No. 91

U N I V E R S I T Y O F C O L O G N E

W O R K I N G P A P E R S E R I E S I N E C O N O M I C S

INCORPORATION DECISIONS AND JOB CREATION IN NEW FIRMS

SUSANNE PRANTL
FREDERIK THEN£E

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

Incorporation Decisions and Job Creation in New Firms*

Susanne Prantl[♯] and Frederik Thenée[§]

January 27, 2017

Abstract:

The regulation of firms can influence job creation in new firms, and these influences may contribute to long-term effects on the economy-wide distribution of jobs across firms. In this paper, we study the influences on job creation in new firms that follow from incorporation law, specifically the law-induced cost at firm entry. To identify these influences empirically, we allow for endogenous decisions on incorporation and, in addition, exploit a natural experiment in regulation that accompanied German reunification. Our empirical findings are in line with predictions that we derive from the model of Lucas (1978) by integrating incorporation decisions. We show that an increase in incorporation-related entry cost reduces the emergence of incorporated firms in a population of new firms. In addition, such an increase leads to higher initial job creation in incorporated than in unincorporated firms and decreases the mass of the entry size distribution across incorporations relative to the distribution across unincorporated firms in the intermediate range.

*Financial support provided by the German Research Foundation (DFG) in the context of the Priority Programme 1764 is gratefully acknowledged.

[♯]University of Cologne (prantl@wiso.uni-koeln.de, http://www.ieam.uni-koeln.de).

[§]University of Cologne (thenee@wiso.uni-koeln.de; http://www.ieam.uni-koeln.de).

1 Introduction

The regulation of firms can influence job creation in new firms, and these influences may contribute to long-term effects on the economy-wide distribution of jobs across firms. In this paper, we set out to study the influences on job creation in new firms that follow from incorporation law, specifically the law-induced level of fixed cost at firm entry. Incorporation-specific entry cost influence entrepreneurs' decision on whether to start an incorporated firm, or not, and they influence job creation at market entry. Studying factors that determine the emergence of incorporated versus unincorporated firms in a population of new firms, as well as initial job creation in these two types of firms, is important. In particular, because incorporated start-ups are likely to qualify as entrepreneurial entrants that tend towards job creation and firm growth while unincorporated start-ups may rather represent entrants with limited prospects (Levine and Rubinstein, 2016, among others).

We provide comprehensive empirical evidence on entrepreneurs' incorporation decisions and on job creation in new firms reacting to the main incorporation law in Germany during the 1990s. Our evidence is consistent with predictions derived from a theoretical framework that builds on Lucas (1978). We model a population of individuals that differ in their entrepreneurial type and we let each individual opt for entrepreneurship or work as an employee. Each entrepreneur chooses whether to incorporate, or not, and decides on the labor input, i. e. the size, of her firm. An initial prediction concerns the influence of the incorporation-related level of entry cost on the extent to which incorporations emerge at market entry in the population of new firms. The main predictions refer to the extent to which initial job creation occurs in incorporated or unincorporated entrants.

In the empirical analysis, we focus on new firms in East and West Germany after the unexpected and major event of German reunification. The shock-inducing reunification event and the legal setting at the beginning of the 1990s benefit our effect identification strategy in three respects.

Entrepreneurs who start a new firm will take risk expectations into account when choosing whether to incorporating or not. The most common type of corporate legal form among start-ups in the 1990s in Germany was the company with limited liability (CLL, Gesellschaft mit beschränkter Haftung, GmbH) and setting-up a CLL implied liability of the firm owners up to their equity share in the firm. Owners of unincorporated firms were, instead, personally liable to an unlimited extent. The relevance of risk expectations in the incorporation decision renders incorporation potentially endogenous in equations explaining job creation at market entry. To address the issue, we exploit a first advantage of our focus on new firms in Germany during the 1990s: data from the time *before* the unexpected, major event of German

reunification provides us with a credible and relevant source of exogenous variation in legal form choices. The excluded instrument that we generate indicates for each post-reunification start-up in our main data a relevant predicted incidence of CLLs among firms that entered before the reunification shock. Using that instrument, we can model firms' legal form choices as endogenous in equations explaining firms' job creation at market entry, and we estimate these equations using two-stage least squares with generated instruments.

A second advantage of the time period under consideration is that entrepreneurs in Germany faced substantially higher entry cost when setting up a new incorporated firm instead of an unincorporated firm. The statutory requirements of the German Law on Companies with Limited Liability imply high financial cost and high cost of time. The latter are high as the path towards establishing the independent legal entity requires various activities involving interactions with lawyers, notaries and administrative authorities. Providing the minimum capital of about $\$30,000$ that was then required at the moment of setting up a CLL implies high financial cost for new firms. Sizable statutory minimum capital requirements were then also common in many other European countries, for example in Austria, Belgium and Greece, and they still are of high importance (Becht et al., 2008, World Bank, 2015). Further sizable financial cost accumulate due to, in particular, high fees for the mandatory services of lawyers and notaries.[1]

The third advantage consists of the natural experiment in regulation which accompanied the unexpected and major event of German reunification. Potential entrepreneurs in East Germany got suddenly confronted with the same, complex legal rules as West Germans when starting a new incorporated firm, and also with identical rules when starting an unincorporated firm.[2] Despite the identical legal rules, we argue that the entry cost following from the law on CLLs were substantially higher in East than in West Germany after reunification, except for entrants in the East and the West that are homogeneous with respect to all cost-related respects.[3]

The main reasons for higher incorporation-related entry cost in East than in West Germany are differences in CLL-specific financial cost, cost of time and opportunity cost. East German entrepreneurs had much lower personal wealth than West Germans, and it is straightforward to show that this can cause higher financial

[1] See Section 3 and Djankov et al. (2002) on the much higher financial cost in Germany compared to the US.

[2] Following the treaty regarding the implementation of a monetary, economic and social union (Vertrag über die Schaffung einer Währungs-, Wirtschafts- und Sozialunion) on May 18^{th}, 1990, core parts of the West German legal framework, including the West German version of the law on CLLs, were enacted in East Germany on July 1^{st}, 1990.

[3] We separate the homogeneous group of entrants from other groups in our main sample to check the plausibility of our identification strategy. See the subsections starting on pages 15 and 23 in Section 4 for details.

cost when needing financial means to fulfill a high minimum equity requirement, or to cover high fees for legal and notarial services. Higher cost of time result because East Germans had no time to prepare for dealing with the complex statutory entry requirements of the law on CLLs. The relevant notaries and lawyers specialized on commercial law were lacking in the East in the years after reunification. Most of the administrative authorities that needed to get involved in the creation of a CLL were undergoing substantial transition. Higher opportunity cost than in the relatively stable West German market environment arise because entering a market at the right moment of time is more likely to be crucial for the profitability of a firm in case of a transition economy.

Against that background, we implement a cross-sectional difference in differences approach, in combination with the already mentioned IV approach, when specifying size equations. Exploiting the natural experiment in regulation, we identify effects of a difference in incorporation-related entry cost by comparing the difference in job creation between firms started as CLL and unincorporated firms in the East with the corresponding difference in the West.

Before turning to the main results, we focus on the prediction regarding the extensive margin: an increase in incorporation-related entry cost is predicted to lower the probability of incorporated relative to unincorporated entrants. What we observe is a lower probability of starting a CLL instead of an unincorporated firm in East than in West Germany after reunification. This is consistent with the prediction building on the natural experiment in regulation according to which the East is the region where CLL-specific entry costs are not only higher than the cost for unincorporated entry, but exceptionally high compared to the West. Providing support for our identification strategy, we observe no difference in the incorporation probability across both regions when focusing on entrants that should be homogeneous in all cost-related respects across the regions.

Our main findings fit with our theory-based hypotheses regarding the impact at the intensive margin. One prediction states that an increase in the incorporation-related entry cost leads to an increase in the entry size difference between incorporated and unincorporated entrants. In line with that, we observe a difference in initial job creation between firms started as CLL and as unincorporated firms that is on average higher in the East than the corresponding difference in the West. The result holds as well if we allow for endogenous legal form choices and no such difference-in-differences arises when focusing on entrants that should be homogeneous in cost-related respects across the regions. The latter result provides again support for our identification strategy. Additional estimates of simple linear probability models, and of a rich set of quantile regression models, provide a further insight which is in line with a related prediction from theory. What we observe is

consistent with the prediction that an increase in the incorporation-specific entry cost decreases the mass of the entry size distribution across incorporations relative to the distribution across unincorporated firms mainly in the intermediate range.

To complement these results, we study legal form changes.[4] Comparing the annual incidence of incorporating in East and West Germany after market entry, we find evidence that is consistent with higher incorporation-related entry cost deterring more entrepreneurs from incorporating at market entry. These entrants are found to be more likely to revise their initial legal form choice later on, but it takes several years until such revisions get implemented to a non-negligible extent. This result fits with high incorporation-related entry cost in general, and, in particular, in the East after reunification.

Our paper relates to several strands of the literature. At first, these are recent contributions that aim at understanding the impact of specific elements of firm regulation, or specific firm entry policies, on firm size distributions and on the composition of new firm populations, in particular Braguinsky et al. (2011), Garicano et al. (2016) and Hombert et al. (2016). Our focus is a related, but a different one as we investigate theoretically and empirically the impact of entry cost following from incorporation law, specifically the German law on CLLs, on the composition of a population of new firms, the initial labor input decisions of these firms and the resulting size distribution across the new firms.

We add to the existing research on the consequences of reforms to the business registration process in specific countries. Bruhn (2011), Kaplan et al. (2011) and Branstetter et al. (2014), among others, show that regulatory interventiona which strongly reduce the registration cost in Mexico and Portugal, respectively, led to an increase in the number of registered firms. The effects on the average size of entrants are, however, less obvious and depend, in particular, on the role of previously informally operating firms play which decide to register after the reform. Informally operating firms played a large role in these countries, in contrast to the US or Germany. For Germany, Prantl and Spitz-Oener (2009) and Prantl (2012) analyze the entry regulation incorporated in the German Trade and Craft Code imposing a mandatory standard on entrants that want to start their legally independent business in one of the regulated occupations. The empirical approach in these two papers is in one respect related to ours. The main findings indicate that entry regulations reduce entry into self-employment, including entry into long-lasting self-employment. We investigate, in contrast, consequences of the entrepreneurial

[4] Descriptive statistics indicate that legal form changes for Germany show several pattern that are similar to those in the US (see Levine and Rubinstein (2016) and Cole and Sokolyk (2015)). They are, however, less common events in Germany than in the US, and uncommon in general. These differences between Germany and the US reflects the fact that legal form choices involving incorporation decisions are quite expensive in Germany.

choice between alternative modes of entry, i.e. different legal forms of organization. We show theoretically and empirically that differences in the average firm size at market entry are directly related to the chosen legal form at market entry which, in turn, depends on the associated entry cost.

Empirical studies by Becht et al. (2008), Dammann and Schündeln (2011, 2012) highlight the entrepreneurial choice between legal form options by showing that entrepreneurs aim at minimizing the cost of incorporation if they can select between different modes of incorporation. We identify, instead, an appropriate empirical context to show that incorporation-related entry cost can discourage entrepreneurs from incorporating in the first place and, thereby, influence the composition of a population of new firms. We also show that entry cost can influence the distribution of initial job creation across new firms.

Another set of papers focuses on the relationship between a firm's legal form and its subsequent economic performance, in particular Harhoff et al. (1998), Levine and Rubinstein (2016), and Cole and Sokolyk (2015). Our work complements their evidence on the link between a firm's legal form and performance measures like firm growth by shedding light on the mechanism driving the impact of incorporation-related entry cost on legal form choices, as well as on initial job creation in new firms. In addition, we add by treating the legal form decision as endogenous in estimation equations explaining firm size.

The remainder of the paper is organized as follows. In the next section, we introduce our stylized theoretical model and derive hypotheses on the impact of incorporation law on the composition of a population of new firms, and on the distribution of the initial job creation in new firms. The relevant aspects of incorporation law and the historical context of our empirical identification strategy are briefly described in Section 3. In Section 4, we explain our empirical models, sketch the data and discuss our estimation results. We summarize and conclude in Section 5. Our Appendices contain details on theoretical proofs, data, and extensions to the empirical analysis.

2 Theoretical motivation

In this section, we sketch the stylized model that we use to derive hypotheses on the impact of incorporation law on the composition of a population of new firms, and on the distribution of firm size at market entry. The detailed exposition and the proofs are provided in Appendix A.

Building on Lucas (1978), we consider a population of individuals that differ in their entrepreneurial type. Each individual is a potential entrepreneur choosing between entrepreneurship or work as dependent employee. We allow for heterogeneity

in entrepreneurial types such that some individuals will enter entrepreneurship and some will prefer to work as employees. An entrepreneur starting a firm project will realize either positive profits or losses in case of firm failure. Taking the profit distribution into account, entrepreneurs choose between limited and unlimited liability by determining the legal form of their firm at market entry. Entrepreneurs of larger firms have a larger incentive to cover themselves against the risk of firm failure by setting-up a company with limited liability, and they expect sufficiently large profits to deal with the higher entry cost associated with their legal form choice.

In Figure 1, we show the sequence of decisions taken by each potential entrepreneur with type $\theta \sim U[0,1]$, where a higher value of θ indicates a more profitable idea for a firm project. The sequence is such that each individual chooses first between (i) entrepreneurship, that is setting up and managing a firm, and (ii) working as an employee in a firm managed by somebody else. When working as an employee, the individual earns a wage $w \in \mathbb{R}_+$ that is independent of his type θ. If the potential entrepreneur opts, on the contrary, for entrepreneurial activity, he can choose the legal form of his firm, denoted by $L \in \{0,1\}$ with $L = 1$ indicating a company with limited liability (CLL) and $L = 0$ representing an unincorporated firm with unlimited liability. In the case of a CLL, the entrepreneur has to bear fixed cost $F_1 \in \mathbb{R}_+$. If he decides to start an unincorporated firm instead, he has to pay $F_0 \in \mathbb{R}_+$. In line with standard statutory requirements of incorporation laws, we assume that the cost associated with setting up a CLL are larger than the corresponding costs for starting an unincorporated firm, i.e. $F_1 > F_0$.

Having opened up a firm under one of the two legal forms, entrepreneurs can hire workers at wage w to produce a homogeneous good.[5] Following Lucas (1978), we model the production function of a firm managed by an entrepreneur θ as $\theta f(n)$, where f shows decreasing returns to scale in the continuous labor input n. More specifically, we assume that f is of the Cobb-Douglas form $f(n) = n^\alpha$ with $0 < \alpha < 1$. The good is sold on a perfectly competitive market at a price normalized to one. Knowing his entrepreneurial type θ, the potential entrepreneur maximizes the expected profit of his firm. The distribution of profits depends on the chosen legal form and the input choice.

More precisely, setting up a CLL comes with the benefit of reduced risk in case of firm failure for the potential entrepreneur. For this reason, we include uncertainty in the model. In particular, we assume that a firm's profit $\Pi(L, n; \theta)$ consists of a deterministic profit part $\pi(n; \theta)$ that is multiplied by an uncertainty parameter ξ_L that has a positive expectation and is independent of θ.[6] As indicated by the sub-

[5] As the good is homogeneous, its quality does not depend on the legal form choice.

