

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 9, 2022, [Current version](#)

Abstract

When employers face a trade-off between being large and paying low wages—and in this sense have monopsony power—some productive employers decide to acquire few customers, forgo sales, and remain small. These decisions have 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 steeper size-wage curves, invest less into marketing, remain smaller, and are less productive. A model with labor market monopsony, product market power, and customer acquisition matching these features of the data predicts ten percent lower aggregate labor productivity in East Germany.

Keywords: aggregate productivity, plant heterogeneity, collective bargaining, monopsony power, size-wage curve, customer capital, size distortions

JEL: E20, E23, E24, J20, J42, J50

*Rüdiger Bachmann: University of Notre Dame, rbachman@nd.edu; Christian Bayer (corresponding author): Rheinische Friedrich-Wilhelms-Universität Bonn, christian.bayer@uni-bonn.de; Heiko Stüber: Institut für Arbeitsmarkt- und Berufsforschung, Heiko.Stueber@iab.de; Felix Wellschmied: Universidad Carlos III de Madrid, fwellsch@eco.uc3m.es.

We thank seminar participants at the 28th CEPR ESSIM, the 2021 Oxford Conference on Firm Heterogeneity and the Macroeconomy, the 2021 NBER-SI, the 2021 Macroeconomics Section Meeting of the German Economic Association, SITE 2021, the CRC TR-224 workshop, the IV Workshop Spanish Macroeconomics Network, the Lisbon Macro Workshop, the FRB of Philadelphia, the IZA workshop on the Macroeconomics of Labor Productivity, and the Sächsische Staatskanzlei as well as at the universities Erlangen-Nürnberg, Bayreuth, Kiel, FU Berlin, Mainz, Kent, Manchester, Stockholm School of Economics, Lund, Linz, Salzburg, WU Vienna, Tübingen, Heidelberg, Leipzig, TU Berlin, and North Carolina. Christian Bayer gratefully acknowledges funding from the European Research Council under the European Union’s Seventh Framework Program (FTP/2007-2013) / ERC Grant agreement no. 282740 and support by the DFG through CRC TR-224. Christian Bayer was a 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 MICIN/ AEI/10.13039/501100011033, grant *PID2020-117354GB-I00*, and the Comunidad de Madrid through research grants EPUC3M11 (V PRICIT) and H2019/HUM-589 and thanks the Department of Economics at ITAM for its hospitality. Heiko Stüber gratefully acknowledges 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. This decline did not happen uniformly but was most pervasive at small plants. In this paper, we investigate the macroeconomic and misallocation consequences when employers strategically adjust to selective union retrenchment. The East and West German labor markets provide a good laboratory to study these consequences. Both regions share the same legal and, by and large, cultural institutions. In East Germany, however, collective bargaining and union membership is more skewed towards large plants. In the communist German Democratic Republic, trade unions did not represent worker interests. As a consequence, after reunification, union membership fell dramatically (see [Schnabel, 2005](#)); most pronounced at small plants.

As a result of this non-uniform union decline, the size-wage curve for plants is steeper in East than in West Germany. This creates disincentives in the East to choose large-scale business models. Consequently, even the most productive plants create relatively small customer networks and hire relatively few workers. 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 Germany, and the disincentives from a steeper size-wage curve explain at least ten percentage points of this gap.

We arrive at this conclusion by employing high-quality administrative wage data which we combine with a new heterogeneous-firm model. In this model, plants have product market power and face an upward-sloping size-wage curve when they decide about entry, their customer networks, and the size of their workforce. Our model thus tractably combines elements from two recent, but separate, strands of the heterogeneous-firm literature: We marry the literature on monopsony power of heterogeneously productive plants (see, e.g., [Berger, Herkenhoff, and Mongey, 2022](#)) with that on customer capital accumulation (see, e.g., [Sedláček and Sterk, 2017](#); [Arkolakis, 2010](#)). This marriage allows us to highlight a new long-term distortion due to monopsony power, i.e., of not being a price taker in the labor market. In our particular application, plants not being price takers results from the unions' focus on plants with business models that require a large scale of operations in the long run. Plants will take this into account and skew their scale decisions towards small-scale business models. Economically, this is a distortion

which not only generates sizable aggregate productivity losses but also explains, parsimoniously, the differences in the plant size distributions between East and West Germany.

We begin by documenting that, in the data, aggregate and industry differences in labor productivity and wages are systematically related to the absence of large plants in East Germany. The share of employment at large plants with more than 249 employees is almost twice as large in the West. In industry-level data, there is a positive correlation between missing large plants and the East-West productivity/wage gap. For example, vehicle manufacturing has both a particularly large East-West gap in labor productivity (36%) and in the concentration of employment at large plants (21 percentage points), while construction has a smaller labor productivity gap (14%) and virtually the same concentration in East and West.

What is more, the lack of large plants, lower productivity, and lower wages are systematically related to differences in size-wage curves. On average, the plant size elasticity of wages is one fifth larger 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 and particularly low average wages. In turn, the steeper size-wage curves plants face in the East can be traced back to differences in collective bargaining: Workers at small plants in East Germany 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 aggregate 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. Plants have the following three-stage decision problem. First, plants decide about market entry. Second, after market entry, they choose how many customers to acquire, trading off additional sales and marketing expenses. This customer base choice also takes into account the labor needed to supply additional customers and, thus, that a larger customer base drives up wages in line with the upward-sloping size-wage curve present in the data. Third and finally, plants decide about prices charged to each individual customer, taking into account their product market power.

In such a model, the described trade-offs adversely affect aggregate productiv-

ity through two channels. First, monopsony power works through a labor allocation channel: It compresses the employment distribution across the heterogeneous plants, and reallocates labor from more to less productive plants. Second, monopsony power works through a network size channel: Plants spend less on customer acquisition leading to less efficient production networks, as the average variety-loving customer bundles from fewer plants. Indeed, the data supports a close relationship between monopsony power and marketing expenses: Marketing expenses are particularly small in those industries in East Germany which have a particularly steeper size-wage curve compared to their West German counterparts.

We calibrate the model to the average plant size and the share of large plants in West Germany. Imposing the steeper size-wage curve from East Germany as a menu to choose from for the plants in our model generates a ten percentage points lower productivity. The network size channel explains half of this number. 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 in that region. 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.

The remainder of the paper is organized as follows: First, we review the literature. Then, Section 2 discusses our data sets. 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.¹ What is more, the aggregate total factor productivity differences are unlikely the result of a higher degree of labor market flexibility in West Germany, nor of differences in industry composition. Also, as Appendix B shows, 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.

¹Hence, 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 our 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 \(2016\)](#). Our paper highlights a new force behind productivity losses from a compressed plant size distribution: steeper size-wage trade-offs. To study this force, the German case is particularly interesting. Government policies (and their enforcement) are essentially uniform across regions but there are East-West German differences in labor market power related to the historically determined concentration of collective bargaining at large plants in East Germany. This selectively increases the steepness of the size-wage curve there.

As we have argued before, steeper size wage trade-offs result in a form of monopsony power that plants have when choosing their scale of operations. Recently, [Berger, Herkenhoff, and Mongey \(2022\)](#) have also highlighted monopsony power as a source of size distortions. Their focus is on the distortions of employment decisions given a plant's business model, while ours is on the distortions affecting the long term choice of the business model itself.² Consequently, they use fluctuations in corporate taxes as shifters of labor demand to identify monopsony power. In our case, higher wages at larger plants do not arise directly from an increased labor demand but indirectly from an increased likelihood of collective bargaining. We view both perspectives on monopsony power as complementary.

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-

²In addition to this more conceptual difference relative to [Berger, Herkenhoff, and Mongey \(2022\)](#), we focus on monopsonistic, as opposed to oligopsonistic, competition. What is more, we restrict the analysis to allocative effects, abstracting from normative efficiency questions.

convergence within the same legal framework.³ Our focus is, thus, different from those earlier studies on the 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 attention in the literature. [Becker, Mergele, and Woessmann \(2020\)](#) and [Sleifer \(2006\)](#) show that East Germany has been only nine percent poorer before World War II. Today the discrepancies are much larger. We explain 40% of today's productivity differences between the two regions or two-thirds of the post World War II increase. [Snower and Merkl \(2006\)](#) study unemployment differences between East and West Germany and relate them to government transfers. Regarding convergence in labor productivity, [Burda \(2006\)](#) emphasizes the role of capital accumulation frictions for the slow convergence 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. [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-border 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.

³Non-convergence can also be found in other countries (Italy's "Mezzogiorno", the US' "Rust-belt", 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 \(2006\)](#).

Lastly, in terms of model ingredients, our paper marries two literatures. We start by drawing from the large literature on monopsony power in the labor market (Lamadon, Mogstad, and Setzler, 2022; Berger, Herkenhoff, and Mongey, 2022; Jäger, Roth, Roussille, and Schoefer, 2021; Card, Cardoso, Heining, and Kline, 2018; Manning, 2011, 2003; Burdett and Mortensen, 1998). We, by contrast, highlight that monopsony power distorts also long-run strategy decisions, namely investments into customer acquisition. Customer acquisition, in addition to differences in technical productivities, is another force the literature has highlighted to explain the size distribution of plants (see Einav, Klenow, Levin, and Murciano-Goroff, 2021; Sedláček and Sterk, 2017; Gourio and Rudanko, 2014; Drozd and Nosal, 2012; Arkolakis, 2010). We show that, combined with a love-of-variety-in-production argument (see, e.g., Bilbiie, Ghironi, and Melitz, 2012), less customer acquisition leads to lower aggregate labor productivity in a framework with monopsony power in the labor market.

2 Data

For our analysis, we use administrative aggregate, industry-level, and micro data at the regional level. 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 regional level.⁴ The micro data sets are, respectively, the German 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

⁴The published regional national account data is only available at the supra-industry level. We thank Dr. Thalheimer from the statistical office of Baden-Württemberg for making data at the industry level available to us.

every four years beginning in 2006. 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, earnings, and contract types. 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. Specifically, the *SES* divides Germany up into 5 regions.⁵ The sample is representative for the universe of all German plants with at least ten employees.⁶ Self-employed workers are not covered.

For our analysis, we employ the 2006, 2010, and 2014 samples, which we pool for most empirical analyses. 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 for our baseline analysis and provide a robustness check including part-time workers. 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 a consistent classification. Table C1 in the Appendix C provides a crosswalk for this merger and shows how it relates to the industries from the national accounts.

The *SES* provides the best available data source for our analysis. First, 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 pay and hours worked of employees. Second, it also provides detailed information on workers' sex, age, education, occupation, tenure, and job levels. Third, the survey has information on about 3.2 million employees from roughly 28,700 plants in 2006, 1.9 million employees from 32,200 plants in 2010, and 0.9 million employees from 35,800 plants in 2014.⁷

⁵North: 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.

⁶This restriction is meant to reduce the administrative burden on small enterprises.

⁷The 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) and IAB Establishment Panel (IAB EP)

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 and which contains daily earnings, not wages, up to the social security cap. The data covers the universe of German plants and is available for both West and East Germany from 1993 until 2014 (see [Stüber and Seth, 2018](#); [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 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.⁹ We use the *AWFP* on occasion because it covers a longer time period than the *SES* and provides supplementary information about plants, but its wage data are inferior to the *SES*.

The *IAB EP* provides additional information for a subset of plants in the *AWFP* (see [Ellguth, Kohaut, and Möller, 2014](#)). The IAB collects this information by means of a survey that it sends yearly to up to 15,500 plants. For our purposes, we use the information on collective bargaining agreements at the plant level contained in the *IAB EP*.

⁸Marginal 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, and the self-employed. East German employees were integrated with the West German social security administration only after 1992.

