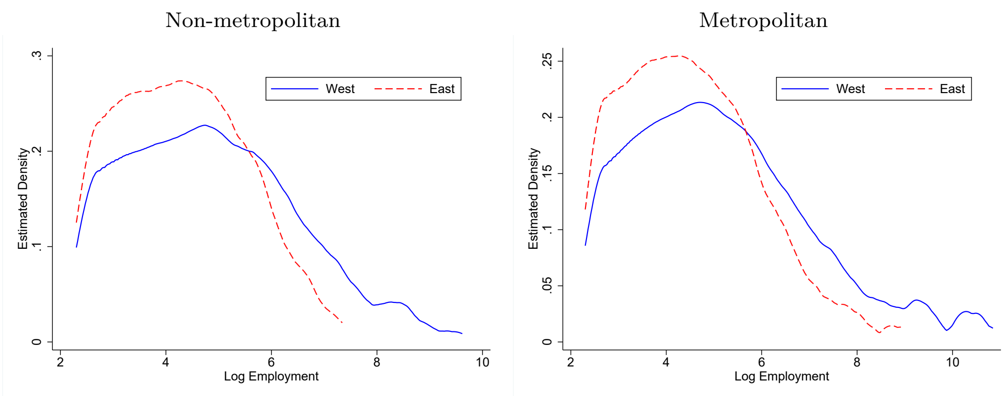
Notes: The figure displays the size distribution in East and West Germany. It displays the density function in the total private, non-primary sector in different years. Calculations are based on the *AWFP* data from the private, non-primary sector.

In this appendix, we show that differences in the plant-size distribution extend to earlier time periods and are not driven by differences in urbanization between East and West Germany. To that end, we employ the *AWFP* data going back to 1994 and use the information on plants' locations at the German "Kreis" (county) level (which are not available in the *SES*).

Figure C1 displays the density of plants over log employment in East and West Germany starting in 1994. The size distribution differences have been fairly stable between 1994 and 2014. Consistent with the *SES* data, there is some, but very little, convergence in the size distribution between 2004 and 2014.

Figure C2 displays the plant size distribution conditional on a plant being located in a metropolitan area. To define these areas, we employ the definition from Dijkstra, Poelman, and Veneri (2019). The figure shows that metropolitan areas have on average more large plants than non-metropolitan areas. Importantly, however, even within each area type, the plant size distribution in East Germany is shifted to the left relative to West Germany and displays a less fat right tail.

Figure C2: Size distribution *AWFP* metropolitan areas, 2014



Notes: The figure displays the size distribution in East and West Germany. It displays the density function in the total private, non-primary sector conditional on plants being located in a non-metropolitan area (left panel) or metropolitan area (right panel). Metropolitan areas are defined in Dijkstra, Poelman, and Veneri (2019) based on functional urban areas. Calculations are based on the *AWFP* data from the private, non-primary sector.

C Industry Classifications

Industry classifications have undergone several revisions since reunification. The *AWFP* data is organized by the so called WZ08 classification. Similarly, the 2010 and 2014 samples of the *SES* use the WZ08 classification. The 2006 sample from the *SES* uses the WZ03 classification. Finally, national accounts are organized by sectors which are based on the WZ08 classification. Table C1 provides a cross-walk across the different classifications.

Table C1: Industry classifications

	SES 2008	*SES* 2003	National accounts
MFT	10–15	15	C (MFG)
MWP	16–18/31–32/58–60	20	C (MFG)
MCP	19–23	22/25–26	C (MFG)
MME	24–25/28	30	C (MFG)
MLE	26–27	32	C (MFG)
MVE	29–30	37	C (MFG)
UTL	35–39	36/43/90	D/E (UTL)
CON	41–42	45	F (CON)
COP	43	46/47	F (CON)
WHC	45–46	48	G (TRD)
RTO	47/33	51	G (TRD)
TRA	49–51/61–63	53–54	H (TRA)
STO	52–53	57	H (TRA)
TUR	55–56	52	I (TUR)
BAN	64	63	K (FIN)
INS	65–66	64	K (FIN)
RNS	68/72–75	71	M (TPS)
TES	69–71	72	M (TPS)
RES	77		N (OPS)
BAC	78–81	78	N (OPS)
OTS	82	93	N (OPS)

Notes: The table provides a crosswalk that maps the 21 industries used in this paper into the industry classifications used by the *SES* in 2008 and 2003 and the sectors from the national accounts. *MFT*: Food and textile manufacturing, *MPW*: Paper and wood manufacturing, *MCP*: Chemical and plastic manufacturing, *MME*: Metal manufacturing, *MEL*: Electronics manufacturing, *MVE*: Vehicle manufacturing, *UTL*: Utilities, *CON*: Construction, *COP*: Construction preparations, *WHC*: Wholesale and car retail, *RTO*: Other retail, *TRA*: Transportation, *STO*: Storage, *TUR*: Tourism, *BAN*: Banking, *INS*: Insurance, *RNS*: Research services, *TES*: Technical services, *RES*: Rental services, *BAC*: Building and area care, *OTS*: Other services.