[6] This implies that entrepreneurs with higher types are more efficient in using their inputs and managing their firms, but they are not better in anticipating or reacting to shocks than entrepreneurs with lower types.

index, the cdf of ξ_L, P_{Ξ_L}, depends on the chosen legal form. The main motivation for starting a CLL is the limited liability of the firm's owner in case of firm failure. We model this limitation of liability by assuming that the distribution of the uncertainty parameter for incorporations and, conclusively, the distribution of the incorporation profits is censored at zero.[7] Therefore, in contrast to the corresponding profit of an unincorporated firm, the profit of a CLL cannot become negative in case of firm failure except for the entry cost F_1 that is sunk.

To determine the entry decisions in the population of potential entrepreneurs, we firstly focus on the entrepreneur's maximization problem conditional on entry. Having entered under legal form L the potential entrepreneur θ decides on the labor input, i.e. the size of his firm, solving the following maximization problem:

$$\max_n E_{\Xi_L}[\Pi(L,n;\theta)] \quad = \quad \max_n E_{\Xi_L}[\xi_L] \underbrace{\pi(n;\theta)}_{=\theta n^\alpha - wn} \tag{MAX}$$

As the expected profit maximizing input choice, n^*, increases strictly in θ, the optimal firm size distribution in the economy results immediately from the distribution of entrepreneurial types across the population of potential entrepreneurs. While this is a standard characteristic of the Lucas framework, in our adapted model, it is also the legal form choice that is uniquely determined by θ.

To analyze how entry costs shape the distributions of optimal firm size for the two legal form modes, it is elementary to understand how the choice of the legal form varies across entrepreneurs. A potential entrepreneur θ will decide to start an unincorporated firm if the expected profit of starting an unincorporated firm is larger than the outside option of becoming a dependent worker (entrepreneurship condition, E_0) and if it is additionally more profitable to set up an unincorporated firm than to set up a CLL (legal form condition, LF_0):

$$E_{\Xi_0}[\Pi(L=0,n^*;\theta)] - F_0 \geq w \tag{E_0}$$
$$E_{\Xi_0}[\Pi(L=0,n^*;\theta)] - F_0 > E_{\Xi_1}[\Pi(L=1,n^*;\theta)] - F_1 \tag{LF_0}$$

The sufficient conditions for an incorporation are:

$$E_{\Xi_1}[\Pi(L=1,n^*;\theta)] - F_1 \geq w \tag{E_1}$$
$$E_{\Xi_1}[\Pi(L=1,n^*;\theta)] - F_1 \geq E_{\Xi_0}[\Pi(L=0,n^*;\theta)] - F_0 \tag{LF_1}$$

To determine how entrepreneurs select the legal form of organization, we analyze how these inequalities change in response to changes in the entrepreneurial type θ.

[7] We assume that $P_{\Xi_1}(\xi) = P_{\Xi_0}(\xi)$ if $\xi > 0$, $P_{\Xi_1}(\xi) = \int_{[\underline{\xi},0]} 1\, dP_{\Xi_0}$ if $\xi = 0$, and $P_{\Xi_1}(\xi) = 0$ if $\xi < 0$, where $\text{supp}\{\Xi_0\} = [\underline{\xi}, \bar{\xi}]$ with $\underline{\xi} < 0 < \bar{\xi}$.

Our assumptions on the profit function, i.e. both the inclusion of the type into the production function and the proportionality of uncertainty to the deterministic profit part, have two important implications. Entrepreneurs with higher types expect larger profits but these are, in turn, also less certain. As a result, entrepreneurs with higher types are more inclined to incorporate as the benefit of reduced risk increases in θ. The legal form constraints reflect this fact. The left-hand side of (LF_1) increases stronger in θ than the right-hand side as $\partial E_{\Xi_1}[\Pi(L=1,n^*;\theta)]/\partial\theta > \partial E_{\Xi_0}[\Pi(L=0,n^*;\theta)]/\partial\theta$. As a consequence, the legal form choice in equilibrium is characterized by thresholds. These are the thresholds θ_0^* and θ_1^* with $0 < \theta_0^* \leq \theta_1^* \leq 1$ such that potential entrepreneurs with $\theta \in [0, \theta_0^*[$ become workers, those characterized by $\theta \in [\theta_0^*, \theta_1^*[$ start unincorporated firms, and potential entrepreneurs with $\theta \in [\theta_1^*, 1]$ incorporate.

In the following, we concentrate on equilibria in which firms of both legal form modes exist as we are interested in the impact that differences in entry cost have on both incorporated and unincorporated firms simultaneously. The most important implication for the characterization of the corresponding equilibrium is that (E_0) and (LF_1) are binding, and define θ_0^* and θ_1^*, respectively.

Having characterized the equilibrium, we can investigate how an increase in the entry cost associated with incorporating, F_1, affects the composition of entrants in the two legal form modes, as well as the distribution of the optimal firm size at market entry.

While (E_0) and, conclusively, θ_0^* are unaffected by the change in CLL-related entry cost, the left-hand side of (LF_1) decreases in F_1 implying an increase in θ_1^*. This increase mirrors the reduced tendency to incorporate under the higher incorporation cost. More precisely, some potential entrepreneurs decide to substitute incorporating by setting up an unincorporated firm. Remembering that the share of potential entrepreneurs that incorporate is equal to $1-\theta_1^*$ and the share of those who start an unincorporated firm is $\theta_1^* - \theta_0^*$, we advance the first hypothesis.

Hypothesis 1. *An increase in the entry cost for CLLs, F_1, decreases the probability of CLLs conditional on entry.*

As we are also interested in implications of the model for the distribution of firm size at market entry, we define the average optimal firm sizes in the two modes of entry as

$$E_\Theta[n^*|L=0] = \int_{\theta_0^*}^{\theta_1^*} n^*(\theta)/(\theta_1^* - \theta_0^*)\, d\theta \qquad (S_0)$$

$$\text{and} \quad E_\Theta[n^*|L=1] = \int_{\theta_1^*}^{1} n^*(\theta)/(1-\theta_1^*)\, d\theta\ . \qquad (S_1)$$

Building on our previous observations that n^* increases strictly in θ and that larger CLL-related entry cost imply a larger CLL threshold θ_1^*, equations (S_0) and (S_1) show that larger CLL-related entry cost lead to an increase in the optimal firm size in both modes of entry. The increase in the average entry size is, however, not identical across legal form modes. In contrast, the rise in the entry size of incorporations outweighs the increase in the entry size of unincorporated firms. We subsume this observation in the following two hypotheses.

Hypothesis 2a. *An increase in the entry cost for CLLs, F_1, leads to an increase in the average optimal size of firms under both legal form modes.*

Hypothesis 2b. *An increase in the entry cost for CLLs, F_1, leads to an increase in the size difference between incorporated and unincorporated firms.*

If F_1 increases, potential entrepreneurs that would otherwise set up a CLL with relatively small size will adapt their legal form choice. They are the marginal entrepreneurs who switch into setting up an unincorporated firm, but do not abstain from starting a firm at all. As the optimal firm size depends positively on θ, an increase in F_1 leads to a set of unincorporated firms that enter the market relatively large and reduces simultaneously the number of CLLs that enter relatively small. The marginal firms of intermediate type are those that are particularly affected by a change in the entry cost difference, which we summarize in the last hypothesis.

Hypothesis 3. *An increase in the entry cost for CLLs, F_1, mainly deters entrants of intermediate optimal firm size from incorporating.*

3 Incorporation law and the cost of firm entry

3.1 German Law on Companies with Limited Liability

An entrepreneur who wants to start a new firm in Germany can choose the legal form of organization, including, in particular, whether to incorporate or not. The most common type of corporate legal form among start-ups in Germany is the company with limited liability (CLL, Gesellschaft mit beschränkter Haftung, GmbH). A CLL is an independent legal entity, representing a non-public corporate company. The relevant law for the 1990s is the German Law on Companies with Limited Liability of 1980 (Gesetz betreffend die Gesellschaften mit beschränkter Haftung, GmbHG). The most common unincorporated legal form among start-ups is the sole proprietorship (Einzelunternehmen, Gewerbebetrieb).[8]

[8]Limited and general commercial partnerships (Kommanditgesellschaft, KG, and Offene Handelsgesellschaft, OHG) and civil law associations (Gesellschaft bürgerlichen Rechts, GbR) are other

Setting-up a CLL implies liability of the firm owners up to their equity share in the firm. Owners of unincorporated firms are, instead, personally liable to an unlimited extent,[9] and they risk all their distrainable personal wealth in case of firm failure. An entrepreneur who starts a new firm will, therefore, take risk expectations into account when choosing whether to incorporate or not. Taking such links between risk expectations and incorporation decisions into account, we will allow for potentially endogenous incorporation decisions in our estimation equations explaining firm size at market entry, and related outcome variables. Linking back to our theoretical motivation in Section 2, the potential losses of a CLL are downward limited at the level of the sunk fixed cost F_1 in case of firm failure.

Limited liability of CLL owners comes along with a high statutory minimum capital requirement of approximately $\$30,000$ in the 1990s that needs to be fulfilled when setting up a CLL.[10] Fulfilling the minimum capital requirement causes financial cost. In addition, sizable financial cost have to be covered as setting up a CLL involves paying various fees, in particular high fees for mandatory services of lawyers and notaries. Altogether, Djankov et al. (2002) estimate for the year 1999 that these fee-related cost amounted to approximately $\$4,000$, representing 15.7% of the German per capita GDP. This is a large number compared to the \$150 required in the US (0.5% of the US American per capita GDP).

In addition to the financial cost, there are high cost of time as a high number of different procedures have to be completed when setting up a CLL. Djankov et al. (2002) counted 10 procedures for the year 1999, defining a procedure as an activity involving interactions with outside entities. Among the CLL-specific procedures were, in particular, the drawing up of an opening balance sheet, the setting up of a partnership agreement where the minimum content is laid down by law, and getting the agreement notarized. Completing all procedures took 42 full-time equivalents of business days in 1999 under the optimistic assumption that potential entrepreneurs needed no time to gather information and took procedures simultaneously whenever possible. In contrast, such cost of time are much lower in the US, that is only 4 procedures were counted in 1999, and their completion required only 4 full-time equivalents of business days.

unincorporated legal forms. Other corporate firm types that are relevant in Germany during the 1990s are the stock company (Aktiengesellschaft, AG) and the commercial partnership limited by shares (Kommanditgesellschaft auf Aktien, KG a. A.). Among start-ups, these corporate legal forms are, however, hardly observed.

[9] In case of a KG, one owner has to have full personal liability, but not all.

[10] The amount of $\$30,000$ corresponds to 25,000 Euro, or 50,000 DM. In principle, the German Law on CLLs allows for allotting (a share of) the required minimum capital in other form than cash, for example in form of adequate firm assets (GmbHG, §5(4) and §9c(1)). Non-cash contributions are, however, of low practical relevance, if at all then of temporary relevance, for the following reason: they require certification of the values of the respective assets, and it is not only time-consuming and costly to get at that, but it may well fail.

Compared to CLLs, all types of unincorporated firms have in common much lower cost of time as only a limited number of basic procedures need to be completed.[11] In addition, the financial cost related to fees are much lower at the moment of set-up, and there is no minimum capital requirement that needs to be fulfilled. Taken together, the preceding discussion shows that entry cost, consisting of financial cost and cost of time, are much higher when setting up a CLL instead of a firm in unincorporated form, and this is directly reflected in our assumption on the level of fixed cost, $F_1 > F_0$ (see Section 2).

Moving beyond these differences between CLLs and unincorporated firms that are part of our theoretical framework, there are additional differences that are likely to foster more firm growth and job creation in CLLs than in unincorporated firms after firm entry. In particular, CLLs are independent legal entities with unlimited lifetime, implementing a separation between ownership and management is straightforward, and firm (share) transfer follows standardized rules in the case of an ownership change.

3.2 Differences in entry cost after German reunification

To identify effects of differences in entry cost on the composition of a population of new firms and on job creation in new firms, we exploit data variation across regions and legal forms resulting from a natural experiment in regulation which accompanied the unexpected and major event of German reunification.

Building on the treaty regarding the implementation of a monetary, economic and social union (Vertrag über die Schaffung einer Währungs-, Wirtschafts- und Sozialunion; May 18^{th}, 1990), core parts of the West German legal framework, including the West German version of the law on CLLs, were enacted in East Germany on July 1^{st}, 1990. Since then, potential entrepreneurs in East and West Germany have faced the same, complex legal rules when starting a new incorporated firm, and also identical rules when starting an unincorporated firm.

Despite identical legal rules, we expect for three reasons that entry cost following from the law on CLLs vary substantially between East and West Germany after reunification.[12] First, the wealth of East Germans was substantially lower in the years

[11] Among the basic procedures that have to be completed when setting up either a corporate or an unincorporated firm are the clearing of the firm's name and purpose at the Chamber of Industry Commerce and the application for a tax number at a tax office. Due to freedom of form when setting up an unincorporated firm, neither a written partnership agreement nor its notarization is required by law.

[12] In the following paragraphs, we focus on East German entrepreneurs who start firms in East Germany, and on Western entrepreneurs who start firms in the West. According to evidence provided in Section 4, these are by far the most common types of new firm entry after German reunification. In addition, we will also provide empirical results when singling out one important group of start-ups with strong East-West links.

after reunification than the wealth of West Germans. The central bank of the Federal Republic of Germany (Deutsche Bundesbank, 1999) reports, for example, East German gross assets of $16,400 per capita in 1990 compared to West German gross assets of $87,800 per capita. The Income and Consumption Survey (Einkommens- und Verbrauchsstichprobe, EVS) indicates a similar difference for 1993: the reported East German gross financial assets per household were about three times smaller than the West German ones (German Federal Statistical Office, Statistisches Bundesamt (1993)). Due to lower personal wealth, potential East German entrepreneurs who wanted to set up a CLL instead of an unincorporated firm faced higher financial cost than West German ones when covering the high CLL-related fees for legal and notarial services and when fulfilling the minimum equity requirement.

Second, the cost of time associated with setting up a CLL instead of an unincorporated firm were also higher in East Germany than in West Germany, not only the financial cost. Due to German reunification, the complex statutory entry requirements of the law on CLLs got unexpectedly and suddenly relevant for potential East German entrepreneurs, but for potential West German entrepreneurs these requirements remained stable. In addition, setting up a CLL involves using the services of notaries and lawyers specialized on commercial law, and specialists of that kind were lacking in the East in the years after reunification, but not so in the West. More generally, many administrative authorities a potential entrepreneur has to interact with in order to complete the CLL-related procedures were less acquainted with these in the East than in the West, making it particularly hard for East Germans in the years after reunification to overcome the relevant bureaucratic hurdles. Note here that this asymmetry induced East German federal states at that time to call for simplifying the law on CLLs (Niederleithinger, 1992).