⁹To 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 Empirical Analysis

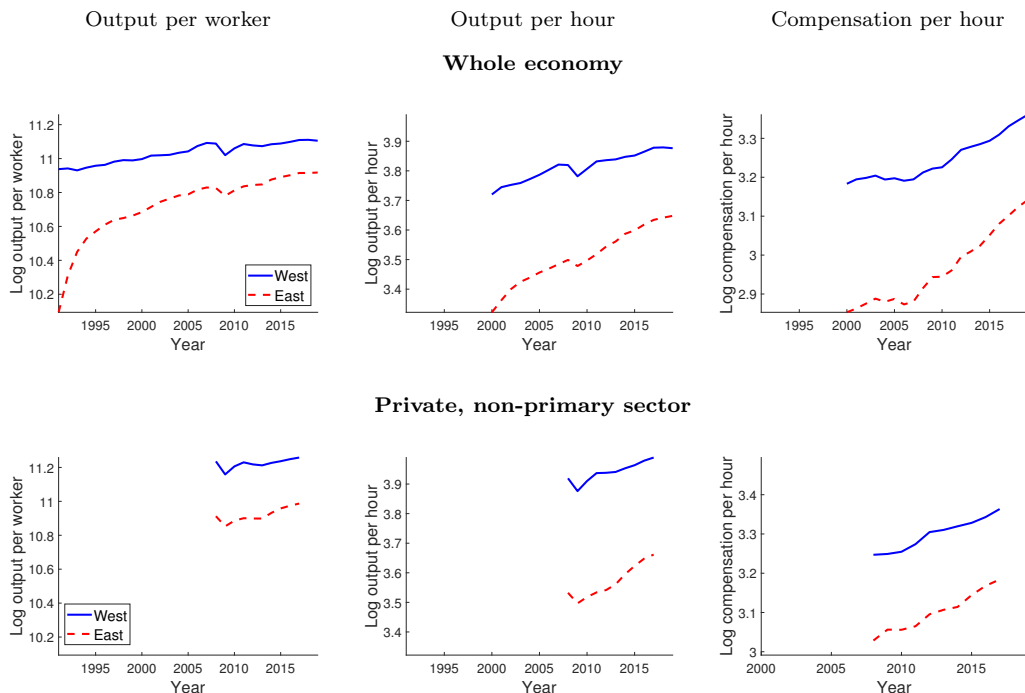
We start this section by documenting that, at an aggregate level, East Germany has lower aggregate labor productivity and labor compensation, whether one includes the public and primary sectors or not. The *SES* data allows us to establish that the lower labor productivity in East Germany is related to missing large plants in the East, which itself is related to a steeper size-wage relationship there, which, finally, is related to regional differences in collective bargaining coverage.

3.1 Aggregate Productivity

In 1990, when centrally planned East Germany reunited with West Germany and became a market economy, a broad range of factors played an important role in depressing labor productivity: Capital was in short supply, machines were outdated, political pressure had plants over-employ labor in the East, and business customer networks evaporated. 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 in East Germany. However, this process of fast growth ended around 1995. Since then, convergence in relative labor productivity and wages has almost come to a halt and the difference remains currently at 18%.¹⁰ What is more, as the bottom panel of Figure 1 shows, the East-West productivity difference remains with 25% even larger in the private (non-primary) sector. Finally, the rightmost panels show a similar magnitude for East-West differences in real wages. This fact, that wage differences mirror productivity differences, also makes the following explanation based on mere accounting unlikely: headquarters of most large firms are located in West Germany, and, hence, the income from unlocalized intangible capital is accounted for there. Given that we measure productivity as value added productivity, this type of accounting would increase measured West German productivity. Yet, it would leave wages unaffected across the two regions. Therefore, without other underlying localized productivity differences, wages across the two regions should be the same.

¹⁰We 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

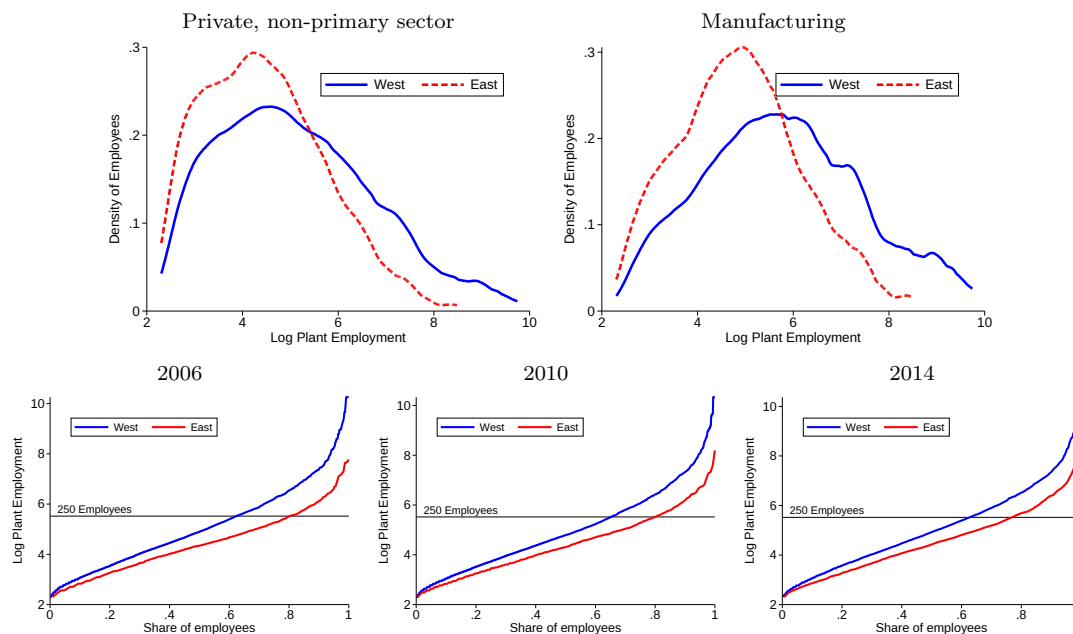


Notes: The figure displays yearly log real output per worker, yearly log real output per hour, and yearly log real labor compensation per hour in East and West Germany. Output is measured as gross value added, which is the GDP concept available at the regional level, because product-specific subsidies and taxes (the difference between the two) are only available at the national level. 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) from 1992 to 2017. 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.

3.2 Missing Large Plants in East Germany

That East Germany has fewer large plants than West Germany in the private, non-primary sector 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 sample years, employment is more concentrated at large plants in the West. We follow the German statistical agency

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

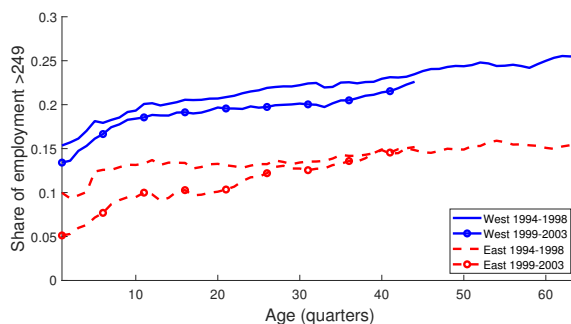


Notes: The figure displays employment-weighted plant size distributions for East and West Germany. The top panels display, respectively, an estimated density function (by a Gaussian kernel smoother) in the private, non-primary sector and in the manufacturing sector. We pool the 2006, 2010, and 2014 samples. The bottom panels display, for different survey years, what fraction of employees is employed at plants up to a certain size as measured by plant log-employment. Data source: *SES* 2006/10/14.

and define plants with more than 249 employees as large. In the West, 39% of employees were 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, we show that this difference in plant size extends back into the 1990s 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 and 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 appear to choose technologies or products that imply a relatively small plant size. 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.

Figure 3: Employment share 250+ by cohort

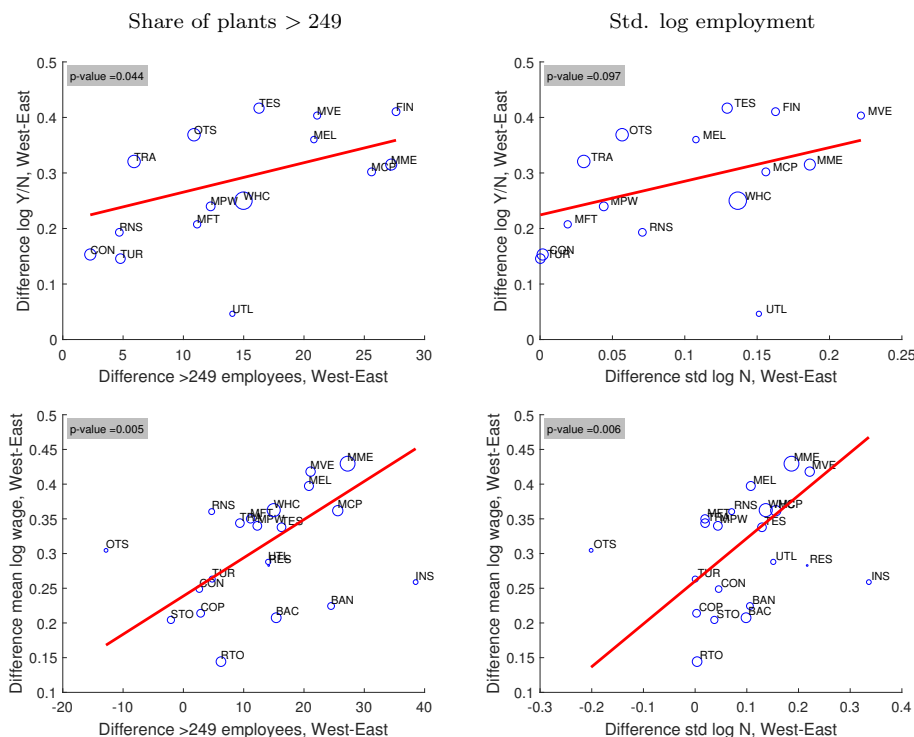


Notes: The figure displays, for the private non-primary sector, for different plant-entry cohorts the share of employment at plants with more than 249 employees over their life-cycles. Data source: *AWFP*.

Returning to Figure 2 and comparing its two top 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 explores the cross-sectional heterogeneity in plant size distributions more systematically at the industry level and relates it to East-West differences in productivity and wages. The left panels use the share of employment at plants with more than 249 employees to compare plant size distributions. The right panels use the standard deviation of log-employment instead. The employment-weighted correlation between productivity differences and plant size distribution differences (top-row) is 0.53 for the 249-share and 0.44 for the standard deviation. Both scatter plots show that those industries where productivity is particularly low in the East are also the industries where particularly fewer workers are employed at large plants in East Germany relative to West Germany.

Relating the size distribution to output per worker has the drawback that it confounds labor share and marginal labor productivity differences across the two regions within industry. To alleviate this concern, the bottom row of Figure 4 relates the plant size distribution to within-industry differences in average log wages across the two regions. Similar to output per worker, we find that those industries where wages are particularly low in the East are also the industries where particularly fewer workers are employed at large plants in East Germany

Figure 4: Productivity and wage differences and large plants by industry



Notes: The top panels relate 2014 log differences in output per worker between West and East Germany within industries to the share of employment at plants with more than 249 employees (left panels) and the standard deviation of log plant employment (right panels). Output is measured as gross value added, which is the GDP concept available at the regional level, because product-specific subsidies and taxes (the difference between the two) are only available at the national level. The lines show (VGR) employment-weighted least squares regressions. The bottom panels relate differences in mean log wages between West and East Germany within industries to the same plant size measures. The lines show (SES) employment-weighted least squares 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, *FIN*: Finance. See Appendix C for the mapping of industries between the *SES* and *VGR*. Data sources: *SES* 2006/10/14 (plant sizes, wages) and *VGR* (labor productivity).

relative to West Germany. The correlations are 0.59 (249-share) and 0.57 (standard deviation), respectively.¹¹

¹¹An additional advantage of using wages is that both the size distribution and wage measures come from the same data source (*SES*) with the same sampling procedures.

3.3 Size-Wage Nexus and Missing Large Plants

These differences in the plant size distribution are, in turn, related to differences in the size-wage curves 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 variable equal to 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 plant-size-related 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, a 1% higher employment is associated with a 0.078% higher wage. The corresponding number for the East is 0.094%, one fifth higher. For example, in West Germany, a business model with 100 employees has to pay 5.6% higher wages than a business model with 50 employees (log difference 0.69). In the East Germany, the same difference in business models comes with 6.7% higher wages. Appendix D.1 shows that the result is robust to including non-linear size terms, which might otherwise drive differences in the average size-wage gradient given the differences in the plant size distributions.¹²

Another concern may be that the steeper size-wage relationship in the East reflects large plants in East Germany attracting a larger share of high-ability workers.

¹²The appendix also extends the analysis to include part-time workers and shows that this, if anything, increases East-West differences in the size-wage nexus. We also estimate a more flexible regression that allows for East/West-specific effects of industry and worker characteristics. This controls for potential East/West-differences in sorting and East/West-specific industry-level demand shocks. Again, we find that the differences in the size-wage elasticities become even a little larger than in our baseline specification.

Table 1: Size-wage elasticities

	Non-primary private sector		
	Baseline	Occupation \times Education	Job level \times Education
Size-Wage 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 \times Education	Job level \times Education
Size-Wage 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	
Size-Wage 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 elasticities 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 \times 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 \times Education:* Controls for a workers' age, sex, education, and job level (five levels, 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* 2006/10/14.

However, 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.¹³ In the next section, we show that the productivity difference rationalized by our model is increasing in the size-wage premium difference between the two regions. To be conservative, we, therefore, choose the regression with the fewest controls as our baseline.