D Robustness of Size-Wage Relationship

This appendix provides a number of robustness checks to our baseline size-wage estimate. We start with worker-level data from the *SES* followed by analyses with plant-level data from the *AWFP*.

D.1 Worker-Level Data

Table D1: More on the size-wage relationship

	Non-primary private sector		
	Quadratic	Cubic	Adding part-time
Difference in elasticities, $\hat{\omega}_E - \hat{\omega}_W$	1.9 (0.3)	1.5 (0.3)	2.0 (0.2)
N (in thousands)	2365	2365	3074

Note: The table displays the estimated difference in the size-wage relationships for the non-primary private sector in West and East Germany. Standard errors are in parentheses. All coefficients are multiplied by 100 for better readability. *Quadratic*: Controls for a workers' age and sex by a full set of dummy-interactions, plus time and industry fixed effects and a quadratic trend in size that is common across the regions. *Cubic*: Controls for a workers' age and sex by a full set of dummy-interactions, plus time and industry fixed effects and a cubic trend in size that is common across the regions. *Adding part-time*: The same as our baseline estimate but including part-time workers in the sample. Data source: *SES*.

In Section 3.3, we assume that the size-wage relationship is log-linear. It is possible that the true relationship is non-linear and the steeper estimate for the size-wage relationship in East Germany simply captures this non-linearity. For instance, if the plant-size relationship was steeper for small plants, the steeper average wage-size relationship in East Germany would simply reflect that there are more small plants there. To allow for this possibility, we augment the regression (1) by a common non-linear term, $F(\ln E_{it})$, that takes the form of either a 2-nd order or 3-rd order polynomial:

$$\ln w_{it} = \beta_0 + \beta_E \, East_i + \hat{\omega}_W \ln E_{it} + F(\ln E_{it}) + (\hat{\omega}_E - \hat{\omega}_W) \, East_i \ln E_{it} + \beta x_{it} + e_{it}. \tag{D.1}$$

The first column of Table D1 shows that allowing for a 2nd order polynomial implies, if anything, an even steeper size-wage curve in East relative to West

Germany. Using instead a 3-rd order polynomial yields almost the same difference between East and West Germany as does the baseline, linear, specification.

Furthermore, recall that we compute the baseline estimate using a sample of full-time workers. The distribution of full-time and part-time workers in East and West Germany is somewhat different, and, hence, it is natural to ask whether our results are robust to including part-time workers. The third column in Table D1 displays estimates of the size-wage relationship in East and West Germany when we include part-time workers. This leads, if anything, again to an even steeper size-wage curve in East relative to West Germany.

D.2 Plant-Level Data

In Section 3.3, we control for worker heterogeneity and sorting by observable worker characteristics: age, sex, education, occupation, and job levels. The plant-level AWFP data together with the work of Bellmann, Lochner, Seth, Wolter et al. (2020) allows us to control for unobserved worker heterogeneity, too. Specifically, Bellmann, Lochner, Seth, Wolter et al. (2020) estimate the following regression for all German plants for three time periods (1998-2004, 2003-2010, and 2010-2014) using the matched employer-employee data from the German social security:

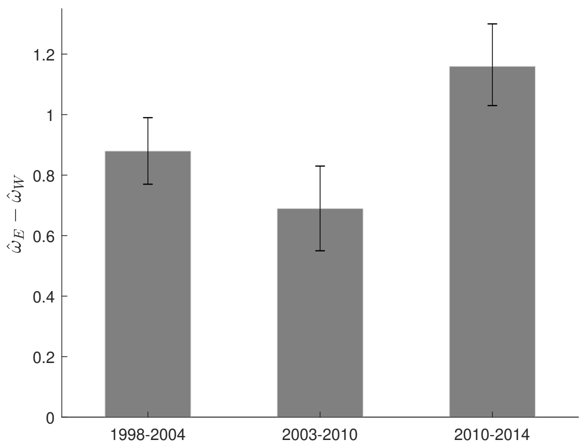
$$\ln w_{ijt} = \alpha_0 + \alpha x_{jt} + \phi_i + \gamma_j + \epsilon_{ijt}, \tag{D.2}$$