Third, East Germany after reunification represented a transition economy. Economic transition in the East started out with a substantial need for restructuring of pre-reunification economic structures and triggered numerous, including highly profitable, entry opportunities. To illustrate the high relevance of entry activities in East Germany, we calculate region-specific entry rates using our main data set.[13] On average, the annual entry rate indicates 2.59 new firms per 1,000 potential entrepreneurs in East German districts during the transition period after reunification. In the more stable West German districts, the average yearly entry rate takes, as expected, a much lower value of 0.62 at that time. Entering the market at the right moment of time is more likely to be crucial for the profitability of a firm in case of the East German transition economy than in the relatively stable West German market

[13]We define the region-specific entry rate as the weighted number of entrants per region normalized by the number of potential entrepreneurs (in thousands) in the respective region. As proxy for the number of potential entrepreneurs serves the respective working age population (see Appendix B.1 for details).

environment. Conclusively, not only the financial cost and the cost of time needed for setting-up a CLL instead of an unincorporated firm but also the CLL-related opportunity cost of time were likely to be larger in the East than in the West.

Taken together, differences in CLL-related financial cost, cost of time and opportunity cost, as well as the underlying differences in the access to financial resources, in the legal and administrative environment and the entry options, were substantial across regions after reunification. Due to these differences, we expect higher entry cost in the East than in the West when setting up a CLL instead of an unincorporated firm after reunification.

We expect this East-West difference despite the fact that the relevant legal rules have been identical in both regions from July 1^{st}, 1990, onwards. Identical legal rules, would only then imply identical entry cost if the population of entrepreneurs, and their entry options, after German reunification were homogeneous in all relevant cost-related respects. This assumption is relevant for our empirical approach towards identifying effects of the German law on CLLs. Accordingly, we will provide empirical tests of its plausibility using a homogeneous sub-population of entrepreneurs in our main data set.[14]

4 Empirical analysis

In this section, we use the empirical context as described above to investigate how entry cost following from incorporation law, specifically the German law on CLLs, shape the composition of a population of new firms and the initial labor input decisions of these firms. First, we focus on the impact of such entry cost on the entrepreneurial decision to start a CLL instead of an unincorporated firm. Then, we turn to the impact of incorporation law on job creation in the new firms at market entry. Finally, we investigate the impact on the incidence of legal form changes after entry.

4.1 Entry as company with limited liability

Empirical model and data

In line with the first hypothesis in Section 2, we expect the probability of starting a CLL instead of an unincorporated firm to decrease in the level of CLL-related entry cost. Taking that expectation to the firm data, we estimate the following simple regression model:

[14] See the subsections starting on pages 15 and 23 in Section 4.

$$\mathsf{CLL}_{idjt} = \beta_0 + \beta_1 \mathsf{E}_d + \eta_j + \vartheta_t + \delta' x_i + \mu' r_d + \varepsilon^1_{idjt}, \tag{1}$$

where we explain the probability of starting a CLL instead of an unincorporated firm using firm-level data. Our empirical context is such that law-induced entry cost when setting up a CLL instead of an unincorporated firm should be higher in the region of East Germany after the reunification shock than in West Germany. For this reason, we expect a negative estimate on the coefficient of the explanatory variable indicating East Germany, denoted by E_d. It is an indicator variable which is coded one for all new firms started in East German districts, and zero otherwise. The parameter β_1 is the coefficient of main interest here. The parameter β_0 is the intercept, and ε^1_{idjt} denotes the error term. The legal form of a new firm i that enters in entry year t, industry j, and district d is indicated by CLL_{idjt}. The indicator is coded one for firms started as companies with limited liability, and zero for unincorporated firms started under a legal form with full liability.

For estimating the model, we use a firm data set on 11,387 new firms that started market activity in East or West Germany after reunification. The data set is well suited for our purposes as it provides firm-level data of the leading German credit rating agency, Creditreform, along with district-level data from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt und Raumforschung, BBSR). An essential advantage of the firm-level data is that it allows for comparing new firms in East and West Germany right after German reunification. Creditreform started to collect data on firms in East Germany early on, using the same standardized data collection and cleaning procedures as in West Germany (see Appendix B and Prantl (2017)). Important is as well that Creditreform collects information on legally independent firms of any size and legal form. Accordingly, our data set covers new firms of all sizes, including many owner-managed entrants without additional workers, and new firms that are started either in unincorporated form or as incorporated entities. In fact, all incorporated firms in our sample are started as CLLs. According to the descriptive statistics in Table 1, CLLs represent 37 percent of the sampled firms.

Sampled firms started market activity between July 1st, 1990, and December 31st, 1993, and they are spread across 22 2-digit industries in manufacturing, construction, trade and services. Approximately half of the firms were started in East Germany, and a large number of districts are covered, 74 West German and 23 East German ones.[15]

[15]Note that we excluded Berlin from our data set. Note also that our results to not depend on that exclusion restriction.

The vector η_j controls for any additive factor specific to 2-digit industries, and ϑ_t for any additive factor specific to entry cohorts. In addition to the data on already mentioned firm characteristics at market entry (legal form, industry classification, date of firm entry, geographical location of the firm), we use further firm data to generate additional explanatory variables, summarized in the control vector x_i. What we use are additional firm characteristics at market entry (number of entrepreneurs in the start-up team, existence of an affiliation to other firms) and data on characteristics of the main entrepreneur per firm (age in years, indicators for educational degrees).[16] The control vector r_d covers a set of district-specific characteristics, that is the population density, shares of workers in manufacturing sector as well as shares of the population at working age (20 to 65 years) proxying the fraction of potential entrepreneurs.[17]

Main estimates

In Table 2, we show linear probability estimates for several variants of the regression model (1).[18] In all variants, observations are weighted to account for the sampling design and to readjust to the structure of the population sampled from (Wooldridge, 2010). Standard errors are clustered to allow for correlation between firms that are active in the same district and industry. In column 1, we show the estimates of the model variant resulting under the assumption that $\delta = 0$ and $\mu = 0$ and that both cohort and industry fixed effects are excluded. In this case, $\widehat{\beta_0}$ reflects the unconditional probability of starting a CLL instead of an unincorporated firm in West Germany (34%) and $\widehat{\beta_0} + \widehat{\beta_1}$ the unconditional relative probability of starting a CLL in East Germany (39%). The negative estimate on the coefficient $\widehat{\beta_1}$ is in line with our expectation and indicates that the unconditional relative CLL probability in the East is 5 percentage points lower than in the West.

In column 2, we take into account that the population of entrepreneurs in our data is heterogeneous with respect to firm- or entrepreneur-specific characteristics, and that therefore our basic results in column 1 may be confounded by the influence of such factors. Accordingly, we let the coefficient δ on the control vector x_i deviate from zero, and add the full sets of controls for 2-digit industry classes and for entry years. In column 3, we add controls for heterogeneity in our data across federal districts, by letting the coefficient μ on the control vector r_d deviate from zero.

[16] We classify university degrees and technical college degrees as high education and vocational training degrees as medium education. The class with low education consists of entrepreneurs who hold neither a vocational training degree nor a higher educational degree.

[17] See Appendix B.2 for further details.

[18] We estimate linear probability models despite the discrete nature of the dependent variable in equation (1), following, among others, Wooldridge (2010). In Appendix Table C1, we also show average marginal effects computed from non-linear probit estimates. Our main findings turn out to be robust to the choice of the econometric method.

Compared to column 1 the negative estimate on the East indicator is smaller after the inclusion of these additional controls, supporting the lower incidence of CLLs in the region with the higher cost of incorporation.

Estimates for a homogeneous group of start-ups

Our empirical context is such that we expect the entry cost when setting up a CLL instead of an unincorporated firm to be higher in the region of East Germany after the reunification shock than in West Germany. We expect that due to substantial East-West differences in CLL-related financial cost, cost of time and opportunity cost and despite the fact that East and West German entrepreneurs have faced identical legal rules from July 1^{st}, 1990, onwards. Identical legal rules should, however, imply identical entry cost across regions if the population of entrepreneurs, and their entry options, after German reunification were homogeneous in cost-related respects.

To provide empirical support for this claim, and thereby for our empirical set-up, we identify a group of start-ups that is rather homogeneous across East and West Germany with regard to access to financial resources, knowledge of legal requirements, access to legal services relevant to setting up a corporation, and expertise in dealing with related administrative hurdles. The group that we identify consists of all start-ups that are affiliated to other firms when entering the market, representing 8.9 percent of the firms in our sample.[19]

For 49 percent of the East German start-ups with affiliations in our data, we know that they are affiliated to a company in West Germany. Due to the transfer of the Western legal and administrative framework to the East, there is no reason to assume that West German companies would be less familiar with the relevant requirements, or less well equipped with access to the relevant legal services, for starting a firm in the East than in the West. In addition, West German companies may directly finance their affiliates, or help them with getting access to financial funds. Some East German start-ups are affiliated to companies that are located outside West Germany. These are also entities who should have general knowledge of legal form choice or entry-related administrative requirements, and may be able to improve an affiliate's access to financial funds. Taken together, affiliated start-ups in East and West Germany have characteristics suggesting similar financial cost, cost of time and opportunity cost.

In column 4 of Table 2, we re-estimate our main model of column 3 on the sub-sample of all non-affiliated firms in East or West Germany. These estimates are similar to our main estimates. In line with expectations, the re-estimate of the coefficient β_1 is more pronounced; the difference is, however, not significant. In

[19]Note that all firms in our sample are legally independent start-ups.

column 5, we turn to re-estimating on the sub-sample of affiliated firms in East or West Germany. The estimate of the coefficient on the indicator for new firms in East Germany is again negative, but small and insignificant. Accordingly, we find no evidence pointing towards influences of sizable differences in entry cost across the two regions for the group of start-ups that is quite homogeneous in cost-related respects. This result is in line with our expectation.

Geographical mobility of entrepreneurs

As mentioned in Section 3.2, we repeatedly focus on locally embedded entrepreneurs: East German entrepreneurs who start firms in East Germany, and Western entrepreneurs who start in the West. Our focus is appropriate in our view as we can show that these are the most common types of firm entry decisions after German reunification. Second, we will derive additional insights from the evidence introduced in the preceding paragraph.

Ideally suited for our first aim are individual-level data from the waves of the 'Career and Qualification Survey' that were carried out in 1991/1992 and in 1998/1999 by the German Federal Institute for Vocational Education and Training (Bundesinstitut für Berufsbildung) and the Research Institute of the Federal Employment Service (Institut für Arbeitsmarkt- und Berufsforschung).[20] Using the individual-level BIBB/IAB-data, we can show that less than five percent of the individuals who take up self-employment between 1990 and 1993 in West Germany are residents in West Germany who spent their youth in East Germany.[21] Data on the size of the ventures started in West Germany indicates that the firms of residents in the West with an Eastern background tend to start smaller than those of residents in the West with a Western background. In addition, we observe that about 2 percent of the self-employed who started a firm between 1990 and 1993 in the East are residents in East Germany who spent their youth in West Germany. Taken together, these pattern suggest that Germans usually stick to their home region when starting as entrepreneur. These pattern arise in the relevant years after the fall of the Berlin Wall despite the fact that all German individuals and legal entities were free at that time to move between East and West Germany. Our findings are well in line with evidence on local embeddedness of entrepreneurs in other countries, for example in Denmark, Italy or the United States (Michelacci and Silva, 2007, Dahl and Sorenson, 2009, 2012).

BIBB/IAB-data on the size of the new ventures that were started in East Ger-

[20] See Appendix B.2 and Prantl and Spitz-Oener (2014).

[21] As we focus here on the migration of entrepreneurs, we complement the work of Hunt (2006) and Prantl and Spitz-Oener (2014) on the emigration from the East German population, or workforce, after German reunification.

many indicates that the few started by residents in the East with a Western background are larger ventures than those of residents in the East with an Eastern background. This observation suggests that this group of individuals overlaps with the group of new firms in the East with affiliations to other firms, including West German ones. As indicated in the last sub-section, keeping such firms in our data set leads, if anything and in line with expectations, to attenuation of our main estimates towards zero, but not significantly so (see columns 3 and 4 of Table 2).

4.2 Effects on firm size at market entry: Main analysis

Empirical model

As stated in the second hypothesis of Section 2, we expect the average entry size of new firms started as CLLs compared to new unincorporated firms to increase in the level of entry cost following from the German law on CLLs. To start the respective empirical analysis, we implement a cross-sectional difference-in-differences approach: Exploiting the natural experiment that arguably lead to higher CLL-related entry cost for new East German than for new West German firms after German reunification, we compare the size difference between firms started as CLL and unincorporated firms in the East with the corresponding size difference in the West.[22] The regression model is:

$$\ln(\mathsf{S}_{idjt}) = \beta_0 + \beta_1 \mathsf{E}_d + \beta_2 \mathsf{CLL}_i + \beta_3 \mathsf{E}_d \times \mathsf{CLL}_i + \eta_j + \vartheta_t + \delta' x_i + \mu' r_d + \varepsilon^2_{idjt} \qquad (2)$$

As dependent variable $\ln(\mathsf{S}_{idjt})$, we use the logarithm of the number of employees, including all working entrepreneurs, in firm i started in district d, industry j and entry year t. The indicator for East Germany E_d, the legal form indicator CLL_{idjt}, the control vectors x_i and r_d, the sets of fixed effects η_j and ϑ_t, and the parameters β_0, β_1, δ, and μ are defined as in Section 4.1. The error term is denoted by ε^2_{idjt}.

Equation (2) allows for two types of additive unobserved factors that can influence the outcome variables of interest. First, the term $\beta_2 \mathsf{CLL}_{idjt}$ will pick up influences of such factors that may vary across firms that are started as CLL or in unincorporated form, but not across regions. Second, there are influences on firm size that may differ systematically across regions while being the same across legal form groups. These will be picked up by the term $\beta_1 \mathsf{E}_d$ in some of our model specifications, otherwise by district-specific fixed effects, entered via the term $\mu' r_d$.

[22] The empirical approach is similar to a standard difference-in-difference approach as we exploit a natural experiment in entry regulation and allow for two types of additive unobserved effects (Blundell and Dias, 2009, Imbens and Wooldridge, 2009, Lechner, 2011). As usually the case in a difference-in-differences framework, we identify an average treatment effect on the treated, in our case the East German firms that are, in contrast to comparison firms in the West, treated with higher law-induced entry cost when starting as CLL instead of starting in unincorporated form.

The coefficient of main interest is the one on the interaction between the legal form indicator, CLL_{idjt}, and the indicator for East Germany, E_d, denoted by β_3. Its estimate will identify the relevant average entry size effect of the East-West difference in the entry cost when setting up a CLL instead of an unincorporated firm under the following assumptions. First, the size difference between firms started as CLL and unincorporated firms in the East would have to be identical to the corresponding size difference in the West if entry cost were identical in East and West Germany. We will provide empirical tests suggesting the plausibility of this assumption in Section 4.3.