The second panel of Table 1 shows that the difference in the size-wage curve between East and West Germany is even more pronounced in the manufacturing sector. The fact that the East-West difference in the size-wage nexus is not uniform across industries generalizes. Importantly, it is also systematically related to industry variation in average wages and the prevalence of large plants, as Figure 5 shows. Industries with a particularly steep size-wage nexus in East Germany are those industries with many missing large plants and lagging wages in the East.

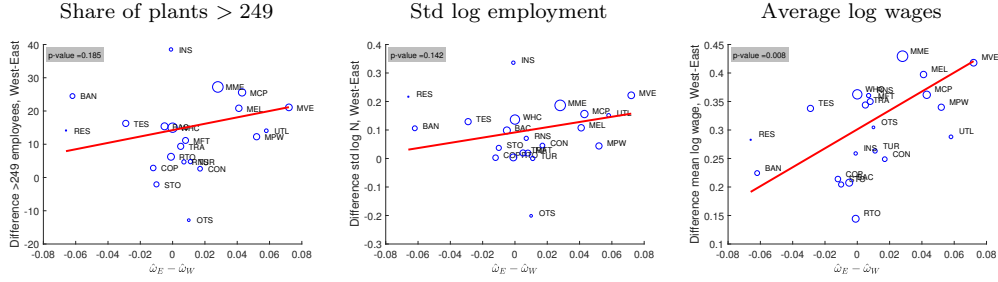
To show this, we estimate Equation (1) for 21 individual industries. In Figure 5, we plot the difference $\hat{\omega}_E - \hat{\omega}_W$ against (a) the difference in the share of employment at large plants, (b) the difference in the standard deviation of log employment, and (c) the difference in the average log wage for each industry. We find that the steeper the size-wage curve is in the East relative to the West, the smaller is the relative share of employment at large plants (employment-weighted correlation of 0.30). The employment-weighted correlation for the standard deviation of log plant employment is 0.33. The correlation between average wages and the size-wage nexus is with 0.56 even stronger. The steeper the size-wage curve is in an East German industry relative its West German “twin”, the more are East wages lagging behind. 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 curves? It could be that East Germans have more specific workplace preferences, leading to lower degrees of substitutability between employers. Instead, we highlight the role of collective wage bargaining and the differences in the role of unions rooted in the

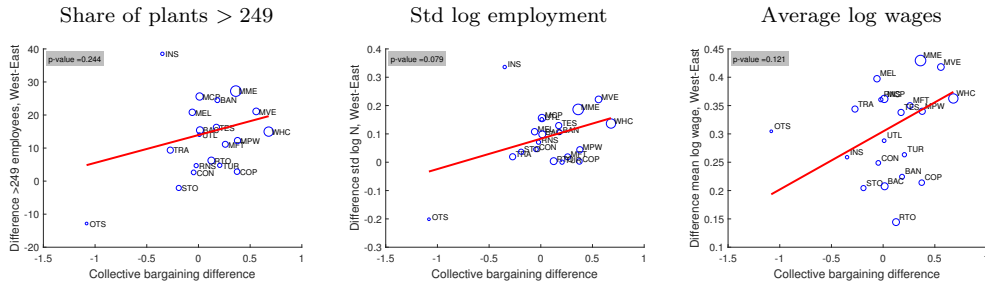
¹³In Appendix D.2 we investigate the issue of selection further by using the social security data which allow us, with the caveat that these are top-coded earnings as opposed to hourly wage data, to use estimates of plant-level fixed effects controlling for worker fixed effects. We find the same pattern of a steeper East German size-wage curve.

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

Nexus between $\hat{\omega}_E - \hat{\omega}_W$ and ...



Nexus between differences in the prevalence of collective bargaining and ...



Notes: The top panel relates differences between West and East Germany in the share of employment at large plants, the standard deviation of log plant employment, and average log wages to differences in size-wage relationships. The bottom panel relates differences between West and East Germany in the share of employment at large plants, the standard deviation of log plant employment, and average log wages to the following double difference: $[\log P(C|L, E) - \log P(C|S, E)] - [\log P(C|L, W) - \log P(C|S, W)]$, 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 (≤ 249 employees) plants in the (E)ast and (W)est. The lines show employment-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. Data source: *SES* 2006/10/14.

different historical developments before 1990. We find that, once we condition on whether individual employment contracts are subject to collective bargaining, the size-wage curve in East and West Germany is basically identical (see the bottom panel of Table 1).¹⁴ At the same time, collectively bargained wages are higher throughout the plant size distribution, and collective bargaining is most prevalent at large plants. When not conditioning on the type of bargaining agreement, the size-wage curve is, therefore, steeper than both conditional size-wage curves. In other words, when one compares large to small plants, wages are higher not only because large plants pay more conditional on the bargaining arrangement, but also because large plants are more likely to be covered by collective bargaining. This composition effect is stronger in East Germany. Small East German plants are relatively unlikely to be covered by a collective bargaining agreement compared to their West German counterparts.¹⁵ Expressed differently, union effort for collective bargaining is, in East Germany, more selectively focused on large plants.¹⁶ The first-mentioned preference-based explanation appears to be difficult to reconcile with this pattern in Germany.

To substantiate this interpretation, the bottom panels of Figure 5 show on the x-axes, for each industry, a double difference in the (log) prevalence of collectively bargained wage contracts between large and small plants and between East and West. That is, it shows the East-West-difference in the elasticity of bargaining prevalence with respect to size. For the majority of industries, this double difference is positive. This means that the fraction of collectively bargained wage contracts increases indeed more in plant size in East than it does in West Germany.

¹⁴The size-wage curve is flatter for collectively bargained wages (e.g., 5.8 vs. 7.7 in the West). This means that collective bargaining raises, in particular, wages at small plants. The literature has documented this also for the U.S. and the U.K. (see [Stewart, 1987](#); [Brown and Medoff, 1989](#); [Blanchflower, Oswald, and Garrett, 1990](#); [Green, Machin, and Manning, 1996](#)). That the size-wage curve for collectively bargained wages is not completely flat has at least two reasons: First, in Germany, unions can negotiate firm-specific wage agreements that then hold for the entire workforce of that firm. 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.

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

¹⁶In Germany, for a plant to be covered by collective bargaining, the employer needs to agree to join an employer association. Workers can, however, pressure employers to do so by striking (see [Jäger, Noy, and Schoefer, 2022](#)). It is natural that unions concentrate such costly efforts on large employers.

This double difference is then 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 (center panel). The relationships between collective-bargaining prevalence differences and our measures of missing large plants is positive with an employment-weighted correlation of 0.27 and 0.40, 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 West Germany, large plants are particularly missing in the East. Finally, the right panel relates the differences in collective bargaining to wage differences across industries. 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 West Germany, wages are particularly low in East relative to West Germany (correlation: 0.36). We view these differences in collective bargaining prevalence as arising from historical developments. In the former socialist East Germany, union membership was high because non-membership was associated with economic and social disadvantages (see [Hans-Böckler-Stiftung, 2022](#)). As a result, unions were not viewed as part of civil society, and union membership fell quickly after reunification. This union retrenchment was particularly pronounced at small plants, leaving collective bargaining concentrated at large plants.

In summary, the data suggest that plants in East Germany face a stronger trade-off between being large and paying low wages. This stronger trade-off appears to originate from the larger concentration of collective bargaining at large plants in East Germany. Most importantly, across industries, the stronger size-wage trade-off in the East correlates with missing large plants and plants paying on average low wages.

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 model where plants chose their optimal scale. The recent literature emphasizes two forces to explain heterogeneous plant scales: productiv-

ity and customer accumulation. For the former, we follow much of the literature, originating in [Hopenhayn \(1992\)](#), that treats technical productivity differences as exogenous. For the latter, we draw on a recent and growing literature that puts some form of customer accumulation at the center stage in addition to productivity differences (see [Einav, Klenow, Levin, and Murciano-Goroff, 2021](#); [Sedláček and Sterk, 2017](#); [Gourio and Rudanko, 2014](#); [Drozd and Nosal, 2012](#); [Arkolakis, 2010](#)). In these models, in order to grow, plants have to make potential customers aware of their products through marketing. Figure 6 shows that marketing expenditures are, indeed, higher in most West German industries compared to their East German counterparts. What is more, the figure shows that these 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 (employment-weighted correlation of 0.71).¹⁷

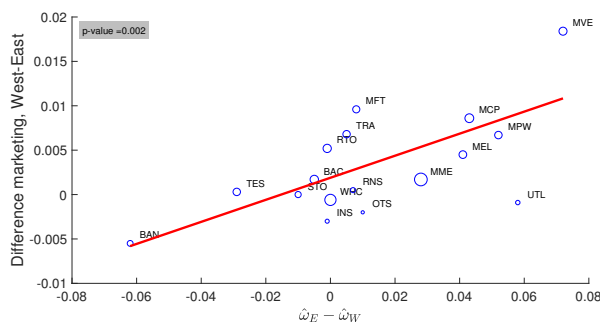
Our model will highlight the following mechanism behind this relationship: plants that face a steeper size-wage trade-off chose business models consistent with a relatively small plant size, economizing on wages. Accordingly, plants in East Germany chose a small business model because, in expectation, they, thus, avoid paying high collectively bargained wages. Such business models also require smaller marketing expenditures. To reiterate, this choice of business model constitutes a form of labor market power.

Concretely, we introduce a size-wage trade-off into the following framework: There are intermediate good producers with heterogeneous productivities using 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 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, perfectly competitive bundlers produce a perfectly substitutable consumption good.¹⁸ Finally, given

¹⁷Data 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. The industry coverage is slightly different to the SES.

¹⁸We emphasize the interaction of customer accumulation and labor market power in shap-

Figure 6: Marketing expenditures and the size-wage nexus



Notes: The figure relates differences between West and East Germany in the ratio of marketing expenditures relative to sales (West minus East) to differences in the size-wage nexus for those industries where marketing expenses are available. The line shows an employment-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 for the mapping of industries to sectors. Data sources: *SES* 2006/10/14 and the *Mannheimer Innovationspanel* 2014.

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 . All bundlers produces a final consumption good, Y_j , using a Dixit-Stiglitz aggregator:

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

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

ing plant size and productivity, and, therefore, we abstract, for tractability reasons, from how interregional trade additionally influences this nexus. We thus model East and West Germany as closed economies each, 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.

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

$$y_{ij} = \gamma_i^\eta \theta_{ij}^\eta \left(\frac{p_{ij}}{\bar{P}_j} \right)^{-\eta} Y_j = \begin{cases} \left(\frac{p_{ij}}{\bar{P}_j} \right)^{-\eta} Y_j & \text{if } \gamma_i = \theta_{ij} = 1, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

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 , the ideal price index, is given by

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

which can be written as

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

because we assume that prices and γ and θ are independent. The latter reflects random matching between intermediate good producers and bundlers, 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 $[\int (\gamma_i \theta_{ij})^\eta di]^{\frac{1}{1-\eta}}$ does not depend on the specific bundler j :

$$\bar{P}_j = (\Gamma \bar{\Theta})^{\frac{1}{1-\eta}} \hat{P}_j, \quad (6)$$

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

$$\hat{P}_j = \left(\int p_{ij}^{1-\eta} di \right)^{\frac{1}{1-\eta}} \quad (7)$$

is the average price charged by intermediate good producers. We focus on the symmetric equilibrium in which $\hat{P}_j = \hat{P}$ and $Y_j = Y$. From this follows, $\bar{P}_j = \bar{P}$.

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 producer i 's idiosyncratic technical productivity (in terms of intermediate goods). As will become clear later, for tractability reasons, we assume that plants learn their productivity after entry. 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 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 mass of bundlers known to that producer.