where w_{ijt} are the daily earnings of worker i at plant j in period t, x_{it} are time-varying worker observables, γ_j is a worker fixed-effect, and ϕ_i is a plant fixed effect. They provide the estimated plant-fixed effect which we match to the AWFP data. This plant fixed effect equals the average wage of a plant controlling for its worker characteristics (observed and unobserved). We then can use this average wage in our size-wage regression. That is, we estimate the following regression:

$$\phi_i = \beta_0 + \beta_E\, East_i + \hat{\omega}_W \ln E_i + (\hat{\omega}_E - \hat{\omega}_W)\, East_i \ln E_i + e_i. \tag{D.3}$$

Figure D1 plots the estimates for $\hat{\omega}_E - \hat{\omega}_W$ all three sample periods. Again, we find that East Germany faces a relatively steeper size-wage, more precisely size-daily-earnings, relationship. In other words, the regression suggests that our baseline

Figure D1: Plant-level size-wage differences



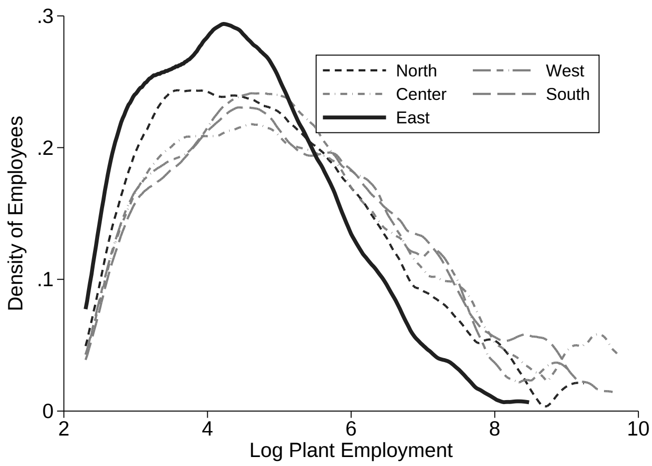
Notes: The figure displays the difference in the size-wage, more precisely size-daily-earnings, relationship between East and West Germany when the size-wage relationship is estimated using plant-level data. It plots the OLS estimate of a regression of the log plant fixed effect of wages on log plant size. Error bands are estimated using asymptotic heteroskedastic robust standard errors. The plant-level fixed effects are provided by the *IAB*. Calculations are based on the *AWFP* data from the private, non-primary sector.

finding is not driven by sorting on unobservables.

Relatedly, one can ask whether through the lens of worker fixed effects (instead of the plant fixed effect) one can find evidence for such sorting. Lochner, Seth, and Wolter (2020) (c.f. their Table B.4) shows that this is not the case. If anything, high-skilled workers sort more into large plants in West Germany which is consistent with our observation in Section 3.3 that the difference in the steepness of the size-wage curve becomes more pronounced the more we control for additional worker observables.

E Analysis with a Finer Regional Resolution for West Germany

Figure E1: Plant-size distributions



Notes: The figure displays the employment-weighted plant size distributions for five German regions, subdividing West Germany in four regions. It displays an estimated density function (by a Gaussian kernel smoother) in the total private, non-primary sector.

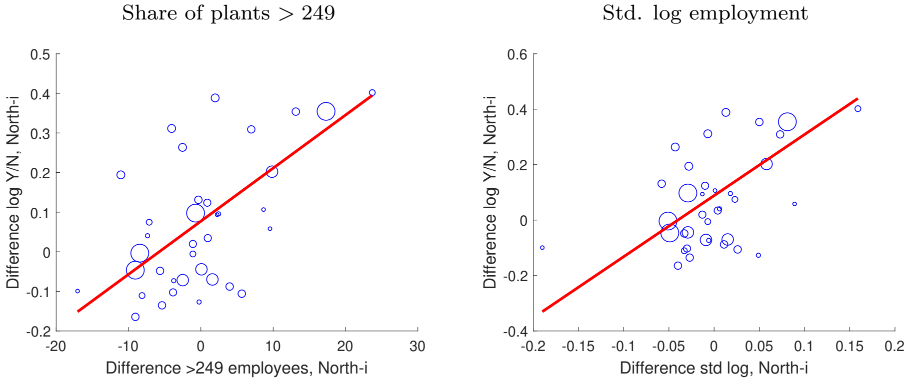
Our baseline analysis distinguishes only between East and West Germany. The *SES* data allows us to distinguish between five regions in total: North, West, Center, South, and East.[24] This appendix extends the analysis and exploits the additional variation coming from the four regions within West Germany.