Second, firms' legal form choices need to be exogenous to firms' entry size decisions, that is each individual in the population of potential entrepreneurs has to develop and start a firm project either as CLL or in unincorporated form independent of the initial choice of firm size. This assumption is likely to be restrictive as potential entrepreneurs may have a preference for limiting their liability depending on unobserved factors that also influence their decisions on firm size at market entry. Core candidates for such factors are unobserved profit and risk expectations of the potential entrepreneur. For example, a potential entrepreneur who wants to start a firm project for which he expects high profit variance will tend to limit his liability by starting a CLL, and he will presumably be more cautious when hiring employees at market entry. To relax the exogeneity assumption on legal form choices in equations explaining firm size at market entry, we will implement an instrumental variables approach.[23]

Basic estimates

In Table 3, we provide ordinary least squares (OLS) and instrumental variables (IV) estimates of the regression model (2) on our main data set as described in Section 4.1.[24] We show OLS estimates from a basic model variant in column 1: neither influences of fixed effect sets nor control vectors are considered. In column 2, we add the industry-specific fixed effect set η_j, and the entry-year-specific set ϑ_t. In addition, we allow the coefficient δ to deviate from zero to account for potential effects of heterogeneity across firms' main entrepreneurs and start-up teams on initial firm size decisions. Furthermore, to control for district-specific characteristics, we let the coefficient μ deviate from zero. In column 3, we replace the variables previously contained in the control vector r_d by the full set of district-specific fixed effects to account for any time-constant effect of heterogeneity across the districts the new

[23]Note for completeness, that also firms' geographical locations have to be pre-determined to the decision on the firm size at market entry. This assumption is, however, unlikely to be restrictive as Germans usually stick to their home region when starting as entrepreneur. See, in particular, the related evidence at the end of Section 4.1.

[24]Observations are weighted and standard errors are clustered as in Section 4.1.

firms are located in.

Before we turn to the results for the main regressor of interest, we explain our findings on the level terms, linking them to our theoretical discussion. The OLS estimates for the coefficient on the level term of the indicator for East Germany, β_1 are positive and significant (see columns 1 and 2 of Table 3).[25] Not only firms started as CLLs in East Germany, but also new unincorporated firms in the East are larger at market entry than the corresponding Western firms. This finding may pick up influences of additive unobserved factors on entry size that differ systematically across regions, but not across legal forms. In addition, it is consistent with the part of the second hypothesis in Section 2, stating that an increase in the average size of incorporated, as well as unincorporated, firms results if the incorporation-related entry cost increase. The underlying mechanism is such that higher CLL-specific entry cost deter more firms from starting as CLLs, but not from entering at all. These entrants are intermediate firm projects choosing a relatively large entry size compared to those firms that would also start in unincorporated form in the regime with low CLL-specific entry cost. The undeterred firms that enter as CLL also choose a relatively large entry size.

The OLS estimates for the coefficient on the level term of the indicator for CLLs, β_2, are also positive and significant (see columns 1 to 3). Using the estimate in column 3, CLLs are, on average, 65% larger than unincorporated firms. These estimates pick up influences of additive unobserved factors on entry size that vary across legal form groups, but not across regions. They fit with the fact that CLL-related entry cost are, in general, higher than the entry cost when starting a firm in unincorporated form. In addition, the pattern is consistent with the following assumption of our theoretical model: The negative income risk in case of firm failure and, in turn, the benefit from limited liability increase in firm size.

In all three columns, we find positive and significant coefficient estimates on our main regressor of interest, the interaction between the legal form indicator, CLL_{idjt}, and the one for East Germany, E_d. The estimates of β_3 indicate that the average entry size difference between firms started as CLL and unincorporated firms in the East is higher than the corresponding size difference in the West. Using the estimate in column 3, the average entry size difference between CLLs and unincorporated firms is 55% higher in the East than in the West.[26]

Taken together, our OLS results are in line with the entry size decisions of those entrepreneurs to be more affected by the statutory requirements of the German law on CLLs for whom we argue the CLL-related financial cost, cost of time and

[25]Note that we allow for district-specific fixed effects via the control vector r_d from column 3 onwards, instead of separately identifying a coefficient on the indicator for East Germany.

[26]Note that $(\exp\{0.4406\} - 1) \cdot 100\% = 55.36\%$.

opportunity cost to be particularly high when setting up a CLL instead of an unincorporated firm: the East German entrepreneurs starting new firms after German reunification in East Germany. This provides support for our main prediction stated in hypothesis 2b in Section 2: An increase in the CLL-related entry cost leads to an increase in the entry size difference between incorporated and unincorporated firms.

The coefficient estimates on the firm-specific and district-specific characteristics that we added in columns 2 and 3 to our basic model variant of column 1 correspond to our expectations, and are in line with the findings of the related literature.When moving from column 1 to 3, the estimate on our coefficient of main interest, β_3, drops in size, but remains economically large and statistically well significant. Following Gelbach's (2016) sequence-independent conditional decomposition of covariate effects, we can show that 54% of the drop in $\widehat{\beta_3}$ when moving from column 1 to 3 is a result of allowing for the industry-specific fixed effects, η_j. Additional 39% of the drop are a result of allowing for influences of the variables summarized in vector x_i (see also Table C2). Note here that we can show that the main estimates in column 3 remain stable if we use alternative, more disaggregate sets of fixed effects (see Appendix Table C3). In particular, we can use disaggregate 3-digit industry-specific and disaggregate country-specific fixed effects as regressors. We can also consider 2 sets of 2-digit industry fixed effects, one set specific to the East and one specific to the West. The last variant, that we want to mention are fixed effect sets that absorb more data variation across entry years of new firms in our sample, that is industry-cohort and district-cohort fixed effects.

Estimates allowing for endogenous choice of the legal form

As explained above, entrepreneurs who start a new firm take risk expectations into account when choosing whether to incorporate or not. Accordingly, incorporation decisions may be endogenous in an estimation equation explaining firm size at market entry. To identify the impact of starting as CLL on entry size against that background, we use data from the time before the unexpected, major event of German reunification as our source of exogenous variation in legal form choices. The excluded instrument that we generate indicates for each start-up i in our main sample on the time period *after* reunification the predicted incidence of CLLs among the firms entering in 1989 *before* reunification, but in the same industrial environment and the same type of geographical region as firm i. To measure the industrial environment, we use disaggregated, 3-digit industry classes. The geographical regions that we separate are regions with high aggregation of economic activity, intermediate aggregation and rural areas. We denote our generated instrument by $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$.[27]

[27]In Appendix B.1 we explain in greater detail how we generate the instrument.

As German reunification was unexpected, the variation of our instrument is not determined by unobserved factors that are specific to the time after reunification, including new market developments or new opportunities of potential entrepreneurs, and the related entry size decisions. We expect the instrument to proxy general industry-specific and agglomeration-related risk factors that influence an entrepreneur's interest in limiting liability. Accordingly, we expect the instrument to be relevant when implementing an IV approach. The partial correlation between $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$ and the propensity of a new firm to start as CLL should be positive. Conditioning on all explanatory variables considered in column 3 of Table 3, we assume that the pre-reunification CLL incidence has no effect on entry size decisions of new firms after reunification other than through its effect on these firms' legal form choices. We are confident to build on that assumption given our discussion in the last paragraph of the preceding section. We reported there that the OLS estimates on the coefficients of main interest, β_2 and β_3, remain virtually unchanged if we allow for alternative sets of disaggregate fixed effects, each picking up different parts of the data variation within 2-digit industries and within districts (see column 1 in Appendix Table C3).

To implement our IV approach, we estimate a two-stage least squares model with generated instruments. Our second stage equation is specified as the cross-sectional difference-in-differences approach in column 3, up to the fact that we endogenize the terms CLL_{idjt} and $\mathsf{East}_d \times \mathsf{CLL}_{idjt}$. To just identify the model, we use the CLL-share generated from pre-reunification data, $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$, directly as excluded instrument, and interacted with the indicator for East Germany, $\mathsf{East}_d \times \mathsf{I}_{jd,\ \textit{pre-reunific.}}$. As we allow for two endogenous regressors in the second stage equation, we estimate two first stage equations. The identification of the entry cost effect on initial firm size in such a model builds on the standard assumptions of IV models with one endogenous regressor, and on the assumption that the effects of the full set of excluded instruments differ across the first stage equations.

The IV estimates in column 4 of Table 3 indicate a positive average effect of incorporating as CLL on entry size, and an average entry size difference between firms started as CLL and unincorporated firms that is higher in the East than the corresponding size difference in the West. These findings confirm our OLS findings.

Moving from the OLS regression in column 3 to the IV regression in column 4, the estimates on β_2 and β_3 double in value. The direction of the changes fits with the view that entrepreneurs who start firm projects for which they expect high profit variance will tend to limit their liability by starting a CLL, and to be more cautious when hiring employees at market entry. Note as well that the pattern of changes is also consistent with attenuation bias towards zero in the OLS regression, caused by classical measurement error. Using the IV estimate on β_2, CLLs are on average

2.7 times larger than unincorporated firms. Using the IV estimate on β_3 in column 4, the average entry size difference between CLLs and unincorporated firms is 2.5 times higher in East than in West Germany.

In columns 1 and 2 of Table 4, we summarize the first stage estimates. The excluded instruments turn out to be strong predictors of the legal form choices of entrepreneurs starting new firms after German reunification. High values of the test-statistic in the F-tests on the irrelevance of the excluded instruments, reported at the bottom of columns 1 and 2, indicate rejection of the null hypothesis. The effects of the excluded instruments differ across the first stage equations in columns 1 and 2. In line with expectations we find the partial correlation between CLL_{idjt} and $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$ to be strongly positive, as well as the one between $\mathsf{East}_d \times \mathsf{CLL}_{idjt}$ and $\mathsf{East}_d \times \mathsf{I}_{jd,\ \textit{pre-reunific.}}$.

We find no indication of weak instrument problems, using two types of statistical tests. First, the Kleibergen-Paap F-test on weak identification of the whole set of endogenous regressors strongly rejects the null-hypothesis.[28] Second, the Sanderson-Windmeijer F-test testing weak identification per individual endogenous regressor rejects as well.[29] When inspecting the unbiased reduced form estimates in column 3 of Table 4 (Angrist and Krueger (2001)), we find again no indication of weak instrument problems. First, the coefficient estimates on $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$ and $\mathsf{East}_d \times \mathsf{I}_{jd,\ \textit{pre-reunific.}}$ differ significantly from zero. Second, the estimates show that the average entry size difference between CLLs and unincorporated firms increases by 20% in the East relative to the West when moving from the lower to the upper quartile in the distribution of the pre-reunification CLL incidence.

4.3 Effects on firm size at market entry: Extensions

Estimates for a homogeneous group of start-ups

With our main estimation approach in Section 4.2, we build on the view that setting up a CLL instead of an unincorporated firm involves higher entry cost in the region of East Germany after the reunification shock than in West Germany. We argued that this gap in entry cost explains why we find a positive difference in the average entry size differences between CLLs and unincorporated firms in the East and the West. That positive difference-in-differences should disappear, in line with our argumentation, if the population of entrepreneurs, and their entry options, after German reunification were homogeneous in entry cost-related respects.

To provide support for this claim, and thereby for our empirical approach, we proceed in a similar manner as in the sub-section starting on page 15. We re-estimate

[28] The test-statistic is provided in the last row of Table 3.
[29] The test-statistics are indicated at the bottom of Table 4.

our main model of column 3 in Table 3 separately for two different firm groups. The two groups result from singling out the new firms in East and West Germany that have affiliations to other firms and, therefore, characteristics suggesting similar financial cost, cost of time and opportunity cost.

First, we provide the OLS and the IV re-estimates of our main model on the group of all non-affiliated firms in East or West Germany (see columns 1 and 2 of Table 5). These estimates are very similar to our main estimates in columns 3 and 4 of Table 3. In column 3 of Table 5, we turn to re-estimating on the complementing sub-sample of affiliated firms in East or West Germany.[30] The reported OLS estimate of the coefficient on the incorporation indicator is positive, large and highly significant for the group of affiliated start-ups, and this is in line with our main findings and our expectation. The interaction between the incorporation indicator and the indicator for East Germany is, in contrast to our main findings, small and insignificant. As expected, we find no indication of sizable differences in entry cost across the two regions in case of a group of start-ups that is quite homogeneous across regions with regard to resource access, legal knowledge, and administrative expertise.

Estimates allowing for heterogeneous effects of the legal form choice on entry size

The average effect estimates in Section 4.2 are, in summary, consistent with the view that the entry size decisions of one group of entrepreneurs are particularly restricted by the statutory requirements of the German law on CLLs. These are the entrepreneurs who face particularly high CLL-related financial cost, cost of time and opportunity cost when starting a CLL instead of an unincorporated firm.

In this section, we proceed with searching for heterogeneity in the effect of the legal form choice along the conditional entry size distribution. We do so as our theory in Section 2 suggests that higher CLL-related entry cost decrease the mass of the CLL-specific entry size distribution relative to the distribution for the unincorporated firms mainly in the intermediate range.

To start with, we estimate simple linear probability models, each explaining the incidence of a new firm to start in a specific entry size category. We consider three size categories[31] and the same set of regressors as in our main model specification in column 3 of Table 3. We report the OLS and IV estimates in Table 6.

The OLS and IV estimates in columns 1 and 2 of Table 6 indicate for both entry cost regimes, the one with relatively high CLL-related entry cost in East Germany and the alternative one in West Germany, that CLLs are less likely than

[30] We abstain from providing IV estimates for the sub-sample of affiliated firms due to the small size of the sub-sample.

[31] 33.7% of the firms in our data start with one employee, 35.4% with 2 or 3 employees, and 30.8% with at least four employees.

unincorporated firms to start in the smallest entry size category. According to the IV estimate for β_3 in column 2 the effect is by 16 percentage points more negative in the East than in the West, but note that the OLS estimate in column 1 does not confirm. In columns 5 and 6, the OLS and IV estimates show that CLLs are more likely than unincorporated firms to start in the highest entry size category. That effect turns out to be more positive in the East than in the West, and the estimates indicate additional 19 to 39 percentage points. The incidence to start in the intermediate entry size category is similar for CLLs than for unincorporated firms according to the OLS estimates in column 3. The IV estimates in column 4 suggest, instead, that it is weakly higher for CLLs. Both types of estimates indicate that the difference in the incidence between CLLs and unincorporated firms is between 23 and 25 percentage points lower in the East than in the West.

The latter pattern stands out and is in line with the following view: relative to unincorporated firms, CLLs have unequivocally a lower incidence to start with intermediate entry size in a regime with relatively high CLL-specific entry cost than in an alternative regime with relatively low CLL-specific entry cost. In case of the incidence to start with small entry size, the pattern is less pronounced, but tends to be similar. In case of the high entry size category, the pattern is in line with CLLs having, relative to unincorporated firms, unequivocally a higher incidence in a regime with relatively high CLL-specific entry cost than otherwise. These findings are in line with our expectation.

We proceed with evidence on heterogeneity in the impact of the legal form choice when estimating average treatment effects across quantiles of the conditional entry size distribution (Koenker, 2005). To that aim, we estimate separate quantile regression equations each expressing a decile of the conditional entry size distribution as a linear function of the same set of regressors as in column 3 of Table 3.[32] In Panel A of Table 7, we provide the estimates, starting with the 30% quantile as 33.7% of the firms in our sample are 1-person firms.[33] To deal with potential endogeneity of the legal form choice in these quantile regressions, we take the straightforward approach of estimating the reduced form variants (see Panel B of Table 7).

All the quantile effect estimates in Table 7 indicate that CLLs start larger than unincorporated firms, and that this entry size difference is more pronounced in the East than in the West. Fitting in that respect with the average effect estimates in Section 4.2, the quantile effect estimates provide an important additional insight: the size of the estimated difference in the entry size differences depends on the decile

[32] Standard errors are robust and observations are weighted as described in Section 4.1.