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

$$w_i = \left(\frac{l_i \Theta_i}{\bar{l} \bar{\Theta}} \right)^{\hat{\omega}} W, \quad (8)$$

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 (8) could be derived from preferences for specific workplaces (see e.g. Berger, Herkenhoff, and Mongey, 2022) or imperfect information about outside options (see e.g. Jäger, Roth, Roussille, and Schoefer, 2021), 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, at least as far as the East-West *differences* are concerned. In turn, these institutional differences are arguably driven by history 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. Then, 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 .¹⁹

$$p_i = \frac{\eta}{\eta - 1} \frac{w_i}{z_i}. \quad (9)$$

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

$$\pi_i(\Theta_i) = \Theta_i \left(p_i y_i - y_i \frac{w_i}{z_i} \right) = \Theta_i \left(y_i \frac{1}{\eta - 1} \frac{w_i}{z_i} \right), \quad (10)$$

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, (3), as well as the optimal price, (9), allows us to express gross profits as a function of known bundlers and marginal costs:

$$\pi(\Theta_i) = \Theta_i \left(\frac{w_i}{z_i} \right)^{1-\eta} \left(\bar{P} \frac{\eta - 1}{\eta} \right)^\eta \frac{Y}{\eta - 1}. \quad (11)$$

4.2.2 Optimal Marketing

The intermediate good producer maximizes gross profits net of marketing costs but takes into account wages as a function of the total number of employees. Therefore, we first need to express wages in (11) as a function of the mass of bundlers known to the producer. To this end, we plug the number of workers, $\frac{y_i}{z_i}$, required to fulfill

¹⁹The intermediate good producers' price-setting can ignore the fact that they are in monopolistic 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.

the demand from each individual bundler, (3), into the size-wage trade-off, (8):

$$w_i = \left(\frac{(p_i)^{-\eta} Y \Theta_i}{z_i \bar{l} \bar{\Theta}} \right)^{\hat{\omega}}. \quad (12)$$

Next, substituting p_i with the optimal pricing decision (9), solving for the wage w_i , and summarizing terms, we obtain wages as a function of the mass of known bundlers as well as productivity and aggregates:

$$w_i = z_i^{\frac{(\eta-1)\hat{\omega}}{1+\eta\hat{\omega}}} \bar{w} \left(\frac{\Theta_i}{\bar{\Theta}} \right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}}, \quad (13)$$

where $\bar{w} = \left[\left(\bar{P} \frac{\eta-1}{\eta} \right)^\eta \frac{Y}{\bar{l}} \right]^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}}$ 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:

$$\Pi_i = \pi(\Theta_i) - \mu \bar{P} \Theta_i. \quad (14)$$

Substituting in gross profits, (11), and the size-wage trade-off, (13), yields

$$\Pi_i = \Theta_i \left(\frac{\Theta_i}{\bar{\Theta}} \right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}(1-\eta)} z_i^{-(1-\eta)\frac{\hat{\omega}+1}{1+\eta\hat{\omega}}} \bar{P}^\eta Y \frac{(\eta-1)^{\eta-1}}{\eta^\eta} \bar{w}^{1-\eta} - \mu \bar{P} \Theta_i. \quad (15)$$

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

$$\frac{1+\hat{\omega}}{1+\eta\hat{\omega}} \frac{Y}{\mu} \frac{1}{\eta} \left(\frac{\bar{P}}{\bar{w}} \frac{\eta-1}{\eta} \right)^{\eta-1} \left(\frac{\Theta_i}{\bar{\Theta}} \right)^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}(1-\eta)} = z_i^{(1-\eta)\frac{1+\hat{\omega}}{1+\eta\hat{\omega}}}, \quad (16)$$

which, solving for Θ_i , simplifies to

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

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 (13) and express (17) in terms of the real wage targeted by a producer:

$$\frac{w_i}{\bar{P}} = \frac{p_i}{\bar{P}} \frac{\eta - 1}{\eta} z_i = \left[\frac{1 + \hat{\omega}}{1 + \eta \hat{\omega}} \frac{Y}{\mu} \frac{1}{\eta} \right]^{\frac{1}{\eta-1}} \frac{\eta - 1}{\eta} z_i. \quad (18)$$

The real wage is proportional to idiosyncratic technical productivity, z_i , which also implies that marginal costs are constant across producers. Resulting from producers' product market power, workers do not receive the full marginal product of labor. Instead, they get a wage equal to the inverse mark-up in the product market, $\frac{\eta-1}{\eta}$, of the producer's technical productivity shifted by the term in squared brackets. This term reflects the efficiency of the producer's network. The price of that producer in terms of final goods, $\frac{p_i}{\bar{P}}$, reflects the marginal value of the producer's output across bundlers. The producer chooses a larger and hence more efficient network, if demand Y divided by the inverse profit margin (in goods) η and the cost of serving one more market μ is higher. In addition, in this choice, producers take into account the effect of their size on their wages and hence their operating profits. This effect becomes stronger when the size-wage trade-off becomes steeper as captured by the elasticity $\frac{1+\hat{\omega}}{1+\eta\hat{\omega}} < 1$, decreasing the network size.

Constant marginal costs across producers, (18), imply, using (9), that all producers charge the same price and sell the same quantity to each bundler they know. In turn, we obtain for the price, using (6),

$$\frac{p_i}{\bar{P}} = \frac{\hat{P}}{\bar{P}} = (\Gamma \bar{\Theta})^{\frac{1}{\eta-1}}, \quad (19)$$

and, for the quantity sold to a single bundler, using (3):

$$l_i z_i = y_i = \left(\frac{p_i}{\bar{P}} \right)^{-\eta} Y = Y (\Gamma \bar{\Theta})^{\frac{\eta}{1-\eta}}. \quad (20)$$

Finally, using (9), the real wage is given by:

$$\frac{w_i}{\bar{P}} = \frac{\eta - 1}{\eta} (\Gamma \bar{\Theta})^{\frac{1}{\eta-1}} z_i. \quad (21)$$

Comparing (21) to (18), we obtain:

$$\bar{\Theta} = \frac{Y/\Gamma}{\mu} \frac{1 + \hat{\omega}}{\eta + \eta \hat{\omega}}. \quad (22)$$

Importantly, the average network size depends negatively on the size-wage elasticity, $\hat{\omega}$, as higher monopsony power discourages customer accumulation, in line with the data (see Figure 6). It depends positively on the market size per producer gained by one unit of marketing costs, $\frac{Y/\Gamma}{\mu}$.

To derive a closed-form solution for the distribution of optimal marketing choices, we need to make a functional form assumption about the distribution of idiosyncratic productivity, z_i . We assume that z_i is log-normally distributed, $z_i \sim LN(\ln \bar{z}, \Sigma^2)$.²⁰ For any log-normally distributed random variable $z \sim LN(\ln \bar{z}, \Sigma^2)$ and real number x , it holds that:

$$E(z^x) = \bar{z}^x \phi^{x^2}, \quad \text{with } \phi = \exp(0.5\Sigma^2). \quad (23)$$

This implies, taking expectations over (17) and using that z_i is assumed to be observed after entry:

$$\left[\frac{1 + \hat{\omega}}{1 + \eta \hat{\omega}} \frac{Y}{\mu} \frac{1}{\eta} \left(\frac{\bar{P}}{\bar{w}} \frac{\eta - 1}{\eta} \right)^{\eta-1} \right]^{-\frac{1+\eta\hat{\omega}}{\hat{\omega}(\eta-1)}} = E \left(z_i^{\frac{1+\hat{\omega}}{\hat{\omega}}} \right) = \bar{z}^{\frac{1+\hat{\omega}}{\hat{\omega}}} \phi^{\left(\frac{1+\hat{\omega}}{\hat{\omega}}\right)^2}, \quad (24)$$

because $\bar{\Theta}$ is the expected value of Θ_i .

Dividing (17) by (24) allows us to express individual marketing choices in a compact form:

$$\frac{\Theta_i}{\bar{\Theta}} = \left(\frac{z_i}{\bar{z}} \phi^{-\frac{1}{\hat{\omega}}} \right)^{\frac{1+\hat{\omega}}{\hat{\omega}}}. \quad (25)$$

²⁰Strictly speaking, we approximate the solution, ignoring 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 (16) produces a $\Theta_i > 1$. However, in our calibration, that fraction is negligible.

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 (25) that $\log \Theta_i$ is normally distributed, too, and has a larger variance than $\log z_i$ because $\frac{1+\hat{\omega}}{\hat{\omega}} > 1$. This means that the distribution of networks, the distribution of Θ_i , is more right skewed than the productivity distribution: The most productive producers build particularly large networks. The endogenous customer acquisition decision amplifies, therefore, productivity heterogeneity. This effect becomes smaller as $\hat{\omega}$ increases: A stronger size-wage trade-off renders the acquisition of additional customers less attractive because wages rise too fast.

Before we turn to the final producer decision, namely, market entry, we point out two properties of the optimal producer size. Combining (20) and (25), producer size is given by:

$$l_i \Theta_i = z_i^{1/\hat{\omega}} Y (\Gamma \bar{\Theta})^{\eta/(1-\eta)} \bar{\Theta} \left(\frac{1}{\bar{z}\phi} \phi^{-\frac{1}{\hat{\omega}}} \right)^{\frac{1+\hat{\omega}}{\hat{\omega}}}. \quad (26)$$

From this equation follows immediately that producer size is increasing in idiosyncratic productivity. This holds true even though sales per bundler, y_i , are constant, and thus workers per bundler, l_i , decrease in idiosyncratic productivity. More productive producers choose to know more bundlers, and the number of bundlers they choose to know increases more than proportionally in their productivity.

Secondly, from (26), we obtain an explicit solution for the standard deviation of log producer employment:

$$std(\log(l_i \Theta_i)) = std\left(\frac{1}{\hat{\omega}} \log z_i\right) = \frac{1}{\hat{\omega}} \Sigma. \quad (27)$$

That is, the distribution of log producer employment is, similarly to the distribution of networks, normally distributed. Its dispersion depends positively on the standard deviation of idiosyncratic productivity, Σ . Importantly, and consistent with the data in Figure 5, it depends negatively on the size-wage elasticity.²¹

²¹Specifically, we refer to the middle-upper panel in Figure 5. Note that Figure 5 displays positive relationships because the y-axis uses West-East and the x-axis East-West differences.

4.2.3 Producer Entry

We assume free producer entry which implies that competition drives average producer profits to zero. Recall 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, (14), equal entry costs:

$$\int \Theta_i y_i \left(p_i - \frac{w_i}{z_i} \right) di - \int \mu \bar{P} \Theta_i di = \lambda \bar{P}, \quad (28)$$

which implies, using (19)–(21):

$$\lambda = \int \left[\Theta_i Y (\Gamma \bar{\Theta})^{\frac{\eta}{1-\eta}} (\Gamma \bar{\Theta})^{-\frac{1}{1-\eta}} \left(1 - \frac{\eta-1}{\eta} \right) \right] di - \mu \int \Theta_i di. \quad (29)$$

This, integrating out Θ_i , simplifies to

$$\frac{Y}{\Gamma} \frac{1}{\eta} = \lambda + \mu \bar{\Theta}. \quad (30)$$

Equation (30) has an intuitive interpretation: the goods sold per producer, Y/Γ , multiplied by the profit margin per goods sold (in terms of goods), $1/\eta$, equal the sum of market entry costs, λ , and expected marketing costs, $\mu \bar{\Theta}$, (both in goods). 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 size-wage trade-off, the more this decision is tilted towards marketing and, thus, the larger the producers become, in particular the most productive ones (see Equation (25)).

4.3 Equilibrium

In equilibrium, the total amount of employment needs to equal aggregate labor supply. We fix the aggregate labor supply at one unit.²² Hence, labor demand of all active producers, (26), integrated over all producers needs to be one:

$$\Gamma \int \Theta_i l_i di = \Gamma \int z_i^{\frac{1}{\hat{\omega}}} Y (\Gamma \bar{\Theta})^{\eta/(1-\eta)} \bar{\Theta} \left(\frac{1}{\bar{z}\phi} \phi^{-\frac{1}{\hat{\omega}}} \right)^{\frac{1+\hat{\omega}}{\hat{\omega}}} di = 1 \quad (31)$$

which, solving for Y , yields:

$$Y = \underbrace{\bar{z}\phi (\Gamma \bar{\Theta})^{\frac{1}{\eta-1}}}_{=\hat{P}/\bar{P}} \phi^{\frac{2}{\hat{\omega}}}. \quad (32)$$

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 mass of intermediate good producers known to the representative bundler, $\Gamma \bar{\Theta}$. This network size effect is important because of love-of-variety in production. The effect can alternatively be expressed as the ratio of the average price charged to a bundler, \hat{P} , and the ideal 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 the ideal price index. Second, the last term, $\phi^{\frac{2}{\hat{\omega}}}$, is a labor allocation effect similar to the 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 . This complementarity can be exploited better when a low $\hat{\omega}$ allows for a higher concentration of labor at the most productive producers, which build larger customer networks to this end.