Figure E1 displays the plant-size distributions for all five regions. It first shows a visible distinction between the East and all West German plant-size distributions. East Germany has, by far, the most missing large plants. Second, there is some variation also among the West German regions, which we exploit in the following analysis.

Figure E2 is the analog to Figure 4 in the main text. Those sector/region combinations that have particularly low output per worker also have relatively

[24]North: Schleswig Holstein, Hamburg, Bremen, Berlin, and Lower Saxony; West: Northrhine-Westphalia; Center: Hesse, Rhineland Palatinate, and Saarland; South: Baden-Württemberg and Bavaria; East: Thuringia, Saxony, Saxony-Anhalt, Mecklenburg Western Pomerania, and Brandenburg.

Figure E2: Productivity differences and large plants by sector



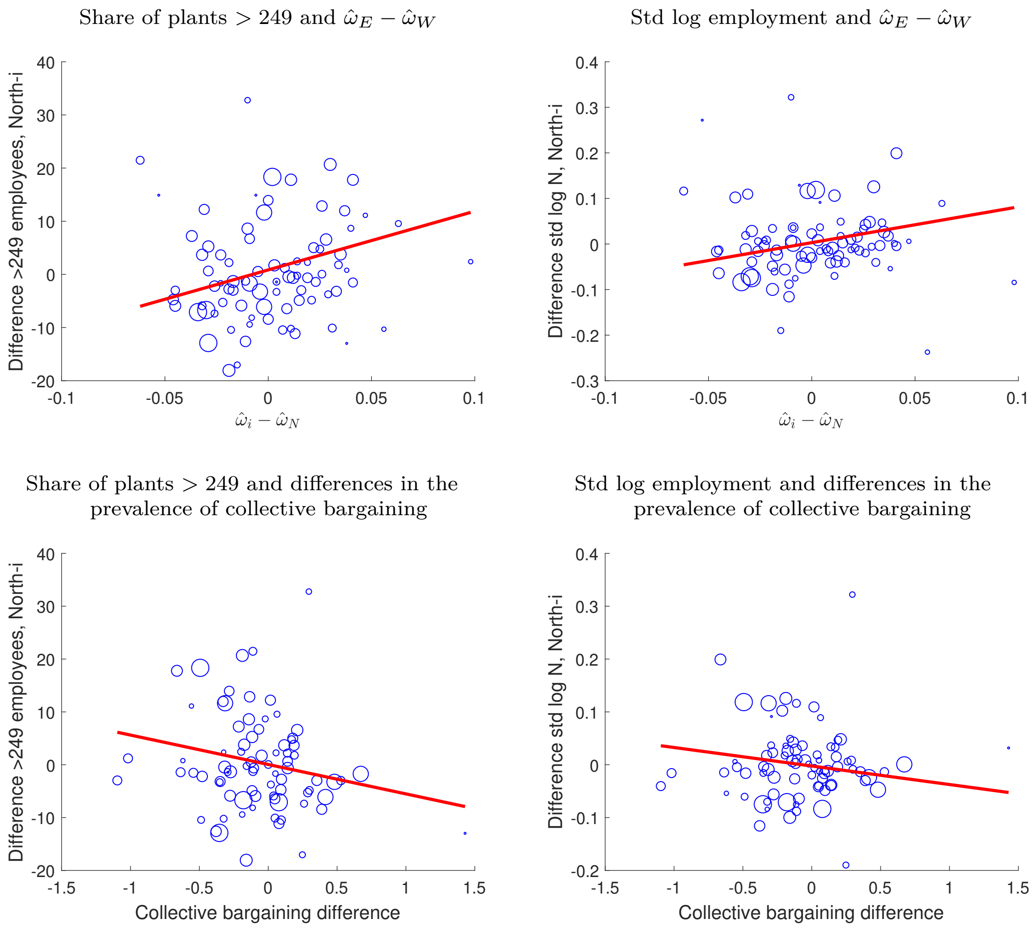
Note: The figures plot log differences in output per worker in 2014 against the share of employment at plants with more than 249 employees (left) and the standard deviation of log plant employment (right). Each dot represents a sector/region combination and displays the difference to the same sector in the North region in Germany. The lines show weighted-least squares regressions. Data sources: *SES* and *VGR*.

few large plants operating in that sector/region. Table E1 shows that these relationships are statistically significant at the 1% level.