[33] Alternatively, we directly account for the discreteness of our entry size variable by implementing the quantile regression approach for count data as proposed by Machado and Silva (2005) in Appendix C (Koenker, 2005, Cameron and Trivedi, 2010).

of the entry size distribution that we focus on. According to the β_3-estimates in Panel A of Table 7, the difference-in-differences is 101% in case of the 40%-quantile, and only 43% in case of the 80%-quantile. The reduced form estimates in Panel A of Table 7 are qualitatively similar, and provide the highest difference-in-differences estimates when we focus on intermediate deciles.[34]

Taken together, the quantile regression results are consistent with our theory. The entry size distribution of CLLs is not only to the right of the entry size distribution of unincorporated firms in both cost regimes. It is also the case that the East-West difference between the entry size distributions of incorporations is relative to the entry size distributions of unincorporated firms particularly pronounced if we focus on entrants at the intermediate size deciles.

Estimates using alternative sample restrictions

The empirical evidence in Section 4.1 indicates that East German individuals hardly relocate to and start a firm in West Germany in the early years after German reunification and, therefore, such geographical mobility is unlikely to influence our main empirical findings. There are, however, social ties that result from Germany's history after World War II and that get valuable after German reunification: Burchardi and Hassan (2013) report positive effects of these ties to East Germany on the personal income of West Germans, and on the development of West German regions. In particular, they show that ties to East Germany as measured at the level of West German regions increase the regional share of West German households who become entrepreneurs, the returns to West Germans' entrepreneurial activity per region, and the likelihood that firms located in a West German region invest in East Germany. Taken together, these findings indicate that differences in social ties linking East and West Germany may influence entrepreneurial activity in West Germany after reunification.

To investigate whether such differences influence our main empirical findings, we take advantage of the data that Burchardi and Hassan (2013) kindly made available. We split our sub-sample of firms in West Germany at the median of one of their district-specific social ties measure, the district-specific share of households with relatives in the East in 1991, and re-estimate our main model variant of column 3 in Table 3 on two different samples.[35] In columns 1 and 2 of Table 8, we estimate on a sample with all East German firms and the West German firms in districts with a social ties level below the median. In column 3 and 4, we use a sample that

[34] The reduced form estimate is 127% in case of the 40%-quantile, and 82% in case of the 80%-quantile.

[35] We find very similar estimation results if we use the alternative social ties measure of Burchardi and Hassan (2013), the share of Soviet expellees in 1961, or if we use their excluded instrument, the share of housing destroyed in 1946.

covers again all East German firms, but now the West German firms in districts with a social ties level at or above the median. In both our alternative samples, the estimates on the coefficients of main interest remain stable and well significant.

Finally, let us note that we didn't find any indication for our empirical results to depend on a few outliers in our main estimation sample (see Appendix Table C5). We can, for example, show that we can drop all new firms with more than 20 employees, or all start-ups with 1 employee. In addition, we can exclude specific 2-digit industries one-by-one, or specific districts.

4.4 Legal form changes

Our empirical findings up to now suggest that more entrants are deterred from incorporating at market entry in East than in West Germany, and this pattern is in line with relatively higher CLL-related entry cost in East than in West Germany. If such entrants are more likely to incorporate later on, we should find a corresponding East-West difference in data on legal form changes.

Using data on all legal form decisions involving young firms in our sample that incorporate before December 31st, 1999, or vice versa, we can provide descriptive statistics on legal form choices at market entry and also on subsequent legal form changes during the first six years after entry. In the first row of Panel A of Table 9, we focus on all the firms that are registered as incorporation at the end of the sixth year after entry. 95.5% of these firms started already as CLL at market entry. In the first year after entry, an additional 1.57% switched from their initially chosen, unincorporated form into CLL, and so on. In Panel B of Table 9, we focus on all the firms that are registered as unincorporated firms at the end of the sixth year after entry and calculate shares as in Panel A.

According to Table 9, legal form changes are no common events in Germany during the 1990s. Specifically, the incidence of legal form changes in our data is only about one third of the incidence reported in recent studies for the US (Levine and Rubinstein, 2016, Cole and Sokolyk, 2015). This difference between Germany and the US reflects the fact that legal form changes involving incorporation decisions are rather expensive in Germany. Apart from that difference in levels, the German pattern in the first rows of Panels A and B compares very well with evidence for the US. Similar to US incorporations, German incorporations unincorporate 7 times less often during the first six years after entry than other German firms change their legal form in the opposite direction during the same time period.[36] Moreover, the incidence of legal form changes decreases across time after market entry, fitting well with the pattern that is typically reported for the US.

[36]The corresponding US estimate indicates 8 times less often (Levine and Rubinstein, 2016).

In the second and third row of Panel A we compare the incidence of incorporating for East and West Germany separately. While 96.59% of the Western firms that are incorporated at the end of the sixth year started already as CLL, the corresponding Eastern share is 2.44 percentage points lower. Accordingly, the incidence of incorporating after entry is higher in the region with relatively higher CLL-related entry cost ($p < 0.01$). When moving to the right in Table 9, we see that this East-West difference is, in particular, driven by firms that changed their legal form in the second or third year after entry. In the first year after entry, the incidence of incorporating is not significantly different between the two entry cost regions. From the fourth year on, there are again no significant differences.

Taken together, the pattern of legal form changes in the East and the West, and how these compare, match our expectation. We observe East-West differences in the incidence of incorporating that are consistent with higher CLL-related entry cost deterring more entrepreneurs from incorporating at market entry and these entrants being more likely to revise their initial legal form choice later on. The fact that the difference in switching activity is strongest in the second and third year after entry fits as well with relatively higher CLL-related entry cost in the East than in the West, and with legal form changes involving incorporations generally being costly in Germany.

5 Conclusion

In this paper, we study influences on job creation in new firms that follow from incorporation law. Adapting the theoretical model of Lucas (1978) to integrate incorporation decisions of entrepreneurs, we derive predictions on influences of incorporation-related entry cost. To empirically identify these influences, we use an instrumental variables approach to allow for entrepreneurs' decisions on incorporation that can be endogenous in equations explaining job creation at firm entry. We integrate that into a cross-sectional difference-in-differences model to take advantage of the natural experiment in regulation that accompanied German reunification. We provide a test of the plausibility of our identification strategy and a set of extensions to our main empirical analysis.

The pattern of our empirical findings are in line with the theoretical predictions and suggest the following. At first, an increase in incorporation-related entry cost affects the composition of a population of new firms by lowering the probability of incorporating relative to starting a firm in an unincorporated form. Investigating, in addition, entrepreneurs' decisions on legal form changes after market entry indicates that entrepreneurs that were deterred from incorporating at market entry are more likely to revise their initial legal form choice later on. It takes, however, several

years until such revisions get implemented to a non-negligible extent.

Most importantly, our main findings indicate that an increase in incorporation-related entry cost leads to an increasing difference in initial job creation between incorporated and unincorporated entrants. Estimates of linear probability models, and of quantile regression models, show furthermore that an increase in incorporation-related entry cost decreases the mass of the entry size distribution across incorporations relative to the one across the unincorporated firms in the intermediate range.

Taken together, we document how incorporation law can influence the extent to which new incorporations emerge at market entry in a population of new firms. We also document the extent to which initial job creation accrues in incorporated or unincorporated entrants. Such law-induced entry effects at the extensive and intensive margin are important as incorporated firms are endowed with law-induced advantages that are likely to foster future job creation and firm growth. Our findings point towards a potential for long-lasting influences on job creation, and on the economy-wide distribution of jobs across firms.

The estimates that we report are also of interest insofar as they are likely to qualify as unintended effects. Specifically, our empirical context is such that we observe, as expected, the strongest restrictions to private entrepreneurial activity in the regime with high incorporation-related entry cost. And this regime happens to represent an economic environment with substantial, shock-induced structural change where private entrepreneurial activity is most needed.

References

Angrist, Joshua D. and Alan B. Krueger, "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 2001, *15* (4).

Becht, Marco, Colin Mayer, and Hannes F Wagner, "Where do firms incorporate? Deregulation and the cost of entry," *Journal of Corporate Finance*, 2008, *14* (3), 241–256.

Blundell, Richard and Monica Costa Dias, "Alternative Approaches to Evaluation in Empirical Microeconomics," *Journal of Human Resources*, 2009, *44* (3), 565–640.

Braguinsky, Serguey, Lee G. Branstetter, and Andre Regateiro, "The Incredible Shrinking Portuguese Firm," Working Paper 17265, National Bureau of Economic Research July 2011.

Branstetter, Lee, Francisco Lima, Lowell J. Taylor, and Ana Venâncio, "Do Entry Regulations Deter Entrepreneurship and Job Creation? Evidence from Recent Reforms in Portugal," *Economic Journal*, 2014, *124* (577), 805–832.

Bruhn, Miriam, "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico," *Review of Economics and Statistics*, 2011, *93* (1), 382–386.

Bundesforschungsanstalt für Landeskunde und Raumordnung, BfLR, *Materialien zur Raumentwicklung. Laufende Raumbeobachtung. Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden 1992/93.* 1995. Heft 67.

Burchardi, Konrad B. and Tarek A. Hassan, "The Economic Impact of Social Ties: Evidence from German Reunification," *Quarterly Journal of Economics*, 2013, *128* (3), 1219–1271.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics using Stata*, revised ed., Stata Press, 2010.

Cole, Rebel A. and Tatyana Sokolyk, "How Do Firms Choose Legal Form of Organization?," SSRN Working Paper 1682263 December 2015.

Dahl, Michael S. and Olav Sorenson, "The embedded entrepreneur," *European Management Review*, 2009, *6* (3), 172–181.

_ **and** _ , "Home Sweet Home: Entrepreneurs' Location Choices and the Performance of Their Ventures," *Management Science*, 2012, *58* (6), 1059–1071.

Dammann, Jens and Matthias Schündeln, "The Incorporation Choices of Privately Held Corporations," *Journal of Law, Economics, and Organization*, 2011, *27* (1), 79–112.

_ **and** _ , "Where Are Limited Liability Companies Formed? An Empirical Analysis," *Journal of Law and Economics*, 2012, *55* (4), 741–791.

Deutsche Bundesbank, "Zur Entwicklung der privaten Vermögenssituation seit Beginn der neunziger Jahre," *Monatsbericht der Deutschen Bundesbank*, 1999, *51*, 33–50.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "The Regulation of Entry," *Quarterly Journal of Economics*, 2002, *117* (1), 1–37.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 2016, *106* (11).

Gelbach, Jonah B., "When Do Covariates Matter? And Which Ones, and How Much?," *Journal of Labor Economics*, 2016, *34* (2), 509–543.

Harhoff, Dietmar, Konrad Stahl, and Michael Woywode, "Legal Form, Growth and Exit of West German Firms–Empirical Results for Manufacturing, Construction, Trade and Service Industries," *Journal of Industrial Economics*, 1998, *46* (4), 453–88.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, "Can Unemployment Insurance Spur Entrepreneurial Activity? Evidence from France.," mimeo July 2016.

Hunt, Jennifer, "Staunchung Emigration from East Germany: Age and the Determinants of Migration," *Journal of the European Economic Association*, 2006, *4* (5), 1014–1037.

Imbens, Guido W. and Jeffrey M. Wooldridge, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 2009, *47* (1).

Kaplan, David S., Eduardo Piedra, and Enrique Seira, "Entry regulation and business start-ups: Evidence from Mexico," *Journal of Public Economics*, 2011, *95* (11), 1501–1515.

Koenker, Roger, *Quantile Regression* Econometric Society Monographs, New York: Cambridge University Press, 2005.

Lechner, Michael, "The Estimation of Causal Effects by Difference-in-Difference Methods," *Foundations and Trends in Econometrics*, 2011, *4* (3), 165–224.

Levine, Ross and Yona Rubinstein, "Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?," *Quarterly Journal of Economics*, 2016. forthcoming.

Lucas, Robert E., "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 1978, *9* (2), 508–523.

Machado, José A. F. and J. M. C. Santos Silva, "Quantiles for Counts," *Journal of the American Statistical Association*, 2005, *100* (472), 1226–1237.

Michelacci, Claudio and Olmo Silva, "Why so many local entrepreneurs?," *Review of Economics and Statistics*, 2007, *89* (4), 615–633.

Niederleithinger, Ernst, "Probleme der Ausdehnung des GmbH-Rechts auf die neuen Bundesländer," *GmbHRundschau*, 1992, pp. 220–224.

Prantl, Susanne, "The impact of firm entry regulation on long-living entrants," *Small Business Economics*, 2012, *39* (1), 61–76.

_ , "Exit Decisions of Young Firms," mimeo, University of Cologne 2017.

_ **and Alexandra Spitz-Oener**, "How does entry regulation influence entry into self-employment and occupational mobility?," *Economics of Transition*, 2009, *17* (4), 769–802.

_ **and** _ , "Interacting Product and Labor Market Regulation and the Impact of Immigration on Native Wages," IZA Discussion Papers 7882, Institute for the Study of Labor (IZA) January 2014.

Statistisches Bundesamt, *Vermögensbestände und Schulden privater Haushalte* 1993. Einkommens- und Verbraucherstichprobe, Fachserie 15, Heft 2.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, 3 ed., Cambridge, MA: MIT Press, 2010.

World Bank, "Doing Business 2016: Measuring Regulatory Quality and Efficiency," 2015.

Figures

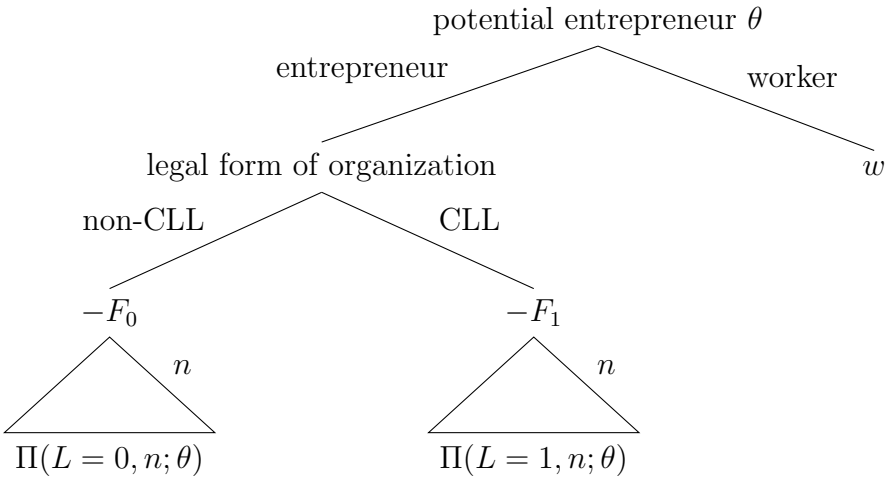


Figure 1: Sequence of decisions a potential entrepreneur θ faces

Tables

Table 1: Definitions of variables and descriptive statistics

Variable	Definition	Mean	Std. Dev.
S_i : firm size	number of employees (incl. working entrepreneurs) at market entry	5.652	19.167
E_d : East Germany	1: firm location in East Germany; 0: otherwise	0.480	
CLL_i : company with limited liability	1: company under the German Law on Companies with Limited Liability (GmbHG); 0: otherwise	0.365	
affiliated firm	1: firm is partly or fully owned by at least one firm; 0: otherwise	0.089	0.284
team size	number of entrepreneurs (see notes)	1.338	0.615
age	age of main entrepreneur (see notes)	38.051	9.584
low education	1: main entrepreneur (see notes) with no vocational training degree; 0: otherwise	0.033	
medium education	1: main entrepreneur (see notes) with vocational training degree; 0: otherwise	0.483	
high education	1: main entrepreneur (see notes) with degree from university; 0: otherwise	0.173	
share working age	district-specific share of people aged 20 to 65 in 1993	0.627	0.014
share manufacturing workers	district-specific share of manufacturing workers in the population in 1993	0.092	0.040
population density	district-specific population density (measured as number of citizens per km^2) in 1993	304.664	400.720

Note: This table provides weighted descriptive statistics for the main sample of 11,387 firms as well as descriptive statistics on district characteristics that we use as controls. The latter are based on data from the BBSR (see Appendix B.2 for further details).
As entrepreneurs we consider all managing owners. The variables age and education refer to the managing owner with the highest equity share if the firm has more than one managing owner. In cases with missing share data and multiple managing owners the entrepreneur-specific information of the oldest entrepreneur has been chosen instead. Information on the entrepreneurs' education is missing for 3,373 firms.