Ultimately, Equation (32) together with the average network size, (22), and producer entry, (30), determine the aggregate equilibrium in the economy. Nor-

²²If we analyzed only one geographical unit, for instance, West Germany, this would be an innocuous normalization. However, as we show below, the economy features an aggregate demand externality whereby higher output increases average productivity. Hence, when we calibrate the model separately for East and West Germany and impose a labor force of equal size in both, we disregard the possibility that East Germany is less productive simply because it has a smaller population size.

malizing average producer productivity $\bar{z}\phi$ to one and solving these equations for aggregate output, the average mass of known bundlers, and the share of active producers yields:

$$Y = \left(\frac{1}{\mu\eta} \frac{1 + \hat{\omega}}{1 + \eta\hat{\omega}} \right)^{\frac{1}{\eta-2}} \left(\phi^{\frac{2}{\hat{\omega}}} \right)^{\frac{1}{\eta-2}} \phi^{\frac{2}{\hat{\omega}}}, \quad (33)$$

$$\bar{\Theta} = \frac{\lambda}{\mu} \left[\frac{1}{\eta-1} \left(\frac{1 + \hat{\omega}}{\hat{\omega}} \right) \right], \quad (34)$$

$$\Gamma = \frac{1}{\lambda} \frac{\eta-1}{\eta} \frac{\hat{\omega}}{1 + \eta\hat{\omega}} Y. \quad (35)$$

Equation (33) shows that output is the product of three terms which are all negatively affected by the size-wage trade-off. The last term, $\phi^{\frac{2}{\hat{\omega}}}$, is the aforementioned labor allocation effect on output that would also be present in a pure monopsony model with heterogeneous producers but without endogenous customer accumulation, as we show in Appendix F.²³

Comparing (32) to (33) (and taking into account the normalization of productivity $\bar{z}\phi = 1$) yields a convenient interpretation of the first two terms: they reflect the efficiency of the transformation of intermediate goods into final goods, a love-of-variety effect. This efficiency depends on the network size of bundlers,

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

and is affected by monopsony power because a steeper size-wage trade-off restricts the varieties available to bundlers. The first term, $\left(\frac{1}{\mu\eta} \frac{1 + \hat{\omega}}{1 + \eta\hat{\omega}} \right)$, reflects the fact that all producers reduce their network size because of their monopsony power. This term would also be present in a model without producer heterogeneity, $\phi = 1$. The second term, $\phi^{\frac{2}{\hat{\omega}}}$, reflects the fact that it is particularly harmful that the most productive producers reduce their network size. The fact that both terms enter the average network size with an exponent, $\frac{\eta-1}{\eta-2}$, larger than one, reflects that there is

²³Whether one interprets the impact of $\hat{\omega}$ on the allocation of labor across differently productive producers—through $\phi^{\frac{2}{\hat{\omega}}}$ —as an inefficiency depends on the ultimate source of $\hat{\omega}$. We have discussed some potential sources in Section 4.2. Given the positive focus of this paper, we do not need to take a stance on this question.

a demand externality in the model, which can also be seen in (35): when aggregate demand is high, more producers enter, the average network becomes bigger and the economy more productive. In turn, output increases further and hence also demand.²⁴

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

$$LC = Y - \Gamma(\lambda + \mu\bar{\Theta}) = Y \left[1 - \left(\frac{\eta - 1}{\eta} \frac{\hat{\omega}}{1 + \eta\hat{\omega}} + \frac{1}{\eta} \frac{1 + \hat{\omega}}{1 + \eta\hat{\omega}} \right) \right] = Y \frac{\eta - 1}{\eta}. \quad (36)$$

This means that it is irrelevant whether we compare Y or LC differences across regions in what follows (assuming η is the same).

5 Implications

With the model solution at hand, we can now quantify the aggregate implications of the differences in monopsony power between East and West Germany that we documented in Section 3. Moreover, we can discuss potential policy implications.

5.1 Quantitative Results

We need to determine 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 to employment, $\hat{\omega}$. Our strategy is to calibrate the model to the West German economy given our baseline estimate of $\hat{\omega}_W = 0.078$ from Section 3.3.

Note from Equation (33) that, once we fix $\hat{\omega}$, the key parameters to understand the relative output between two regions are product market power, η , and the dispersion of idiosyncratic productivities, Σ . Bundesbank (2017) finds an average

²⁴This demand externality is one important difference to Kroft, Luo, Mogstad, and Setzler (2020), who discuss the effects of simultaneous labor and product market power in a model in which each producer serves a single product market. They find that product and labor market power dampen each other. The demand externality implies that an increase in product market power, which comes with an increase in love of variety, makes the distortions of the production network size that come from labor market power more detrimental.

price-cost margin of 1.4 in Germany, and, therefore, we set $\eta = 3.5$. We calibrate the dispersion for idiosyncratic productivities to match the share of employment at large plants, that is, one with more than 249 employees. To this end, 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). Recall that the data is truncated at plants with at least ten employees, and we impose the same truncation in the model simulation when computing these moments. However, we compute output in the model following (33), i.e., using the non-truncated producer distribution, when we compare it to national accounts data that is based on the universe of producers.

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 data we use in Figure 6 pins down the marketing cost, 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 the private, non-primary sector in East Germany ($\hat{\omega}_E = 0.094$). The top panel of Table 2 displays the results of this exercise, in the column titled “*Model East.*”

First and importantly, by varying only $\hat{\omega}$, the model matches the moments of the plant size distribution (that were targeted for 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 (36) it follows that the model also rationalizes a ten percentage points lower labor compensation in the East relative to the West.

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* (pooled 2006/10/14) 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.

Section 3.3 shows that East-West size and productivity differences are particularly large in manufacturing. To investigate whether the model is able to match this stylized fact, we next, keeping the general calibration strategy the same, 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 $\eta = 6$.

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

Table 3: Decomposition of output loss

	Private non-primary	Manufacturing
Total productivity difference	10.3%	15.5%
Network size effect, $(\Gamma\bar{\Theta})^{\frac{1}{\eta-1}}$	5.4%	5.6%
sans heterogeneity, $\left(\frac{1}{\mu\eta} \frac{1+\hat{\omega}}{1+\eta\hat{\omega}}\right)^{\frac{1}{\eta-2}}$	1.9%	2.9%
cum heterogeneity, $\phi^{\frac{2}{\hat{\omega}} \frac{1}{\eta-2}}$	3.5%	2.7%
Labor allocation effect, $\phi^{\frac{2}{\hat{\omega}}}$	5.2%	10.6%

Notes: The table displays the output loss per worker in East relative to West Germany, $1 - Y_{east}/Y_{west}$ from Table 2, decomposed, to the first order, into the three channels highlighted in the discussion of Equation (33).

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 in East Germany is 84 percent of output in West Germany, in the data it is 70 percent.

Table 3 decomposes the predicted output losses into the two channels we have highlighted in Equation (33). In the private, non-primary sector, the total output effect is split roughly half into the network size and the labor allocation effect. In manufacturing, the share of the effects is roughly one third and two thirds, respectively.

Of the two terms that constitute the network size effect, in the private non-primary sector, the term arising from heterogeneity is quantitatively larger than the 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 discourages the most productive producers to choose a business model with many customers, rendering the entire production network in the economy less efficient. In the manufacturing sector, the split is more even.

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 originating from a distorted choice in the scale of a plant's operation: 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 of output loss 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 demand 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 policymakers to help distressed regions catch up. This can be seen in Equation (33), where the entry costs, λ , do not appear. In the model, increasing the number of active producers through entry subsidies crowds out network investments of existing producers one-for-one so that the equilibrium production network size, $\Gamma\bar{\Theta}$, remains unaffected, (see Equation 34). 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.

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.²⁵

²⁵We note that this externality aspect of marketing choices carries over to a setup that explicitly motivates monopsony power by a love of variety by households for employers. Appendix H shows that a positive subsidy not only increases output net of cost but also the number of active firms and thus varieties in terms of employers. Thus, the laissez-faire equilibrium in such a setup is inefficient.

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 the average network size, 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.

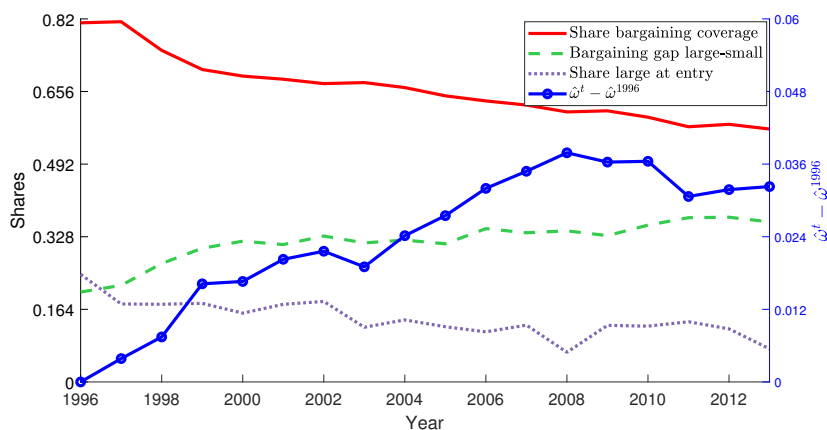
Finally, there is also the question whether minimum wages could mitigate the monopsony power related to the choice of business models which we highlight in this paper. To the extent that minimum wages substitute for low collective bargaining coverage at small plants and, thereby, flatten the size-wage curve, our framework would indeed suggest that they increase aggregate productivity. In particular, minimum wages could, thus, incentivize producers to invest into larger production networks and reallocate workers towards the most productive producers. However, differently from marketing subsidies, they may not reduce the size-wage tradeoff for the most productive producers who pay wages in excess of the minimum wage.

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 particular, the unions' focus on large plants in their efforts to implement collective bargaining steepens the size-wage curves.

Are such changes in the prevalence and selectiveness of collective bargaining also reflected in changes in the steepness of the size-wage curve and plant sizes at entry over time? Figure 7 provides suggestive time series evidence in favor. First, it shows that, in Germany, collective bargaining has substantially declined over

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



Notes: On the left axis, the figure displays for all of Germany, private non-primary sector, over time the share of workers covered by a collective bargaining agreement (share bargaining coverage), the difference in the probability to be covered by collective bargaining for workers at plants with at least 250 employees relative to workers at plants with fewer employees (Bargaining gap), and the share of employment at plants of at least 250 employees in an entering cohort of plants, 4 quarters after entry (share large at entry). On the right axis, it displays the East-West difference in the steepness of the size-wage curve minus its steepness difference in 1996 ($\hat{\omega}^t - \hat{\omega}^{1996}$). Data sources: *AWFP* for $\hat{\omega}^t - \hat{\omega}^{1996}$ and “share large entry” and *IAB Establishment Panel*, for “Share bargaining coverage” and “Bargaining gap”.

time. In 1996, more than 80% of workers were covered by a collective bargaining agreement. This number decreases to less than 60% by 2013. What is more, collective bargaining declined foremost at small plants (for evidence on selective retrenchment in collective bargaining see also [Jäger, Noy, and Schoefer, 2022](#)). In 1996, workers at plants with more than 249 employees had a 20% higher probability to be covered by collective bargaining compared to workers at plants with fewer employees. This gap rose to 36 percentage points by 2013. Second, and in line with what one would expect from our cross-sectional evidence, this selective 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 employees of an entry cohort used to be at large plants in 1996. This share has declined to around 12% by 2013.

Our paper highlights that collective bargaining retrenchment, in general, and selective retrenchment, in particular, decreases aggregate productivity through dis-

torting plants' long run business-model choices. In an influential study, [Brown and Medoff \(1978\)](#) show that, at the industry level in the U.S., high unionization rates are positively associated with labor productivity. Subsequent studies fail to confirm this earlier finding using within-industry, firm-level data (see [Hirsch, 2004](#), for a survey). Our analysis suggests that these results may not be contradictory after all. At an aggregate (industry) level, an increase in unionization of small plants flattens size-wage curves, making it more attractive to choose business models that require large operations. As we have argued, this raises aggregate labor productivity. However, given that the threat of unionization affects business model choices at entry, productivity differences need not manifest themselves when individual unionized and non-unionized firms are compared within industry.

6 Conclusion

Large aggregate labor productivity differences persist across regions where government policies (and legal institutions enforcing these) are almost identical. We consider the case of Germany where, more than two decades after reunification, the East German private, non-primary sector remains about 25% less productive than its West German counterpart. We show that this difference in productivity is closely linked to differences in the size distribution of plants, which are, in turn, related to differences in collective bargaining coverage. In East Germany, collective bargaining is much more concentrated at large plants than it is in West Germany. This selective difference in collective bargaining coverage creates incentives to chose business models in East Germany where the plant stays small to avoid paying high wages, i.e., it effectively creates additional labor market monopsony power in East relative to West Germany. Finally, these East-West differences in monopsony power correlate with average-wage and productivity differences.

We develop a model that merges labor market power and customer acquisition and show that labor market power distorts the size distribution and lowers, thereby, aggregate labor productivity. When firms face steeper size-wage curves and thus have more labor market power, they decide to forgo customer acquisition because otherwise they would require a larger workforce, which raises wages. This leads to long-run business models relying on smaller production networks in general and to

a smaller concentration of workers at the most productive producers in particular. Both channels affect aggregate labor productivity adversely. The model, calibrated to the estimated difference in monopsony power, explains about 40 percent of the observed lower labor productivity in East Germany.