Next, we produce with Figure E3 the analog to Figure 5 in the main text. That is, we use data for 21 industries paired with the five regions to revisit the relationship between missing large plants and a steeper size-wage relationships. The figure displays in the two top panels for each sector within each region the difference in the size-wage relationship against the difference in the share of employment at large plants (left) and the difference in the standard deviation of log employment (right). Those industry-region combinations that have particularly steep size-wage curves also have relatively few large plants operating in that industry/region. Table E1 shows that these relationships are again statistically significant.

The bottom panels of Figure E3 show on the x-axes, for each industry, a double difference in the prevalence of collectively bargained wage contracts between large and small plants. This double difference is plotted against our two measures of differences in the plant size distribution: the share of employment at large plants (left panel) and the standard deviation of log plant-level employment (right panel). The relationship between collective-bargaining prevalence differences and plant-size differences is negative. Industry-region combinations in which the prevalence

Figure E3: The share of large plants, the size-wage relationship, and collective bargaining



Note: The top panels plot the share of employment at plants with more than 249 employees (left) and the standard deviation of log plant employment (right) against the size-wage relationship. Each dot represents a sector/region combination and displays the difference to the same sector in the North region in Germany. The bottom panel relates sectoral differences across regions in the share of employment at large plants and the standard deviation of log plant employment to the following double difference: $\log P(C|L, R_1) - \log P(C|S, R_1) - [\log P(C|L, R_i) - \log P(C|S, R_i)]$, where $P(C|\cdot)$ is the conditional probability of a worker being subject to collective bargaining in our sample in (L)arge (>249 employees) or (S)mall ($\leq$ 249 employees) plants in region 1 and region i. The lines show weighted-least square regressions. Data sources: Calculations using the *SES*.

Table E1: P-values

	East-West	East-West + finance dummy
Y/N, > 249	0.024	0.048
Y/N, Std log	0.037	0.070
> 249, $\hat{\omega}_E - \hat{\omega}_W$	0.185	0.021
Std log. $\hat{\omega}_E - \hat{\omega}_W$	0.142	0.030
> 249, Collective	0.244	0.155
Std log, Collective	0.079	0.055
	All regions	All regions + finance dummy
Y/N, > 249	0.000	0.000
Y/N, Std log	0.000	0.000
> 249, $\hat{\omega}_E - \hat{\omega}_W$	0.007	0.005
Std log, $\hat{\omega}_E - \hat{\omega}_W$	0.011	0.006
> 249, Collective	0.072	0.070
Std log, Collective	0.130	0.110

Notes: The table displays p-values (two-sided tests) from the regression lines in Figures 4, 5, E2, and E3. The last column repeats the regressions from the second column adding a dummy for the financial sector, taking into account that this sector is particular in terms of its branching structure and therefore the definition of a plant as a production unit.

of collectively bargained wages increases relatively more in plant size are those industry-region combinations where large plants are particularly missing. Table E1 shows that the relationships are marginally statistically significant.

F A Simple Model of Monopsony Power

This appendix is to argue that the combined presence of customer accumulation, imperfect competition in the goods markets, and endogenous producer entry has additional output effects over and above those present in a simple model of monopsony power in the labor market. For this purpose, consider the simplified version of our model of Section 4 without imperfect competition in the goods market, no customer accumulation, and no endogenous producer entry. Producers hire labor, l_i, and combine it with their idiosyncratic productivity, z_i, to produce a homogeneous output good, y_i. We assume again that a producer's wage, relative to the average wage, is log-linear in its size, l_i:

$$w_i = \left(\frac{l_i}{\bar{l}}\right)^{\hat{\omega}}, \tag{F.1}$$

where again we normalize the wage at the average plant size to unity, making labor the numeraire. Hence, producers' profits are given by their revenues minus labor costs:

$$\Pi_i = Pz_i l_i - l_i \left(\frac{l_i}{\bar{l}}\right)^{\hat{\omega}}. \tag{F.2}$$

Taking the first-order condition with respect to labor and rearranging gives a producer's optimal size as a function of its idiosyncratic productivity:

$$l_i = \bar{l} z_i^{\frac{1}{\hat{\omega}}} \left(\frac{P}{1+\omega}\right)^{\frac{1}{\hat{\omega}}}. \tag{F.3}$$

Labor market clearing implies that total labor demand is equal to the total labor supply of one. Hence, taking expectations of (F.3), where we again assume that z_i is log-normally distributed, yields

$$\int l_i di = \bar{l} \bar{z}^{\frac{1}{\hat{\omega}}} \phi^{\frac{1}{\hat{\omega}^2}} \left(\frac{P}{1+\omega}\right)^{\frac{1}{\hat{\omega}}} = 1. \tag{F.4}$$