Table 2: Incidence of starting a CLL under different CLL-related entry cost

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Dependent variable:			CLL_i		
Constant	0.3891***				
	(0.0120)				
E_d	-0.0497***	-0.0913***	-0.0933***	-0.1004***	-0.0298
	(0.0155)	(0.0112)	(0.0127)	(0.0135)	(0.0217)
Firm & Entrep. Controls	No	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
District Controls	No	No	Yes	Yes	Yes
R^2_{adj}	0.003	0.590	0.591	0.505	0.923
N	11387	11387	11387	10319	1068

Note: The table shows ordinary least squares estimates of variants of regression model (1). Columns 1 to 3 are based on the main sample of 11,387 entrants. Column 4 focuses on the restricted sample of firms without any affiliation to a parent company, column 5 on the opposite sample of affiliated firms.

The dependent variable, CLL_i, is an indicator that takes the value one for entrants starting a firm under the German limited liability company law. While column 1 includes the East indicator, E_d, as the only explanatory variable apart from the intercept, in column 2 firm and entrepreneurial controls as well as industry and year fixed effects are added. Columns 3 to 5 additionally including district-specific controls show the estimates of the model formalized in equation (1). All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 3: Firm size at market entry under different CLL-related entry cost

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Dependent variable:		$\ln(\mathsf{S}_i)$		
E_d	0.3119***	0.2308***		
	(0.0256)	(0.0232)		
CLL_i	0.6995***	0.5006***	0.4981***	0.9962***
	(0.0279)	(0.0260)	(0.0257)	(0.2318)
$\mathsf{E}_d \times \mathsf{CLL}_i$	0.6241***	0.4452***	0.4406***	0.9192***
	(0.0412)	(0.0377)	(0.0372)	(0.1096)
Firm & Entrep. Controls	No	Yes	Yes	Yes
District Controls	No	Yes	No	No
Industry FE	No	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes
District FE	No	No	Yes	Yes
R^2_{adj}	0.312	0.401	0.405	0.316
N	11387	11387	11387	11387
Kleibergen-Paap F				31.919

Note: Columns 1 to 3 provide ordinary least squares estimates of variants of our main regression model formalized in equation (2) for the main sample of 11,387 entrants. In column 4, we provide instrumental variable estimates.

The dependent variable, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. Column 1 shows the estimates of a regression of logarithmic entry size on the indicator for firms starting as limited liability companies, CLL_i, the indicator for firms entering East Germany, E_d, as well as their interaction, $\mathsf{E}_d \times \mathsf{CLL}_i$. In column 2 firm and entrepreneurial controls as well as district controls and industry and year fixed effects are added. Columns 3 and 4 show the estimates coming from a regression identical to equation (2) including district fixed effects. The excluded instruments in column 4 are the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$, and its interaction with the indicator for East Germany as explained in detail in Section 4.2. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 4: Firm size at market entry—first stages and reduced form

Dependent variable:	(1) IV, 1st-stage CLL_i	(2) IV, 1st-stage $\mathsf{E}_d \times \mathsf{CLL}_i$	(3) IV, red. form $\ln(\mathsf{S}_i)$
$\mathsf{I}_{jd,\ pre\text{-}reunific.}$	0.4446***	-0.1900***	0.2683**
	(0.0552)	(0.0324)	(0.1058)
$\mathsf{E}_d \times \mathsf{I}_{jd,\ pre\text{-}reunific.}$	-0.1025**	0.7720***	0.6075***
	(0.0417)	(0.0315)	(0.0818)
Firm & Entrep. Controls	Yes	Yes	Yes
District FE, Industry FE, Year FE	Yes	Yes	Yes
R^2_{adj}	0.366	0.425	0.327
N	11387	11387	11387
Angrist-Pischke F	65.808	550.947	
Sanderson-Windmeijer F	63.645	306.940	
F of excl. instr.	32.821	303.622	

Note: Columns 1 and 2 show the first stage regression estimates of the instrumental variables approach introduced in Section 4.2 for the full sample of 11,387 entrants. Column 3 depicts the estimates of the corresponding reduced form.

The dependent variable in column 1, CLL_i, is an indicator for entrants starting a firm under the German limited liability company law. In column 2, the dependent variable, $\mathsf{E}_d \times \mathsf{CLL}_i$, is the interaction of CLL_i with the indicator for entry in East Germany. The dependent variable in column 3, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. $\mathsf{I}_{jd,\ pre\text{-}reunific.}$ is the predicted industry- and region-specific incidence of CLLs before reunification. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 5: Firm size at market entry in a homogeneous group of start-ups

	(1) OLS	(2) IV	(3) OLS
Dependent variable:		$\ln(\mathsf{S}_i)$	
CLL_i	0.4981***	0.9878***	0.7533***
	(0.0254)	(0.2177)	(0.2072)
$\mathsf{E}_d \times \mathsf{CLL}_i$	0.4235***	1.0317***	0.1144
	(0.0376)	(0.1291)	(0.2480)
Firm & Entrep. Controls	Yes	Yes	Yes
District FE, Industry FE, Year FE	Yes	Yes	Yes
R^2_{adj}	0.375	0.247	0.297
N	10319	10319	1068
Kleibergen-Paap F		30.312	

Note: Columns 1 and 3 provide ordinary least squares estimates of our main regression model formalized in equation (2). In column 2, we provide instrumental variable estimates. While columns 1 and 2 use the restricted sample of non-affiliated firms, column 3 depicts the estimates in the opposite sample of all affiliated firms.

The dependent variable, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. The explanatory variable, CLL_i, is an indicator for whether the entrepreneur starts his firm under a legal form with limited liability. The interaction, $\mathsf{E}_d \times \mathsf{CLL}_i$, takes one for all entrants that start a limited liability company in East Germany. The excluded instruments in column 2 are the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ pre\text{-}reunific.}$, and its interaction with the indicator for East Germany as explained in detail in Section 4.2. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 6: Incidence of starting in a specific entry size category under different CLL-related entry cost

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Dependent variable:	$1\{1\ \mathsf{empl}_i\}$	$1\{1\ \mathsf{empl}_i\}$	$1\{2\text{-}3\ \mathsf{empl}_i\}$	$1\{2\text{-}3\ \mathsf{empl}_i\}$	$1\{\geq 4\ \mathsf{empl}_i\}$	$1\{\geq 4\ \mathsf{empl}_i\}$
CLL_i	-0.2534***	-0.5795***	0.0260	0.2790*	0.2274***	0.3005***
	(0.0128)	(0.1360)	(0.0161)	(0.1507)	(0.0138)	(0.1123)
$\mathsf{E}_d \times \mathsf{CLL}_i$	0.0574***	-0.1579***	-0.2518***	-0.2314***	0.1943***	0.3893***
	(0.0164)	(0.0578)	(0.0206)	(0.0613)	(0.0180)	(0.0509)
Firm & Entrep. Controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE, Industry FE, Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2_{adj}	0.234	0.104	0.061	0.017	0.309	0.282
N	11387	11387	11387	11387	11387	11387
Kleibergen-Paap F		31.919		31.919		31.919
Mean dep. var.	0.337	0.337	0.354	0.354	0.308	0.308

Note: The table shows ordinary least squares (columns 1, 3, and 5) and instrumental variable (columns 2, 4, and 6) linear probability estimates for the full sample of 11,387 entrants.

The dependent variables $1\{1\ \mathsf{empl}_i\}$, $1\{2\text{-}3\ \mathsf{empl}_i\}$, and $1\{\geq 4\ \mathsf{empl}_i\}$ are indicators for 1-person firms, 2-to-3-person firms and firms with an entry size of 4 or more employees including the working entrepreneurs. The explanatory variable, CLL_i, is an indicator for whether the entrepreneur starts his firm under a legal form with limited liability. The interaction, $\mathsf{E}_d \times \mathsf{CLL}_i$, takes one for all entrants that start a limited liability company in East Germany. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 7: Heterogenous effects across quantiles of the entry size distribution

Dependent variable:	(1) $Q_{\ln(\mathsf{S}_i)}(0.30)$	(2) $Q_{\ln(\mathsf{S}_i)}(0.40)$	(3) $Q_{\ln(\mathsf{S}_i)}(0.50)$	(4) $Q_{\ln(\mathsf{S}_i)}(0.60)$	(5) $Q_{\ln(\mathsf{S}_i)}(0.70)$	(6) $Q_{\ln(\mathsf{S}_i)}(0.80)$	(7) $Q_{\ln(\mathsf{S}_i)}(0.90)$
Panel A							
CLL_i	0.2695*	0.3782***	0.5082***	0.5477***	0.5797***	0.7087***	0.8686***
	(0.1436)	(0.0226)	(0.0193)	(0.0221)	(0.0321)	(0.0384)	(0.0409)
$\mathsf{E}_d \times \mathsf{CLL}_i$	0.8291***	0.6988***	0.4717***	0.4575***	0.4724***	0.3543***	0.1698***
	(0.1159)	(0.0363)	(0.0354)	(0.0372)	(0.0387)	(0.0509)	(0.0618)
Panel B:							
$\mathsf{I}_{jd,\ \textit{pre-reunific.}}$	0.0051	0.1297**	0.1808**	0.2115**	0.3190***	0.4011***	0.6966***
	(0.0458)	(0.0582)	(0.0762)	(0.0868)	(0.1034)	(0.1272)	(0.1564)
$\mathsf{E}_d \times \mathsf{I}_{jd,\ \textit{pre-reunific.}}$	0.6798***	0.8183***	0.7314***	0.7544***	0.6749***	0.6006***	0.5137***
	(0.0950)	(0.0777)	(0.0828)	(0.0812)	(0.0945)	(0.1030)	(0.1112)
Firm & Entrep. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE, Industry FE, Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11387	11387	11387	11387	11387	11387	11387

Note: The table reports quantile regression estimates of the main model for the full sample of 11,387 entrants.
The dependent variables in columns 1 to 7 are the 30%-, 40%-, 50%-, 60%-, 70%-, 80%-, and 90% quantiles of $\ln(\mathsf{S}_i)$, respectively. While Panel A shows the results from the model formalized in equation (2), Panel B depicts the quantile regression estimates from the reduced form where the legal form indicator, CLL_i, is replaced by the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$.
Sampling weights are used as explained in detail in Appendix B.2. Robust standard errors are reported in parentheses. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 8: Firm size at market entry in a group of start-ups with strong social ties to the East

	(1) OLS	(2) OLS	(3) IV	(4) IV
Dependent variable:		$\ln(\mathsf{S}_i)$		
CLL_i	0.4928***	0.4896***	1.3998***	1.0285***
	(0.0362)	(0.0375)	(0.2838)	(0.2804)
$\mathsf{E}_d \times \mathsf{CLL}_i$	0.4151***	0.4352***	0.8657***	0.9833***
	(0.0452)	(0.0465)	(0.1345)	(0.1343)
Firm & Entrep. Controls	Yes	Yes	Yes	Yes
District FE, Industry FE, Year FE	Yes	Yes	Yes	Yes
R^2_{adj}	0.433	0.415	0.233	0.295
N	8288	8225	8288	8225
Kleibergen-Paap F			27.803	23.921

Note: The table reports estimates of the main model formalized in equation (2). Columns 1 and 2 show ordinary least squares estimates, columns 3 and 4 depict the results of the instrumental variable approach discussed in Section 4.2. The sub-samples in columns 1 and 3 consist of all firms from the main sample that enter in East Germany or in a Western district with weak, i.e. below median, social ties to the East. The sub-samples in columns 2 and 4 consist of all firms that enter in East Germany or in a Western district with strong social ties to East Germany. Social ties are measured in West Germany by the district-specific share of households in 1991 with relatives in East Germany (Burchardi and Hassan, 2013).

The dependent variable, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. The explanatory variable, CLL_i, is an indicator for whether the entrepreneur starts his firm under a legal form with limited liability. The interaction, $\mathsf{E}_d \times \mathsf{CLL}_i$, takes one for all entrants that start a limited liability company in East Germany. The excluded instruments in columns 3 and 4 are the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ pre\text{-}reunific.}$, and its interaction with the indicator for East Germany as explained in detail in Section 4.2. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table 9: Legal form decisions from market entry to the sixth year after entry

Panel A:							
	Years until incorporating:						
	0	1	2	3	4	5	6
Germany	95.50%	1.57%	1.13%	0.75%	0.59%	0.27%	0.20%
East	94.16%	1.69%	1.65%	1.30%	0.77%	0.23%	0.20%
West	96.59%	1.47%	0.70%	0.29%	0.44%	0.30%	0.20%
Δ_d	−2.44%***	0.23%	0.95%***	1.01%***	0.32%	−0.07%	0.00%
t-stat.	−3.55	0.59	2.68	3.35	1.21	−0.40	0.02
Panel B:							
	Years until unincorporating:						
	0	1	2	3	4	5	6
Germany	99.39%	0.12%	0.16%	0.07%	0.07%	0.09%	0.10%

Note: The table summarizes how many years firms that are incorporated (Panel A) or unincorporated (Panel B) six years after entry were active under the respective other legal form type before they switched the legal form of organization. Panel A includes all 4,894 (2,770 Western and 2,124 Eastern) firms that are incorporated at the end of the sixth year after entry or that exit the market as incorporation before the end of the sixth year. Panel B contains all 6,477 (3,207 Western and 3,270 Eastern) firms that are unincorporated six years after entry and those that exit the market as unincorporated firm with less than six years market activity. 16 firms from the main sample are excluded due to inconsistent legal form change information.

While the first rows of Panel A and B show the corresponding fractions for all entrants, rows 2 and 3 of Panel A depict the incidence of legal form changes to a CLL for the East and West separately. The difference in the share of legal form changes, Δ_d, is reported in the forth row and the t-statistic testing the equality of means is reported in the last row of Panel A.

Sampling weights are used as explained in detail in Appendix B.2. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

A Theory-related Appendix

In this Appendix, we extend the discussion of our theoretical model introduced in Section 2. To avoid redundancy, we build on the assumptions stated in the main part and start with the derivation of the equilibrium.