We finally show that neither hiring nor entry subsidies can remedy this productivity loss. A marketing subsidy can mitigate but not fully cure the problem which is ultimately linked to non-uniform collective bargaining retrenchment across plants.

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*.
- BECKER, S. O., L. MERGELE, AND L. WOESSMANN (2020): “The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism,” *Journal of Economic Perspectives*, 34, 143–71.
- BELLMANN, L., B. LOCHNER, S. SETH, AND S. WOLTER (2020): “AKM Effects for German Labour Market Data,” *Institut für Arbeitsmarkt-und Berufsforschung (IAB) Methodenreport*.
- BERGER, D. W., K. F. HERKENHOFF, AND S. MONGEY (2022): “Labor Market Power,” *American Economic Review*, 112, 1147–93.
- 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,” *NBER Working Paper No. 17265*.
- BROWN, C. AND J. MEDOFF (1978): “Trade Unions in the Production Process,” *Journal of Political Economy*, 86, 355–378.
- (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.
- DAUTH, W., S. FINDEISEN, S. LEE, AND T. PORZIO (2022): “Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany,” *mimeo*.
- DIJKSTRA, L., H. POELMAN, AND P. VENERI (2019): “The EU-OECD Definition of a Functional Urban Area,” *OECD Regional Development Working Papers No. 2019/11*.

- 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,” *NBER Working Paper No. 29561*.
- ELLGUTH, P., S. KOHAUT, AND I. MÖLLER (2014): “The IAB Establishment Panel—methodological essentials and data quality,” *Journal for Labour Market Research*, 47, 27–41.
- 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.
- GOMES, P. AND Z. KUEHN (2017): “Human Capital and the Size Distribution of Firms,” *Review of Economic Dynamics*, 26, 164–179.
- 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.

- HANS-BÖCKLER-STIFTUNG (2022): “Zentrale Einheitsgewerkschaft – Die Strukturen des FDGB,” <https://www.gewerkschaftsgeschichte.de/ost-strukturen-des-fdgb.html>, accessed 15. July 2022.
- 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*.
- HIRSCH, B. (2004): “What Do Unions Do for Economic Performance?” *Journal of Labor Research*, 25, 415–456.
- HOPENHAYN, H. A. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 1127–1150.
- 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.
- JÄGER, S., S. NOY, AND B. SCHOEFER (2022): “The German Model of Industrial Relations: A Primer,” *mimeo*.
- JÄGER, S., C. ROTH, N. ROUSSILLE, AND B. SCHOEFER (2021): “Worker Beliefs About Rents and Outside Options,” *Mimeo*.
- JOHNSON, P. AND C. PAPAGEORGIOU (2020): “What Remains of Cross-Country Convergence?” *Journal of Economic Literature*, 58, 129–75.
- KROFT, K., Y. LUO, M. MOGSTAD, AND B. SETZLER (2020): “Imperfect competition and rents in labor and product markets: The case of the construction industry,” *NBER Working Paper No. 27325*.
- LAMADON, T., M. MOGSTAD, AND B. SETZLER (2022): “Imperfect Competition and Rent Sharing in the US Labor Market,” *American Economic Review*, 112, 169–212.
- 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, vol. 4, Part B, 973–1041.
- MERTENS, M. AND S. MÜLLER (2022): “The East-West German Gap in Revenue Productivity: Just a Tale of Output Prices?” *Journal of Comparative Economics*.
- 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.
- SLEIFER, J. (2006): *Planning Ahead and Falling Behind: The East German Economy in Comparison with West Germany 1936-2002*, Jahrbuch für Wirtschaftsgeschichte: Beiheft, Akademie Verlag.
- 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 (2018): “The Administrative Wages and Labor Market Flow Panel,” *FAU Discussion Papers in Economics No. 01/2017 (updated Dec. 2018)*.
- 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 or differences in total factor productivity (TFP). TFP differences, in turn, could result from differences in access to technology or institutions (which is unlikely to be the case 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 East Germany—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 appendix, we establish that, first, differences in factor inputs are unlikely the reason behind the observed differences in output per worker. In other words, it has to be TFP.²⁶ Second, we show that the labor market in East Germany is at least as dynamic as the West German labor market, meaning that labor market sclerosis is not to blame, either.

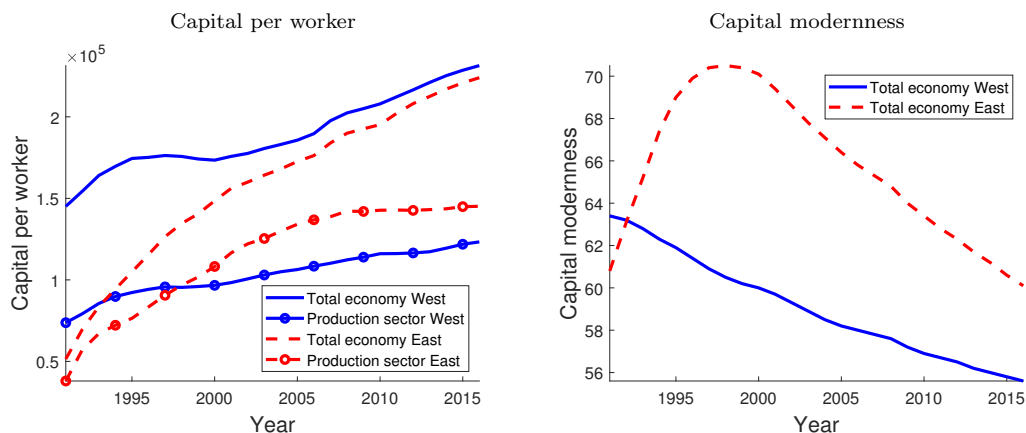
A.1 Capital and Labor Inputs

Burda (2006) puts forward an explanation for low labor productivity where capital accumulation is subject to frictions. East Germany had a lower capital stock in 1992, implying low initial labor productivity, and if it takes time for the East to accumulate capital, this would explain a persistent productivity gap. Figure A1 (left panel) compares the (net, i.e., after depreciation) capital stock per worker in East Germany to that in West Germany. It shows that the capital stock per worker was indeed much lower initially, but, 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

²⁶Mertens and Müller (2022) also argue that the East-West German aggregate output per worker difference in manufacturing can only be explained by TFP differences.

Figure A1: Capital stock



Notes: The left panel displays the net capital stock (after depreciation) per worker in East and West Germany for the total economy and the production sector (manufacturing, mining, utilities, and construction). The right panel displays the modernness of the capital in East and West Germany (net capital divided by gross capital). Data source: *VGR*.

detailed industry and region. It does, however, provide data on the production sector (manufacturing, mining, utilities, and construction), and Figure A1 (left panel) shows that, in that sector, East Germany has even overtaken the West German economy in terms of capital intensity by 1998.

In this comparison, a confounding factor could be capital quality, if East German plants still produced with outdated capital from before the reunification. Figure A1 in the right panel displays the modernness of capital, i.e., net capital divided by gross capital. Consistent with the large catch-up in capital accumulation shown in the left panel, the capital stock is of a rather young vintage in East Germany suggesting that, if anything, it is of higher quality.

Another potential explanation for the lower labor productivity in East Germany could be lower quality of labor inputs. 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. At first inspection, for education and occupation, Table A1 does not suggest that differences in worker quality are a likely explanation. Considering formal education, East Germany tends to have, if anything, a more skilled workforce with fewer workers without formal training. Considering tasks, there is some evidence that workers in West

Table A1: Worker skills

Education shares							
Low		Medium				High	
West	East	West	East	West	East	West	East
11.58	3.66	74.79	83.75	13.63	12.59		
Work task shares							
Low		Medium		Semi-high		High	
West	East	West	East	West	East	West	East
40.09	44.31	42.98	41.99	7.68	5.79	9.24	7.91

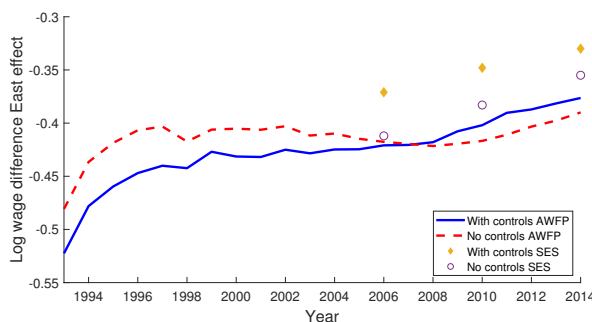
Note: The table displays the shares of education and occupation categories in West and East Germany. Education: Low: Workers without formal vocational training. Medium: Workers with formal vocational training and/or higher education entrance qualification. High: Workers with a university degree. Task: 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. (see Blossfeld, 1987) Data source: *AWFP* 2006–2014.

Germany perform tasks that require somewhat more skills but differences are minor. To analyze this more formally, we estimate in the *SES*, at the worker-level, the following regression for the years 2006, 2010, and 2014:

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

where $East_i$ is a dummy variable equal to one if a worker works at a plant that 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 any observables. This restricted regression simply estimates the mean log-wage differences between East and West Germany for each year. The regression with observables does the same but controlling for different worker skill distributions in East and West Germany. Figure A2 compares the two regressions.

Figure A2: 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 ages, sex, education, and task categories at the plant level.

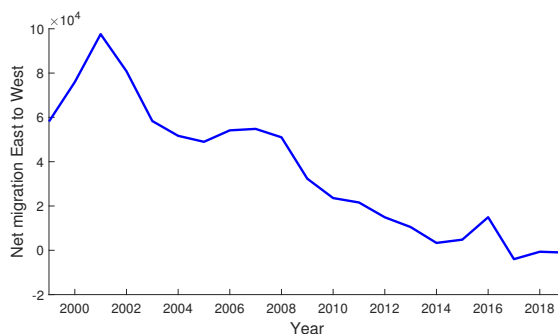
It shows that the mean difference in log wages and the mean difference in log wages after controlling for observable worker characteristics are very similar. Controlling for worker observables explains some of the lower wages in East Germany but, even among observationally identical workers, wages are about 0.35 log points lower in East Germany.

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 average earnings and plant-level shares of worker observables. In addition, the *AWFP* summarizes occupations in four broad groups called work tasks. This leads to the following plant-level regression for each year:

$$\ln w_{jt} = \alpha_0 + East_j + age_{jt} + sex_{jt} + educ_{jt} + task_{jt} + \epsilon_{jt}, \quad (\text{A.2})$$

where $\ln w_{jt}$ is the log average wage at plant j in year t , $East_j$ is a dummy variable equal to 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

Figure A3: 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).

mean and estimate again two versions of the regression, one with the covariates of worker observables and one without it.

Again, as Figure A2 shows, worker observables explain little of the wage differences. In fact, during the early years after reunification, worker characteristics have been somewhat better in East relative to West Germany.²⁷ 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 A3 shows that net-outflows from the East to the West have essentially converged to zero by 2013.

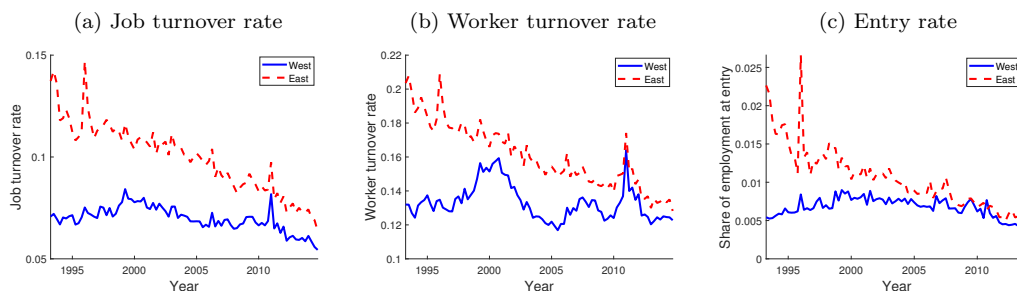
In line with the above, Fuchs-Schündeln and Izem (2012) and Heise and Porzio (2021) find that plant or job characteristics, rather than worker characteristics, explain the bulk of wage differences between East and West Germany even when unobserved worker heterogeneity is controlled for.

A.2 Missing Reallocation

Given that it appears to be neither capital nor the quality of labor that explains productivity differences, the 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,

²⁷We note that a similar quality of the workforce also suggests that East German plants do not remain small because they cannot find high-skilled workers, as in Gomes and Kuehn (2017).