Dividing (F.3) by (F.4) to eliminate P and rearranging yields:

$$l_i = z_i^{\frac{1}{\hat{\omega}}} \bar{z}^{-\frac{1}{\hat{\omega}}} \phi^{-\frac{1}{\hat{\omega}^2}}. \tag{F.5}$$

It follows that the output of each producer is:

$$y_i = z_i l_i = z_i^{\frac{1+\hat{\omega}}{\hat{\omega}}} \bar{z}^{-\frac{1}{\hat{\omega}}} \phi^{-\frac{1}{\hat{\omega}^2}}. \tag{F.6}$$

Finally, taking expectations and normalizing average productivity, $\bar{z}\phi$, to one as in the main text, gives total output as:

$$Y = \int y_i di = \phi^{\frac{2}{\hat{\omega}}}, \tag{F.7}$$

which is the analog to (37)

$$Y = \left(\frac{1}{\mu\eta} \frac{1+\hat{\omega}}{1+\eta\hat{\omega}} \right)^{\frac{1}{\eta-2}} \left(\phi^{\frac{2}{\hat{\omega}}} \right)^{\frac{1}{\eta-2}} \phi^{\frac{2}{\hat{\omega}}},$$

which determines output in our main model. Comparing the two equations highlights two mechanisms of the baseline model. First, without customer search and endogenous entry, there is no output distortion resulting from the average producer building too small a customer network, $\left(\frac{1}{\mu\eta} \frac{1+\hat{\omega}}{1+\eta\hat{\omega}} \right)^{\frac{1}{\eta-2}}$, which itself is a result of monopsony power. Second, the pure-monopsony model misses the interaction of product market power and labor market power, such that the term $\left(\phi^{\frac{2}{\hat{\omega}}} \right)^{\frac{1}{\eta-2}}$ is absent. It does, however, feature the interaction between heterogeneity with labor market power, such that labor market monopsony reduces output also by $\phi^{\frac{2}{\hat{\omega}}}$. As a corollary, it follows that, with homogenous producers and fixed labor supply, there is no output loss from monopsony power in the labor market.

G A Wage Subsidy

The standard output loss associated with monopsony power is underemployment. Given our assumption of exogenous labor supply this is absent in our model. Instead, Section 4 identifies two additional output losses: Allocation of workers away from the most productive producers and underinvestment into producer networks. This section shows that the standard policy tool to overcome the problem of underemployment, a (proportional) wage subsidy, fails to address these two additional output losses in our model. The intuition for this is, before laying out the argument formally, as follows: With constant elasticity in goods-demand, all producers charge the same markup and thus all prices (relative to wages) move down proportionally with the subsidy. This leaves the share of an individual producer in total output of a bundler unchanged, if the individual producer's relative wage does not change. This also means that individual employment per known bundler is constant relative to total. With the isoelastic producer-specific labor supply it also turns out that the individual share of known bundlers relative to the average is constant. In the end, all incentives to accumulate customers change proportionally with the subsidy. Altogether, this means that the individual share in total employment remains unchanged and hence, because this share is the only determinant of an individual producer's relative wage, these relative wages indeed remain unchanged. This leaves entry as the only potential margin to be effected by the subsidy. The subsidy increases, ceteris paribus, the profits of active producers and should thus spur entry. However, with fixed labor supply, average wages adjust one-for-one with the subsidy eliminating the extra entry incentive as well as any aggregate incentive to accumulate more customers.

The formal exposition of this argument follows closely the model of Section 4 and, thus, we will be brief here. Producers receive a proportional wage subsidy, τ. Hence, they set prices as a mark-up over their real marginal costs

$$p_i = \frac{\eta}{\eta - 1} \frac{w_i}{z_i}(1 - \tau), \tag{G.1}$$

i.e., the wage subsidy raises the labor demand of each producer for each final good producer that it knows. From this follows the gross profits as a function of known bundlers:

$$\pi(\Theta_i) \;=\; \Theta_i \left(\frac{w_i}{z_i}\right)^{1-\eta} \bar{P}^{\eta} Y \frac{(\eta-1)^{\eta-1}}{\eta^{\eta}} (1-\tau)^{1-\eta}. \tag{G.2}$$