To derive the decision of the rational potential entrepreneur, the sequential game is solved backwards. The potential entrepreneur forms expectations about his profits in the respective entry modes if he decides to become an entrepreneur. More specifically, having entered with legal form L the potential entrepreneur θ decides on the labor input solving the following maximization problem:

$$\max_n E_{\Xi_L}[\Pi(L,n;\theta)] \quad = \quad \max_n E_{\Xi_L}[\xi_L] \underbrace{\pi(n;\theta)}_{=\theta n^\alpha - wn} \tag{MAX}$$

The solution to this maximization problem is summarized in the following lemma.

Lemma A1. *Both* $n^* := \arg\max_n E_{\Xi_L}[\Pi(L,n;\theta)]$ *and* $E_{\Xi_L}[\Pi^*] := \max_n E_{\Xi_L}[\Pi(L,n;\theta)]$ *increase strictly in* θ*.*

Proof. Since $f(n)$ is strictly concave increasing and $E_{\Xi_L}[\xi_L] > 0$ for $L \in \{0,1\}$ the first order condition

$$n^* = \left(\frac{\alpha\theta}{w}\right)^{1/(1-\alpha)} \tag{FOC}$$

is necessary and sufficient. Plugging n^* into the objective function yields

$$E_{\Xi_L}[\Pi^*] = E_{\Xi_L}[\xi_L] \underbrace{\left(\frac{\alpha}{w}\right)^{\alpha/(1-\alpha)} (1-\alpha)\theta^{1/(1-\alpha)}}_{=:\,\pi(n^*;\theta)} \tag{3}$$

□

So, as in the standard Lucas (1978) model, the distribution of firm sizes results immediately from differences in the type of potential entrepreneurs in the population. But it is not only the optimal firm size that is determined by the type, specific entrepreneurs select the legal form dependent on their type, too. The mechanism becomes clearer when scrutinizing the entrepreneur's decision on the legal form.

As already explained, a potential entrepreneur will decide to form an unincorporated firm if the expected profit of starting an unincorporated firm is larger than the expected profit under either of the outside options, i.e. becoming a dependent worker (entrepreneurship condition, E_0) or incorporating (legal form condition, LF_0).

$$E_{\Xi_0}[\Pi(0,n^*;\theta)] - F_0 \geq w \tag{E_0}$$

$$E_{\Xi_0}[\Pi(0,n^*;\theta)] - F_0 > E_{\Xi_1}[\Pi(1,n^*;\theta)] - F_1 \tag{LF_0}$$

If, in contrast, the expected profit of incorporating is larger than both the wage and the expected profit from setting up an unincorporated firm, the same entrepreneur starts a CLL:

$$E_{\Xi_1}[\Pi(1, n^*; \theta)] - F_1 \geq w \qquad (E_1)$$

$$E_{\Xi_1}[\Pi(1, n^*; \theta)] - F_1 \geq E_{\Xi_0}[\Pi(0, n^*; \theta)] - F_0 \qquad (LF_1)$$

Finally, if neither (E_0) nor (E_1) hold, the potential entrepreneur prefers to be employed as dependent worker.

Both the entrepreneurship and legal form conditions are illustrated in Figure A1(a). The coordinate system spans the space of expected profits in the two legal form regimes. While the horizontal axis depicts the expected profit of a potential entrepreneur starting a firm without limited liability, the vertical axis reflects the corresponding profit with limited liability. According to equations (E_0) to (LF_1), the space is partitioned into regions that determine the entry decision. Entrepreneurs that expect profits in the vertically ruled part of the graph decide to set up a CLL as, here, both (E_1) and (LF_1) are fulfilled, whereas entrepreneurs with profit tuples lying in the horizontally ruled subset prefer to become owners of an unincorporated firm. Finally, entrepreneurs in the lower left rectangle that receive expected profits smaller than the outside option under either of the entry modes decide to work as employees.

Based on these considerations and the optimal input choice, we are now able to derive the firm size distribution in the economy in the two legal form regimes. For this reason, we characterize the entry decision of each entrepreneur who decides to open up a firm. Conditional on becoming an entrepreneur, the decision on the optimal legal form follows from (LF_0) and (LF_1). Which of this two mutually exclusive inequalities holds, depends, in particular, on the difference in expected profits under the two legal forms

$$E_{\Xi_1}[\Pi(1, n^*; \theta)] - E_{\Xi_0}[\Pi(0, n^*; \theta)] = \left(\frac{E_{\Xi_1}[\xi_1]}{E_{\Xi_0}[\xi_0]} - 1 \right) \cdot E_{\Xi_0}[\Pi(0, n^*; \theta)] \quad .^{37} \qquad (4)$$

Equation (4) implies that a profit maximizing potential entrepreneur θ who expects a profit of $E_{\Xi_0}[\Pi(0, n^*; \theta)]$ as the owner of an unincorporated firm expects a profit of $\frac{E_{\Xi_1}[\xi_1]}{E_{\Xi_0}[\xi_0]} \cdot E_{\Xi_0}[\Pi(0, n^*; \theta)]$ when starting his firm under the law for CLLs instead. Translating this consideration to all potential entrepreneurs $\theta \in [0, 1]$, equation (4) describes the set of expected profit tuples that can be achieved in the population of expected profit maximizing potential entrepreneurs. This subset is sketched by the maximum expected profit tuple line $(MEPL)$ in Figure A1(a) which goes through

[37] Note that this equality is a direct implication of equation (3).

the origin and has a slope of $\frac{E_{\Xi_1}[\xi_1]}{E_{\Xi_0}[\xi_0]}$.[38]

Lemma A1 implies that the $MEPL$ diverges from the identity function as θ, n^*, and, conclusively, $E_{\Xi_0}[\Pi(0, n^*; \theta)]$ increase. As a result, points lying further to the upper right of the $MEPL$ correspond to entrepreneurs of larger type who start larger firms conditional on entry.

The point at the intersection of the $MEPL$ and (E_0), for example, reflects the decision of a particular entrepreneur who—when choosing his labor input conditional on his expectations optimally—expects to earn $(w + F_0)$ as the owner of an unincorporated firm while expecting a profit of $\frac{E_{\Xi_1}[\xi_1]}{E_{\Xi_0}[\xi_0]} \cdot (w + F_0)$ from owning a CLL. According to the regime each entrepreneur's maximum expected profit tuple lies in, i.e. which point on the $MEPL$ represents his expected profit under an optimal input choice, he decides on whether to become a worker or an entrepreneur and conditional on entrepreneurial activity on his firm's legal form of organization.

For this reason, it is not only the functional form of the $MEPL$ that is relevant for the distribution of firm sizes in the two legal form regimes but also the weight that is assigned to each point on it. The density function of tuples on the $MEPL$ is a projection from the type space into the expected profit-tuple space given that potential entrepreneurs maximize profits conditional on entry. As it is not clear that all points on the $MEPL$ have a non-zero weight, let us give a sufficient condition for the existence of both CLLs and unincorporated firms.

Lemma A2 (Sufficient Condition)**.** *If (i)* $E_{\Xi_0}[\Pi(0, n^*; 1)] > w+F_0$*, (ii)* $E_{\Xi_1}[\Pi(1, n^*; 1)] - E_{\Xi_0}[\Pi(0, n^*; 1)] > F_1 - F_0$ *and (iii)* $\frac{E_{\Xi_1}[\xi_1]}{E_{\Xi_0}[\xi_0]} < \frac{w+F_1}{w+F_0}$*, then workers and entrepreneurs of both unincorporated firms and CLLs exist.*

Proof. It follows directly from (i) and (ii) that entrepreneur $\theta = 1$ will decide to start a CLL. Moreover, entrepreneur $\theta = 0$ prefers to become an employee as $w > 0$ and $F_0 > 0$.

Entrepreneur $\hat{\theta} \in]0, 1[$ who is characterized by $E_{\Xi_0}[\Pi(0, n^*; \hat{\theta})] = w + F_0$ exists as the left-hand side is continuously increasing in θ (see Lemma A1) and starts an unincorporated firm as (E_0) is obviously fulfilled and (LF_0) holds due to

$$\begin{aligned} & \tfrac{w+F_1}{w+F_0} && > \tfrac{E_{\Xi_1}[\xi_1]}{E_{\Xi_0}[\xi_0]} \\ \Leftrightarrow \quad & 1 + \tfrac{F_1 - F_0}{E_{\Xi_0}[\Pi(0, n^*; \hat{\theta})]} && > \tfrac{E_{\Xi_1}[\xi_1]\pi(n^*; \hat{\theta})}{E_{\Xi_0}[\xi_0]\pi(n^*; \hat{\theta})} \\ \Leftrightarrow \quad & E_{\Xi_0}[\Pi(0, n^*; \hat{\theta})] - F_0 && > E_{\Xi_1}[\Pi(1, n^*; \hat{\theta})] - F_1 \end{aligned}$$

where the first equivalence follows from the definition of $\hat{\theta}$ and the second from

[38] The expected profit tuple in the origin corresponds to the entrepreneur of type $\theta = 0$. He maximizes his profit by not hiring any worker and receives a zero profit in either of the two legal forms if he decides to become an entrepreneur.

equation (3). □

The intuition for Lemma A2 is as follows. While the first inequality states that the entrepreneur of the highest type finds it profitable to start a firm, the second inequality complements as it requires that this potential entrepreneur with the largest firm, who also faces the highest negative income risk in case of firm failure, prefers to cover against this risk. Finally, condition (iii) might look cumbersome on the first glance but has an intuitive interpretation. It says that—given that any potential entrepreneur decides to set up a firm—unincorporated firms can only exist if the relative benefit of forming a CLL due to the reduced risk is smaller than the relative loss associated with the higher entry cost. Note that the existence of workers is a direct implication of the standard Lucas framework as the potential entrepreneur of the lowest type always prefers earning a positive wage over becoming an entrepreneur.

Having discussed the sufficient condition for the existence of CLLs and unincorporated firms, let us now turn our view to firms' characteristics. As, according to Lemma A1, the expected gain from limited liability (described in equation (4)) is strictly increasing in θ, the equilibrium decision of the potential entrepreneur is characterized by thresholds.

Lemma A3 (Threshold equilibrium)**.** *Suppose the sufficient condition in Lemma A2 holds. Then* θ_0^* *and* θ_1^* *with* $0 < \theta_0^* < \theta_1^* < 1$ *exist such that in equilibrium potential entrepreneurs with* $\theta \in [0, \theta_0^*[$ *become workers, those characterized by* $\theta \in [\theta_0^*, \theta_1^*[$ *form unincorporated firms, and entrepreneurs with* $\theta \in [\theta_1^*, 1]$ *form CLLs. The equilibrium labor inputs* $n^*(\theta)$ *are defined in* (FOC).

Proof. An entrepreneur θ expects (conditional on entry) a profit of $E_{\Xi_0}[\Pi(0, n^*; \theta)]$ under unlimited liability and $E_{\Xi_1}[\Pi(1, n^*; \theta)]$ under limited liability.
As equation (3) implies that $\frac{d}{d\theta} E_{\Xi_1}[\Pi(1, n^*; \theta)] > \frac{d}{d\theta} E_{\Xi_0}[\Pi(0, n^*; \theta)] > 0$ and that both derivatives are continuous in θ, there exists θ_1^* such that for all $\theta < \theta_1^*$ (LF_0) holds while for all $\theta \geq \theta_1^*$ (LF_1) is true.
Moreover, as $\frac{d}{d\theta} E_{\Xi_0}[\Pi(0, n^*; \theta)] > 0$ there exists θ_0^* such that for all $\theta \geq \theta_0^*$ (E_0) holds. Furthermore, the last step in the proof of Lemma A2 implies that if (E_0) is true then (LF_0) holds as well and the intersection is non-empty.
To show that the described threshold equilibrium exists it suffices to show that (LF_1) implies (E_1). Consider $\hat{\theta}$ with $E_{\Xi_1}[\Pi(1, n^*; \hat{\theta})] - E_{\Xi_0}[\Pi(0, n^*; \hat{\theta})] = F_1 - F_0$, i.e. the entrepreneur with the smallest θ such that (LF_1) holds.
As $\hat{\theta} > \theta_0^*$, (E_0) is true for $\hat{\theta}$, i.e. $E_{\Xi_0}[\Pi(0, n^*; \hat{\theta})] > w + F_0$. But this also implies that $E_{\Xi_1}[\Pi(1, n^*; \hat{\theta})] = E_{\Xi_0}[\Pi(0, n^*; \hat{\theta})] + F_1 - F_0 > w + F_1$ which is sufficient for (E_1). □

Note that the proof implies that (E_0) and (LF_1) are the binding constraints if entrants under both legal forms exist. Based on this observation we can give an explicit definition of the equilibrium thresholds θ_0^* and θ_1^*.

Corollary A1. *θ_0^* is characterized by*

$$\begin{aligned} & E_{\Xi_0}[\Pi(0, n^*; \theta)] && = w + F_0 \\ \Leftrightarrow \quad & \pi(n^*; \theta) && = \tfrac{w+F_0}{E_{\Xi_0}[\xi_0]} \\ \Leftrightarrow \quad & \theta_0^* && = \left(\tfrac{w+F_0}{E_{\Xi_0}[\xi_0]}\right)^{1-\alpha} \left(\tfrac{w}{\alpha}\right)^{\alpha} \left(\tfrac{1}{1-\alpha}\right)^{1-\alpha} \end{aligned}$$

and θ_1^* by

$$\begin{aligned} & E_{\Xi_0}[\Pi(0, n^*; \theta)] - F_0 && = E_{\Xi_1}[\Pi(1, n^*; \theta)] - F_1 \\ \Leftrightarrow \quad & \pi(n^*; \theta) && = \tfrac{F_1-F_0}{E_{\Xi_1}[\xi_1]-E_{\Xi_0}[\xi_0]} \\ \Leftrightarrow \quad & \theta_1^* && = \left(\tfrac{F_1-F_0}{E_{\Xi_1}[\xi_1]-E_{\Xi_0}[\xi_0]}\right)^{1-\alpha} \left(\tfrac{w}{\alpha}\right)^{\alpha} \left(\tfrac{1}{1-\alpha}\right)^{1-\alpha} \end{aligned}$$

Having derived the equilibrium thresholds, we can analyze how an increase in the CLL entry costs, F_1, affects the composition of entrants in the two modes of entry. The formal change of (LF_0), (LF_1), and (E_1) resulting from an increase in F_1 is illustrated when moving from Figure A1(a) to A1(b). While the $MEPL$ is unaffected by the fixed cost change, the vertically ruled part of the graph—indicating the expecting profit tuples leading to a choice in favor of a CLL—shrinks. This directly implies a reduction in the share of potential entrants choosing a CLL conditional on entry. We summarize this observation in the following proposition which is closely related to hypothesis 1.

Proposition 1. *Consider an equilibrium in which there exist workers and both types of entrepreneurs given fixed cost (F_0, F_1). An increase in the entry cost for CLLs, F_1, decreases the share of CLLs and increases the share of unincorporated firms conditional on entry.*

Proof. It follows from the characterization in Corollary A1 that the increase in F_1 has a positive effect on θ_1^* only.[39] Conclusively, $Pr\{\text{non-CLL}|\text{CLL} \vee \text{non-CLL}\} = (\theta_1^* - \theta_0^*)/(1 - \theta_0^*)$ increases while $Pr\{\text{CLL}|\text{CLL} \vee \text{non-CLL}\} = (1 - \theta_1^*)/(1 - \theta_0^*)$ decreases. □

Our second set of hypotheses focused on the composition of entrants with regard to the size. In our empirical analysis, we identify the effect of entry cost by comparing differences in the legal form-specific entry decisions between the two cost regions.