Figure A4: Job and worker turnover rates



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. All three panels: private non-primary sector. Data source: *AWFP*.

unproductive plants towards more productive plants.²⁸ Using the *AWFP*, we show, however, 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 A4 (a) displays the job turnover rates for East and West Germany. Job reallocation in East Germany has been relatively high following the years after reunification, likely contributing to the rapid productivity growth during these 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 between 1993 and 2015.

An economy may reallocate workers across plants also without reallocating jobs, for example, to improve match quality between existing jobs and workers. Figure A4 (b) shows that East Germany 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

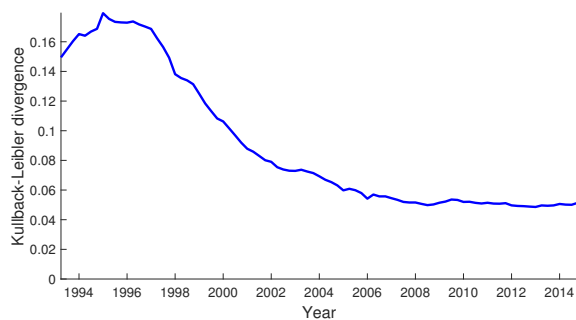
²⁸[Boeri and Terrell \(2002\)](#) find that such job reallocation has indeed been important in understanding productivity growth in former Soviet Republic countries. Even for the U.S., the evidence suggests that much of long-run productivity growth is driven by the reallocation of jobs from less to more productive plants (see [Foster, Haltiwanger, and Krizan, 2001](#)).

the West level afterward. [Dauth, Findeisen, Lee, and Porzio \(2022\)](#) show that the high labor reallocation after the reunification was, indeed, from low- to high-paying plants, contributing to the initial rapid wage growth in East Germany.

The third panel, [Figure A4 \(c\)](#), considers another notion of 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, East Germany is the economy with more reallocation.

Yet another notion of reallocation is the growing and shrinking of industries. Since the industry composition has been significantly different in East Germany 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 A5](#) plots the Kullback-Leibler divergence as a measure of the distance between the West and East German employment distributions 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 convergence in industry structure, nor does the period of convergence in industry structure, that is 1995 to 2008, exhibit particularly strong 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 industries as in the economy as a whole: East Germany is less productive in each industry, 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.

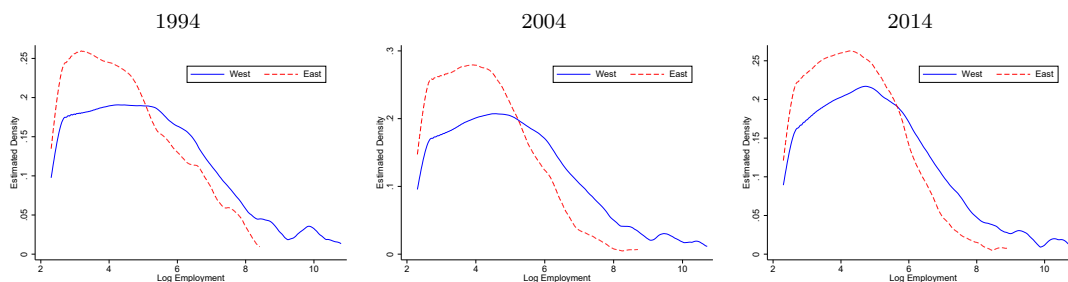
Figure A5: Industry convergence



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

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

Figure B1: Size distribution *AWFP*



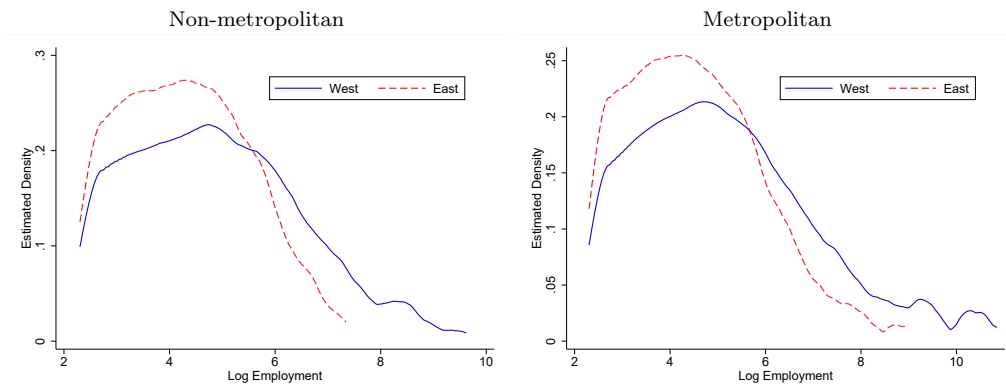
Notes: The figure displays employment-weighted plant size distributions in form of an estimated density function (by a Gaussian kernel smoother) in the total private, non-primary sector in 1994, 2004, and 2014. Data source: *AWFP*.

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 use 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 B1 displays the density of plants over log employment in East and West Germany starting in 1994. The East-West size distribution differences have been fairly stable between 1994 and 2014.

Figure B2 displays plant size distributions conditional on a plant being located in a metropolitan area. To define these areas, we use the definition from [Dijkstra, Poelman, and Veneri \(2019\)](#). The figure shows that metropolitan areas have more employment at large plants than non-metropolitan areas (the estimated employment density in the right panel has a fatter right tail). 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 B2: Size distribution *AWFP* metropolitan areas, 2014



Notes: The figure displays employment-weighted plant size distributions in form of an estimated density function (by a Gaussian kernel smoother) in the total private, non-primary sector, splitting the sample by plants being located in a non-metropolitan area (left panel) or metropolitan area (right panel). Metropolitan areas are defined as in [Dijkstra, Poelman, and Veneri \(2019\)](#), based on functional urban areas. Data source: *AWFP*.

C Industry Classifications

Table C1: Industry classifications

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

Notes: The table provides a crosswalk that maps the 21 industries used in this paper into the industry classifications used by the *SES*, the 2003 and 2008 industry classifications, and the SNA-ISIC-A38-level industry classification from the national accounts (also based on WZ08). The latter is less detailed, and we have to group some industries. Parentheses show the name we give to the respective industry group. *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.

Industry classifications have undergone several revisions since reunification. The *AWFP* data contains WZ08 2-digit “Abteilungen” as industry classification. Also, the 2010 and 2014 samples of the *SES* use the 2-digit WZ08 classification. The 2006 sample from the *SES* uses the WZ03 classification. Finally, national accounts are organized by the SNA-ISIC-A38 level of the WZ08 classification. The latter is less fine grained than the 2-digit level. Table C1 provides a cross-walk across the different classifications.

D Robustness of the Size-Wage Nexus

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 analysis 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
Difference $\hat{\omega}_E - \hat{\omega}_W$	1.9 (0.3)	1.5 (0.3)
N (in thousands)	2365	2365
	Adding part-time	More region-specific controls
Difference $\hat{\omega}_E - \hat{\omega}_W$	2.0 (0.2)	2.0 (0.3)
N (in thousands)	3074	2365

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 second order polynomial 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 3rd order polynomial 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. *More region-specific controls:* The same as the baseline estimate but allows workers' age and sex as well as industry effects to be region specific. Data source: *SES* 2006/10/14.

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 size-wage relationship in East Germany would simply reflect that there are more small plants there. To allow for this possibility, we replace the common linear term $\omega_W \ln E_{it}$ in regression (1) by a common non-linear term, $F(\ln E_{it})$,

that takes the form of either a 2nd-order or 3rd-order polynomial:

$$\ln w_{it} = \beta_0 + \beta_E \text{East}_i + F(\ln E_{it}) + (\hat{\omega}_E - \hat{\omega}_W) \text{East}_i \ln E_{it} + \beta x_{it} + e_{it}. \quad (\text{D.1})$$

The first column of Table D1 shows that allowing for a 2nd-order polynomial leads to, if anything, an even steeper size-wage curve in East relative to West Germany. Using instead a 3rd-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 bottom panel of 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. Finally, we allow worker characteristics and industry effects to have heterogeneous effects on wages across the two regions (“More region-specific controls”). That is, we allow in a flexible way for worker sorting based on observables to have different wage effects in the two regions and industry-level demand to be different across the regions. Again, we find that the differences in the size-wage elasticities become even a little larger than in our baseline specification.

D.2 Plant-Level Data

In Section 3.3, we control for worker heterogeneity and worker sorting by observable worker characteristics: age, sex, education, occupation, and job levels. The plant-level AWP data together with Bellmann, Lochner, Seth, and Wolter (2020) allows us to control for unobserved worker heterogeneity, too. Specifically, Bellmann, Lochner, Seth, and Wolter (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:

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

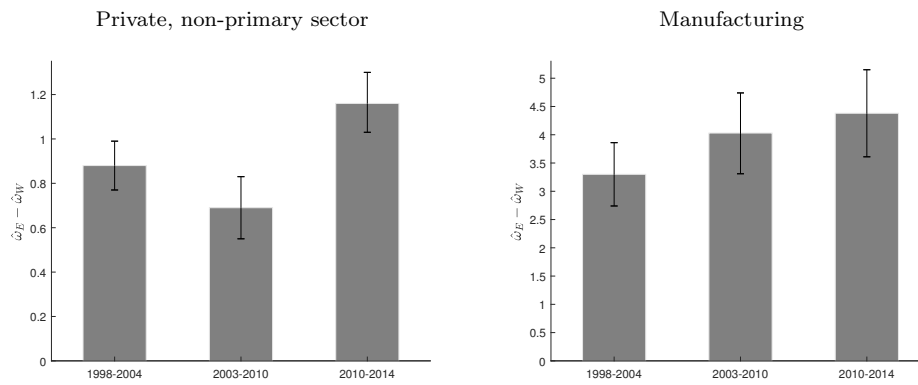
where w_{ijt} are the daily earnings of worker i at plant j in period t , x_{it} are time-varying worker observables, γ_j is a worker fixed effect, and ϕ_i is a plant fixed effect. They provide an estimate of the plant fixed effect, $\hat{\phi}_i$, which we match to the AAFP data. This plant fixed effect 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:

$$\hat{\phi}_i = \beta_0 + \beta_E East_i + \hat{\omega}_W \ln E_i + (\hat{\omega}_E - \hat{\omega}_W) East_i \ln E_i + e_i. \quad (D.3)$$

The left panel of Figure D1 plots the estimates for $\hat{\omega}_E - \hat{\omega}_W$ for the private, non primary sector for all three sample periods. Again, we find that East Germany faces a relatively steeper size-wage, more precisely size-daily-earnings, relationship. The right panel repeats the analysis but restricts it to the manufacturing sector. Reassuringly and as in our baseline results, size-wage differences are particularly pronounced in manufacturing. In other words, these regressions suggests that our baseline findings are not driven by sorting on unobservables. We note that the estimated East-West elasticity difference for the private, non-primary sector is somewhat smaller compared to our baseline (an elasticity of one vs 1.6 percent). This baseline uses practically uncensored hourly wages. The alternative interprets daily top-coded earnings as wage data. This means that deviations from full time hours lead to measurement error in wages in the alternative data set. This measurement error can be expected to vary systematically with plant size stemming from the larger flexibility of work hours at larger plants. This effect is likely to be less important in the manufacturing sector, where we find very similar elasticities across the two approaches: workers in that sector are more likely to work full time.

Conversely, one can ask whether our result of a steeper size-wage curve is driven by high-skill workers in East Germany sorting more into larger plants. [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.

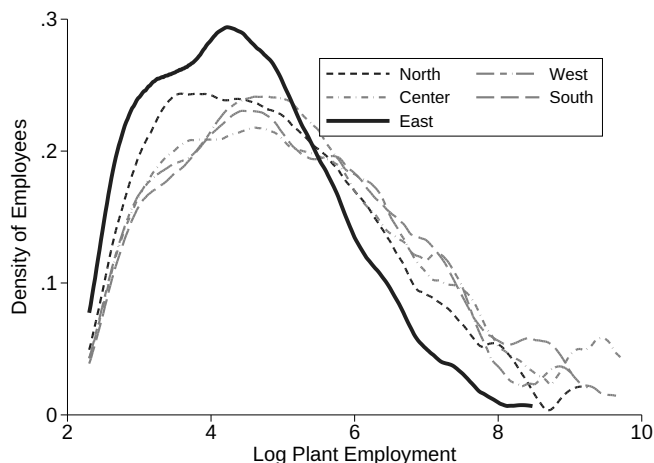
Figure D1: Plant-level size-wage differences



Notes: The figure displays, the difference in the size-wage, more precisely the 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*. Data source: *AWFP*.