Moreover, using the wage equation, we can derive again wages as a function of the number of known bundlers as well as productivity and aggregates:

$$w_i = z_i^{\frac{(\eta-1)\hat{\omega}}{1+\eta\hat{\omega}}} \bar{w} \left(\frac{\Theta_i}{\bar{\Theta}}\right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}} (1-\tau)^{-\frac{\eta\hat{\omega}}{1+\eta\hat{\omega}}}, \tag{G.3}$$

where $\bar{w} = \left(\bar{P}^{\eta} Y \left(\frac{\eta}{\eta-1}\right)^{-\eta} / \bar{l}\right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}}$ summarizes the other aggregate terms that affect wages. Using this together with the gross profits, (G.2), and subtracting marketing expenditures yields the operating profits:

$$\Pi_i = \Theta_i \left(\frac{\Theta_i}{\bar{\Theta}}\right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}(1-\eta)} z_i^{-(1-\eta)\frac{\hat{\omega}+1}{1+\eta\hat{\omega}}} \bar{P}^{\eta} Y \frac{(\eta-1)^{\eta-1}}{\eta^{\eta}} \bar{w}^{1-\eta} (1-\tau)^{\frac{1-\eta}{1+\eta\hat{\omega}}} - \mu \bar{P} \Theta_i. \tag{G.4}$$

Solving the associated first order condition for Θ_i yields again a relationship between the optimal amount of known bundlers to a producer's idiosyncratic productivity:

$$\frac{\Theta_i}{\bar{\Theta}} = z_i^{\frac{1+\hat{\omega}}{\hat{\omega}}} \left[\frac{Y}{\mu} \frac{1+\hat{\omega}}{1+\eta\hat{\omega}} \frac{1}{\eta} \left(\frac{\bar{P}}{\bar{w}} / \frac{\eta}{\eta-1} \right)^{\eta-1} \right]^{\frac{1+\eta\hat{\omega}}{\hat{\omega}(\eta-1)}} (1-\tau)^{-\frac{1}{\hat{\omega}}}. \tag{G.5}$$

This equation, at first glance, seems to suggest that a wage subsidy indeed increases relative customer accumulation proportionally for all firms. However, logically this is impossible and thus, by using the definition of $\bar{\Theta}$, the subsidy term drops and we get back to the same equation (c.f. equation 22) that determines the relative size of the customer network:

$$\frac{\Theta_i}{\bar{\Theta}} = \left(\frac{z_i}{\bar{z}\phi} \phi^{-\frac{1}{\hat{\omega}}} \right)^{\frac{1+\hat{\omega}}{\hat{\omega}}} \tag{G.6}$$

Using this equation, we can now derive the optimal producer-level behavior:

$$w_i = z_i \bar{w} \bar{z}^{-\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}} \phi^{-\frac{(1+\hat{\omega})^2}{\hat{\omega}(1+\eta\hat{\omega})}} (1-\tau)^{\frac{-\eta\hat{\omega}}{1+\eta\hat{\omega}}} \tag{G.7}$$

$$p_i = \frac{\eta}{\eta-1} \bar{w} \bar{z}^{-\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}} \phi^{-\frac{(1+\hat{\omega})^2}{\hat{\omega}(1+\eta\hat{\omega})}} (1-\tau)^{\frac{1}{1+\eta\hat{\omega}}} \tag{G.8}$$

$$\frac{p_i}{\bar{\bar{P}}} = \frac{\hat{P}}{\bar{\bar{P}}} = (\Gamma\bar{\Theta})^{-1/(1-\eta)} \tag{G.9}$$

$$l_i z_i = y_i = \left(\frac{p_i}{\bar{\bar{P}}}\right)^{-\eta} Y = Y(\Gamma\bar{\Theta})^{\eta/(1-\eta)}. \tag{G.10}$$

From (G.10) follows that relative output and hence relative employment per bundler do not depend on the wage subsidy. Together with (G.6) this implies that total relative emoployment remains unchanged. Hence, the subsidy cannot cure the output loss resulting from a reallocation of labor away from more to less productive producers.

It still could be that the subsidy promotes entry. The producers' free entry condition reads:

$$\int \Theta_i y_i \left(p_i - \frac{w_i}{z_i}(1-\tau)\right) - \int \mu \bar{P} \Theta_i = \lambda \bar{P}, \tag{G.11}$$

which, after aggregation and using (G.7) - (G.10), yields:

$$\lambda = \frac{Y}{\Gamma}\frac{1}{\eta} - \mu\bar{\Theta}. \tag{G.12}$$

Similarly, we can derive again the average network size:

$$\bar{\Theta} = \frac{Y/\Gamma}{\mu}\frac{1}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}, \tag{G.13}$$

where again τ does not show up explicitly.

Finally, labor market clearing implies that also Y is independent of τ because

$$\Gamma \int \Theta_i l_i di = \Gamma \int z_i^{\frac{1}{\hat{\omega}}} Y (\Gamma\bar{\Theta})^{\eta/(1-\eta)} \bar{\Theta} \left(\frac{1}{\bar{z}\phi}\phi^{-\frac{1}{\hat{\omega}}}\right)^{\frac{1+\hat{\omega}}{\hat{\omega}}} di = 1 \tag{G.14}$$

yields for Y:

$$Y = \bar{z}\phi(\Gamma\bar{\Theta})^{\frac{1}{\eta-1}}\phi^{\frac{2}{\bar{\omega}}}. \tag{G.15}$$

This means that τ does not show up in the equilibrium conditions (G.12), (G.13), and (G.15), which are the same as without the subsidy. This concludes the argument.

H A Marketing Subsidy

In Section 4, producers maximize profits given their private marketing costs μ. Yet, individual private marketing expenditures create a positive externality by increasing the network size that producers build and, thus, increase the productivity of the bundlers which also means that all producers in the network become more productive in producing final output. A Ramsey planner that can freely choose the marketing costs that private producers face, $\tilde{\mu}$, keeping the physical marketing costs μ in place, would maximize output minus real costs, i.e., labor compensation:

$$LC = Y - \Gamma(\lambda + \mu\bar{\Theta}), \tag{H.1}$$

subject to the optimal employment, customer accumulation, and entry decision of producers:

$$Y = \left(\frac{1}{\tilde{\mu}\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)^{\frac{1}{\eta-2}} \left(\phi^{\frac{2}{\hat{\omega}}}\right)^{\frac{1}{\eta-2}} \phi^{\frac{2}{\hat{\omega}}} \tag{H.2}$$

$$\bar{\Theta} = \frac{\lambda}{\tilde{\mu}}\left[\frac{1}{\eta-1}\left(\frac{1+\hat{\omega}}{\hat{\omega}}\right)\right] \tag{H.3}$$

$$\Gamma = \frac{1}{\lambda}\frac{\eta-1}{\eta}\frac{\hat{\omega}}{1+\eta\hat{\omega}}Y. \tag{H.4}$$

Combining these equations yields:

$$LC = Y - Y\left(\frac{\eta-1}{\eta}\frac{\hat{\omega}}{1+\eta\hat{\omega}} + \frac{\mu}{\tilde{\mu}}\frac{1}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right). \tag{H.5}$$

The corresponding first-order condition is given by:

$$\frac{\partial Y}{\partial\tilde{\mu}} - \frac{\partial Y}{\partial\tilde{\mu}}\left(\frac{\eta-1}{\eta}\frac{\hat{\omega}}{1+\eta\hat{\omega}} + \frac{\mu}{\tilde{\mu}}\frac{1}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right) + \frac{1}{\tilde{\mu}^2}\frac{\mu}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}Y = 0, \tag{H.6}$$

where, using (H.2),

$$\frac{\partial Y}{\partial\tilde{\mu}} = -\frac{1}{\eta-2}Y\frac{1}{\tilde{\mu}}, \tag{H.7}$$

and, hence,

$$1-\left(\frac{\eta-1}{\eta}\frac{\hat{\omega}}{1+\eta\hat{\omega}}+\frac{\mu}{\tilde{\mu}}\frac{1}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)-\frac{\mu}{\tilde{\mu}}\frac{\eta-2}{\eta}\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}=0. \tag{H.8}$$

Rearranging yields:

$$\frac{\mu}{\tilde{\mu}}=\frac{\eta}{\eta-1}\frac{1+\hat{\omega}(\eta-\frac{\eta-1}{\eta})}{1+\hat{\omega}}. \tag{H.9}$$

That is, the optimal subsidy is positive ($\frac{\eta}{\eta-1}>1$ and $\frac{1+\hat{\omega}(\eta-\frac{\eta-1}{\eta})}{1+\hat{\omega}}>1$ if $\eta>2$) and grows in $\hat{\omega}$.

However, it only addresses the too small average network size. It does not affect the allocation of workers to relatively unproductive plants. This follows from the observation that the first-order condition is independent of the Oi-Hartman-Abel allocation effect, $\phi^{\frac{1}{\hat{\omega}}}$, as well as the interaction effect between producer heterogeneity, product market power, and labor market power, $\phi^{\frac{1}{\hat{\omega}}\frac{1}{\eta-2}}$. Put differently, a subsidy on marketing expenditures cannot alter the output losses associated with these two effects.