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. Data source: *SES* 2006/10/14.

Our baseline analysis distinguishes only between East and West Germany. The *SES* data allow us to further distinguish between five regions in total: North, West, Center, South, and East.²⁹ This appendix extends the analysis and exploits the additional variation coming from the four regions within West Germany. We find the same qualitative patterns as in the main text, however, the relationships have a higher statistical significance.

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.

²⁹North: 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.

Table E1: P-values

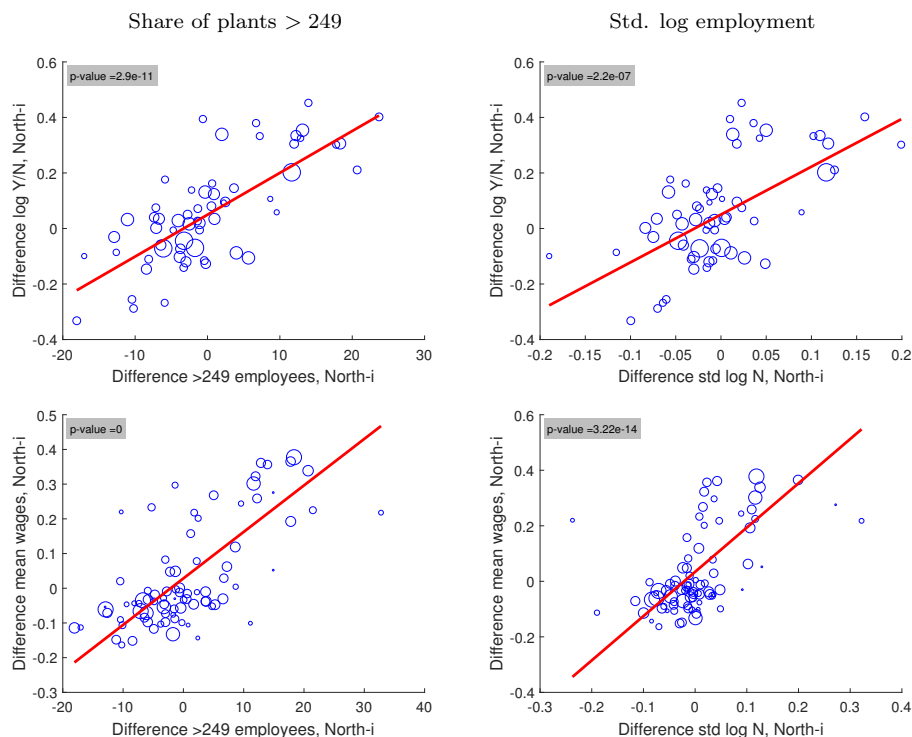
	East-West	East-West + finance dummy
$Y/N, > 249$	0.044	0.091
$Y/N, \text{Std log}$	0.097	0.149
Wages, > 249	0.005	0.000
Wages, Std log	0.006	0.002
$> 249, \hat{\omega}_E - \hat{\omega}_W$	0.185	0.021
Std log, $\hat{\omega}_E - \hat{\omega}_W$	0.142	0.030
Wages, $\hat{\omega}_E - \hat{\omega}_W$	0.008	0.017
$> 249, \text{Collective}$	0.244	0.155
Std log, Collective	0.079	0.055
Wages, Collective	0.121	0.157
	Finer West regions	Finer West regions + finance dummy
$Y/N, > 249$	0.000	0.000
$Y/N, \text{Std log}$	0.000	0.000
Wages, > 249	0.000	0.000
Wages, 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
Wages, $\hat{\omega}_E - \hat{\omega}_W$	0.044	0.043
$> 249, \text{Collective}$	0.073	0.070
Std log, Collective	0.133	0.110
Wages, Collective	0.004	0.004

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, in terms of its definition of a plant as a production unit. Data sources: *SES* 2006/10/14 and *VGR*.

Figure E2 is the analog to Figure 4 in the main text. Those industry/region combinations that have particularly few large plants operating also have low output per worker and low average wages. 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.

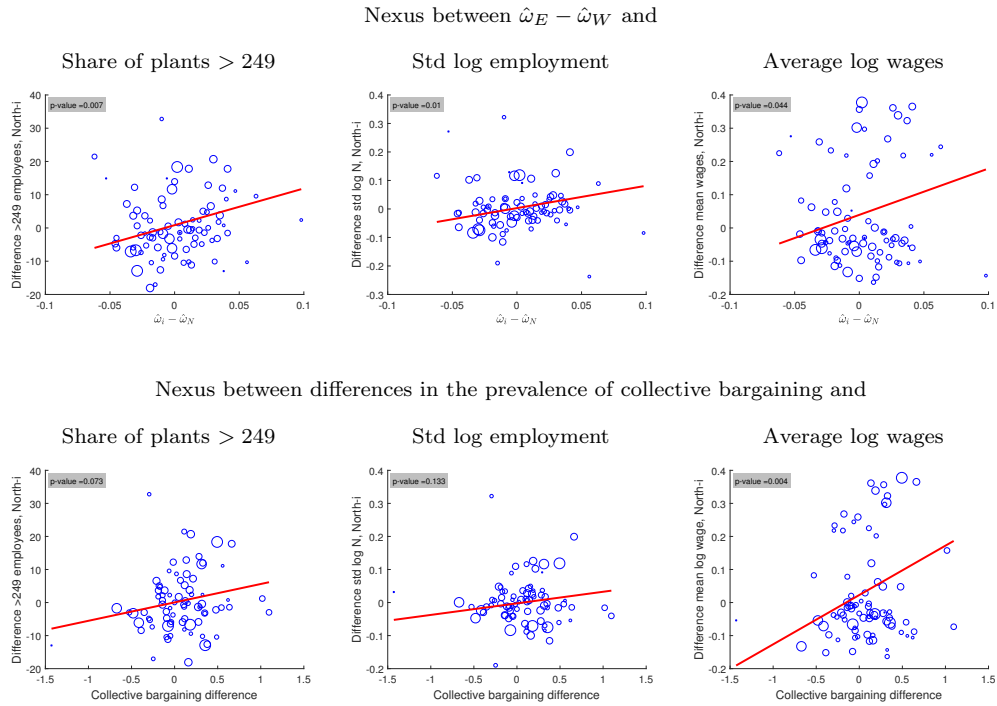
Figure E2: Productivity and wage differences and large plants by industry



Notes: Each dot represents an industry/region combination and displays the difference to the same industry in the North region in Germany. The top panels relate 2014 log differences in output per worker to the share of employment at plants with more than 249 employees (left) and the standard deviation of log plant employment (right). Output is measured as gross value added, which is the GDP concept available at the regional level, because product-specific subsidies and taxes (the difference between the two) are only available at the national level. The lines show (VGR) employment-weighted least squares regressions. The bottom panels relate differences in mean log wages to the same plant size measures. The lines show (SES) employment-weighted least squares regressions. Data sources: *SES* 2006/10/14 (plant sizes, wages) and *VGR* (labor productivity).

That is, we use data for 21 industries paired with the five regions to revisit the relationship between steeper size-wage curves and missing large plants and low wages. The figure displays in the two top panels for each industry within each region the difference in the size-wage nexus against the difference in the share of employment at large plants (left panel), the difference in the standard deviation of log employment (center panel), and the difference in mean log wages (right panel). Those industry/region combinations that have particularly steep size-wage curves also have relatively few large plants operating and have low wages in that

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



Note: Each dot represents an industry/region combination and displays the difference to the same industry in the North region in Germany. The top panel relates differences in the share of employment at large plants, the standard deviation of log plant employment, and average log wages to differences in size-wage relationships. The bottom panel relates differences in the share of employment at large plants, the standard deviation of log plant employment, and average log wages to the following double difference: $[\log P(C|L, R_i) - \log P(C|S, R_1)] - [\log P(C|L, R_i) - \log P(C|S, R_1)]$, 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 (≤ 249 employees) plants in region 1 (North) and region i . The lines show weighted least square regressions. Data source: SES 2006/10/14.

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, the double difference in the prevalence of collectively bargained wage contracts between large and small plants and between regions. We again plot this double difference 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 (center panel). Moreover, the right panel shows the relationship with industry/region differences in average log wages. The relationship between collective bargaining prevalence differences and plant size differences is positive. Industry-region combinations in which the prevalence of collectively bargained wages increases relatively more in plant size are those industry-region combinations where large plants are particularly missing. Similarly, the relationship between collective bargaining prevalence differences and differences in average log wages is positive. Industry-region combinations in which the prevalence of collectively bargained wages increases relatively more in plant size are those industry-region combinations where plants pay lower average wages. Table E1 shows that these relationships are statistically significant.

F A Simple Model of Monopsony Power

This appendix makes explicit the additional output effects arising from the combined presence of customer accumulation, love-of-variety in production, and endogenous producer entry 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 love-of-variety in production, 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 :

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

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

$$\Pi_i = P z_i l_i - l_i \left(\frac{l_i}{\bar{l}}\right)^{\hat{\omega}}. \quad (\text{F.2})$$

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

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

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

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

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

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

It follows that the output of each producer is:

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

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

$$Y = \int y_i di = \phi^{\frac{2}{\hat{\omega}}}, \quad (\text{F.7})$$

which is the analog to (33):

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

which determines output in our main model. Comparing the two equations highlights the importance of producer networks in our baseline model. The monopsony model with heterogeneous producers does, however, feature the labor allocation effect, $\phi^{\frac{2}{\hat{\omega}}}$. 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 sources of output loss: Allocation of workers away from the most productive producers and underinvestment into producer networks. This appendix shows that the standard policy tool to overcome the problem of underemployment, a (proportional) wage subsidy, fails to address these two additional sources of output loss. The intuition for this result, before laying out the argument formally, is 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 the total output of a bundler unchanged *if the individual producer's wage does not change relative to other producers*. This also means that individual employment per known bundler is constant relative to total employment. With 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, confirming the conjecture above. This leaves entry as the only potential margin to be affected 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

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

i.e., the wage subsidy raises the labor demand of each producer for each bundler

that it knows. From this follows the gross profits as a function of known bundlers:

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

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:

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

where $\bar{w} = \left[\left(\bar{P} \frac{\eta-1}{\eta} \right)^\eta \frac{Y}{l} \right]^{\frac{\hat{\omega}}{1+\eta\hat{\omega}}}$ summarizes the other aggregate terms that affect wages. Using (G.3) together with the gross profits, (G.2), and subtracting marketing expenditures yields the operating profits:

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

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

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

This equation, at first glance, seems to suggest that a wage subsidy indeed increases relative customer accumulation proportionally and for all firms. However, this is logically 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 25) that determines the individual producer's size of the customer network relative to the average:

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

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

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

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

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

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

From (G.10) follows that the distribution of output per bundler and, hence, employment per bundler is unchanged compared to the main text. In particular, they do not depend on the wage subsidy. Together with (G.6) this implies that the distribution of employment across plants, $l_i \Theta_i$, 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:

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

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

$$\frac{Y}{\Gamma} \frac{1}{\eta} = \lambda + \mu \bar{\Theta}. \quad (\text{G.12})$$

Similarly, we can derive again the average network size:

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

where again τ does not show up explicitly.

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

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

yields for Y :

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

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. This also means that all producers in the network become more productive in producing final output. What is more, the increase in output increases aggregate demand leading to a yet larger optimal network size. A Ramsey planner that can subsidize marketing and thereby freely choose the marketing costs, $\tilde{\mu}$, that private producers take into account, while the planner still has to pay the physical marketing costs, μ , would maximize output minus real costs, i.e., consumption, which also equals labor compensation:

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

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

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

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

$$\Gamma = \frac{1}{\lambda} \frac{\eta-1}{\eta} \frac{\hat{\omega}}{1 + \eta\hat{\omega}} Y. \quad (\text{H.4})$$

Combining these equations yields:

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

The corresponding first-order condition is given by:

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

where, using (H.2),

$$\frac{\partial Y}{\partial \tilde{\mu}} = -\frac{1}{\eta-2} Y \frac{1}{\tilde{\mu}}, \quad (\text{H.7})$$

and, hence,

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

Rearranging yields:

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

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, the size-independent marketing subsidy only addresses the effect *sans* heterogeneity on network size. It does not remedy the allocation of workers to relatively unproductive plants. This follows from the observation that the first-order condition is independent of the labor allocation effect, $\phi_{\tilde{\omega}}^{\frac{2}{\eta}}$. By extension, the effect *cum* heterogeneity on network size, $\phi_{\tilde{\omega}}^{\frac{2}{\eta-2}}$, is also not remedied. In other words, a size-independent subsidy on marketing expenditures cannot alter the output losses arising from the compressed distribution of labor across producers.