[39] The first-order derivative is $\frac{\partial \theta_1^*}{\partial F_1} = \frac{1-\alpha}{E_{\Xi_1}[\xi_1]-E_{\Xi_0}[\xi_0]} \left(\frac{F_1-F_0}{E_{\Xi_1}[\xi_1]-E_{\Xi_0}[\xi_0]}\right)^{-\alpha} \left(\frac{w}{\alpha}\right)^{\alpha} \left(\frac{1}{1-\alpha}\right)^{1-\alpha} > 0.$

To derive theoretical predictions for the effect differences in entry cost have on the legal form-specific size at firm entry, we defined the expected legal form-specific entry sizes in equations (S_0) and (S_1) as $E_\Theta[n^*|L=0] = \int_{\theta_0^*}^{\theta_1^*} n^*(\theta)/(\theta_1^* - \theta_0^*)\, d\theta$ and $E_\Theta[n^*|L=1] = \int_{\theta_1^*}^{1} n^*(\theta)/(1-\theta_1^*)\, d\theta$. Based on this definition, the following proposition results immediately and motivates hypotheses 2a and 2b.

Proposition 2. *Consider an equilibrium in which there exist workers and both types of entrepreneurs given fixed cost (F_0, F_1). An increase in the entry cost for CLLs, F_1, leads to an increase in the average size of both firm types. This increase is disproportionately larger for CLLs.*

Proof. Obviously, both $E[n^*(\theta)|\theta \in [\theta_0^*, \theta_1^*]]$ and $E[n^*(\theta)|\theta \in [\theta_1^*, 1]]$ increase in θ_1^* as $n^*(\theta)$ increases in θ.
W.l.o.g. assume that the increase in the entry cost for CLLs, F_1, from $\underline{F}_1$ to $\bar{F}_1$ implies an increase in θ_1^* from $\underline{\theta}_1^*$ to $\bar{\theta}_1^*$. Moreover, w.l.o.g. assume that $\theta_0^* = 0$.
Then it holds for any continuous real valued function $\tilde{n}(\theta)$ that

$$\begin{aligned}
& E\left[\tilde{n}(\theta)|\theta \in [0, \bar{\theta}_1^*]\right] - E\left[\tilde{n}(\theta)|\theta \in [0, \underline{\theta}_1^*]\right] \\
= & \int_0^{\bar{\theta}_1^*} \tilde{n}(\theta) \frac{1}{\bar{\theta}_1^*}\, d\theta - \int_0^{\underline{\theta}_1^*} \tilde{n}(\theta) \frac{1}{\underline{\theta}_1^*}\, d\theta \\
= & \int_{\underline{\theta}_1^*}^{\bar{\theta}_1^*} \tilde{n}(\theta) \frac{1}{\bar{\theta}_1^*}\, d\theta + \int_0^{\underline{\theta}_1^*} \tilde{n}(\theta) \underbrace{\left(\frac{1}{\bar{\theta}_1^*} - \frac{1}{\underline{\theta}_1^*}\right)}_{=\frac{\underline{\theta}_1^* - \bar{\theta}_1^*}{\bar{\theta}_1^* \underline{\theta}_1^*}} d\theta \\
= & \int_{\underline{\theta}_1^*}^{\bar{\theta}_1^*} \tilde{n}(\theta) \frac{1}{\bar{\theta}_1^*}\, d\theta - \int_0^{\underline{\theta}_1^*} \tilde{n}(\theta) \left(\frac{\bar{\theta}_1^* - \underline{\theta}_1^*}{\bar{\theta}_1^* \underline{\theta}_1^*}\right) d\theta
\end{aligned}$$

and

$$\begin{aligned}
& E\left[\tilde{n}(\theta)|\theta \in [\bar{\theta}_1^*, 1]\right] - E\left[\tilde{n}(\theta)|\theta \in [\underline{\theta}_1^*, 1]\right] \\
= & \int_{\bar{\theta}_1^*}^{1} \tilde{n}(\theta) \frac{1}{1-\bar{\theta}_1^*}\, d\theta - \int_{\underline{\theta}_1^*}^{1} \tilde{n}(\theta) \frac{1}{1-\underline{\theta}_1^*}\, d\theta \\
= & \int_{\bar{\theta}_1^*}^{1} \tilde{n}(\theta) \underbrace{\left(\frac{1}{1-\bar{\theta}_1^*} - \frac{1}{1-\underline{\theta}_1^*}\right)}_{=\frac{\bar{\theta}_1^* - \underline{\theta}_1^*}{(1-\bar{\theta}_1^*)(1-\underline{\theta}_1^*)}} d\theta - \int_{\underline{\theta}_1^*}^{\bar{\theta}_1^*} \tilde{n}(\theta) \frac{1}{1-\underline{\theta}_1^*}\, d\theta \\
= & \int_{\bar{\theta}_1^*}^{1} \tilde{n}(\theta) \left(\frac{\bar{\theta}_1^* - \underline{\theta}_1^*}{(1-\bar{\theta}_1^*)(1-\underline{\theta}_1^*)}\right) d\theta - \int_{\underline{\theta}_1^*}^{\bar{\theta}_1^*} \tilde{n}(\theta) \frac{1}{1-\underline{\theta}_1^*}\, d\theta \quad .
\end{aligned}$$

From these two equations it follows that

$$E\left[\tilde{n}(\theta)|\theta \in [\bar{\theta}_1^*, 1]\right] - E\left[\tilde{n}(\theta)|\theta \in [\underline{\theta}_1^*, 1]\right] \;>\; E\left[\tilde{n}(\theta)|\theta \in [0, \bar{\theta}_1^*]\right] - E\left[\tilde{n}(\theta)|\theta \in [0, \underline{\theta}_1^*]\right]$$

is equivalent to

$$\chi(\tilde{n}) := \int_{\bar{\theta}_1^*}^{1} \tilde{n}(\theta) \left(\frac{\bar{\theta}_1^* - \underline{\theta}_1^*}{(1-\bar{\theta}_1^*)(1-\underline{\theta}_1^*)} \right) d\theta - \int_{\underline{\theta}_1^*}^{\bar{\theta}_1^*} \tilde{n}(\theta) \underbrace{\left(\frac{1}{1-\underline{\theta}_1^*} + \frac{1}{\bar{\theta}_1^*} \right)}_{= \frac{1+\bar{\theta}_1^* - \underline{\theta}_1^*}{(1-\underline{\theta}_1^*)\bar{\theta}_1^*}} d\theta$$
$$+ \int_{0}^{\underline{\theta}_1^*} \tilde{n}(\theta) \left(\frac{\bar{\theta}_1^* - \underline{\theta}_1^*}{\bar{\theta}_1^* \underline{\theta}_1^*} \right) d\theta \qquad > 0 \quad .$$

To show that the entry cost induced size increase is disproportionately larger for CLLs, we still need to prove that $\chi(n^*) > 0$ where $n^*(\cdot)$ is defined in (FOC).

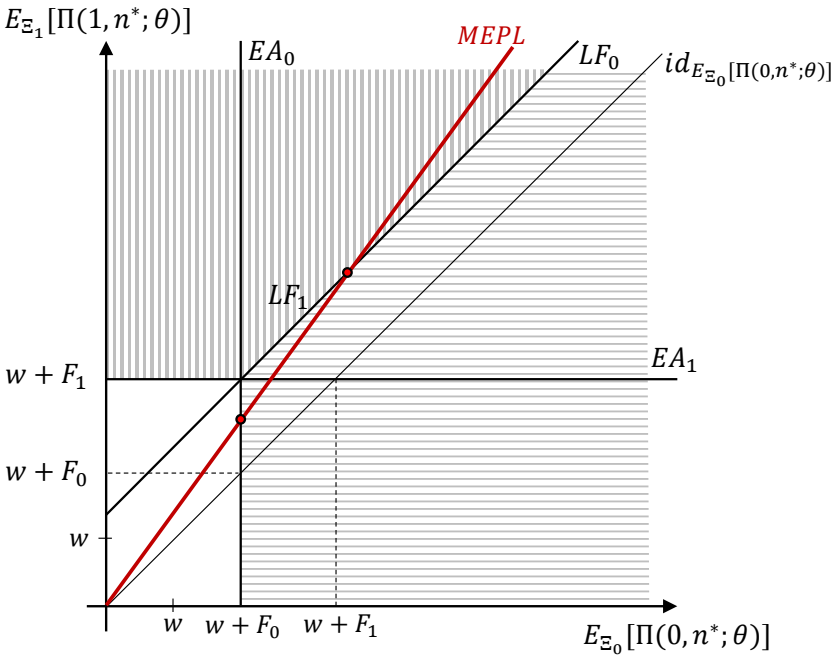
It is straightforward to show that $\chi(l) = 0$ for any affine function $l(\cdot)$ and, in particular, for $l(\theta) := n^*(\underline{\theta}_1^*) + \frac{n^*(\bar{\theta}_1^*) - n^*(\underline{\theta}_1^*)}{\bar{\theta}_1^* - \underline{\theta}_1^*} \cdot \theta$. Conclusively, $\chi(n^*) > 0$ is equivalent to $\chi(\tilde{n}) > 0$ with $\tilde{n}(\theta) := n^*(\theta) - l(\theta)$. But $\chi(\tilde{n}) > 0$ follows immediately from the definition of $l(\theta)$ and the convexity of $n^*(\theta)$ as

$$\begin{aligned} & n^*(\theta) > l(\theta) \quad \text{if } \theta \in]0, \underline{\theta}_1^*[\ \cup \ ]\bar{\theta}_1^*, 1[\\ \text{and} \quad & n^*(\theta) < l(\theta) \quad \text{if } \theta \in]\underline{\theta}_1^*, \bar{\theta}_1^*[\quad . \end{aligned}$$

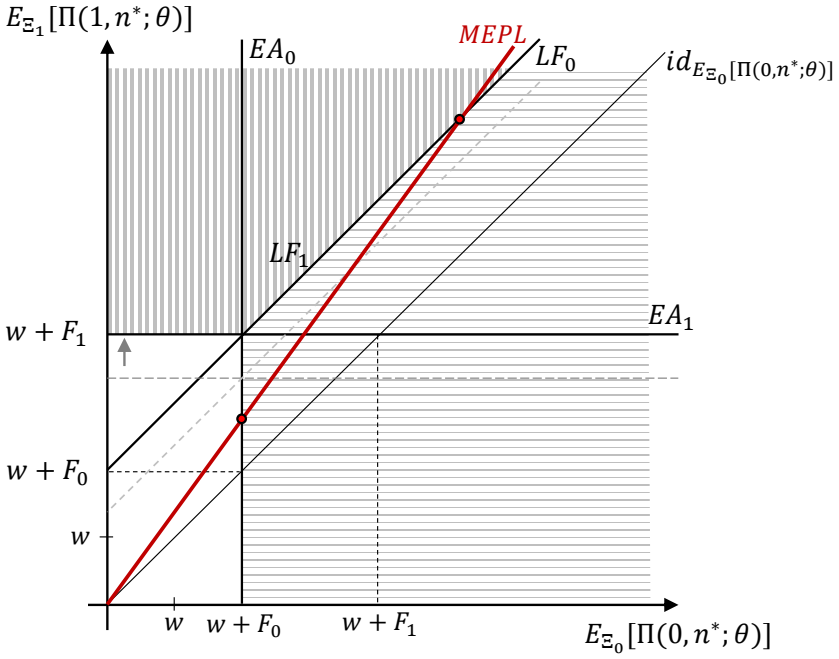
□

The last hypothesis 3 is a by-product of the preceding proofs and the mechanism driving the described size effects. Assuming that an increase in the CLL-specific entry cost, F_1, from $\underline{F}_1$ to $\bar{F}_1$ implies an increase in the CLL threshold, θ_1^*, from $\underline{\theta}_1^*$ to $\bar{\theta}_1^*$, all *non-marginal* potential entrepreneurs of type $\theta \in [0,1] \setminus [\underline{\theta}_1^*, \bar{\theta}_1^*[$ do not adapt their decision on the legal form. Only *marginal* entrepreneurs—characterized by $\theta \in [\underline{\theta}_1^*, \bar{\theta}_1^*[$—are deterred from incorporating and substitute setting up a CLL by setting up an unincorporated firm.

If both firm types exist under low as well as high CLL-related entry costs, then it holds that $\theta_0^* < \underline{\theta}_1^* < \bar{\theta}_1^* < 1$ and, conclusively, $n^*(\theta_0^*) < n^*(\underline{\theta}_1^*) < n^*(\bar{\theta}_1^*) < n^*(1)$ due to Lemma A1. Defining firms of intermediate size as those with a size between $n^*(\underline{\theta}_1^*)$ and $n^*(\bar{\theta}_1^*)$, it is only firms of intermediate size that are deterred from incorporating. We summarized this observation in hypothesis 3.



(a) Small CLL-entry cost



(b) Large CLL-entry cost

Figure A1: Entrepreneurship conditions, legal form conditions and the maximum expected profit tuple line

B Data

This data appendix contains information on the construction of additional variables in Section B.1. Section B.2 refers to the data sources.

B.1 Variables

Predicted incidence of CLLs among entrants before reunification

To generate our main instrument, we use information on a representative sample of 2,586 firms, denoted by p. The firms are originate or derivative start-ups in West Germany in the pre-reunification period in 1989. The data source is again Creditreform, as in case of our main firm sample.

We, first, estimate a linear probability model to predict for a firm p the choice of CLL, using indicators of the firm's industrial environment and the type of geographical region it is located in. To measure the industrial environment, we use disaggregated, 3-digit industry classes (Klassifikation der Wirtschaftszweige 1979) if we have legal form information on at least 20 obervations in the respective class. Otherwise we cluster closely related 3-digit industry classes, that is within the same 2-digit industry, to bundles with at least 20 observations. We distinguish between geographical regions with high aggregation of economic activity, intermediate aggregation and rural areas. To separate between these different types of regions, we use BBSR's classification of settlement structure at the county-level (Kreise). Then we rely on the estimates on the sample of pre-reunification firms to predict for each post-reunification firm i in our main sample the incidence of CLLs *before* reunification, but in the same industrial environment and the same type of geographical region. We denote the generated pre-reunification CLL-incidence as $I_{jd,\ pre\text{-}reunific.}$ and use it directly as excluded instrument in our two-stage least squares approach, or interacted with the indicator for East Germany, $\mathsf{East}_d \times \mathsf{I}_{jd,\ pre\text{-}reunific.}$.

Note that robust 2-SLS standard errors are asymptotically valid under weak conditions independent of whether the excluded instrument is a generated measure or not (Wooldridge, 2010).As finite sample properties can differ nevertheless, we also calculated non-parametric bootstrapped standard errors based on 400 replications and found these to be similar to the robust 2-SLS standard errors.

Entry Rate

In Section 4.1 we use a measure of the region-specific entry rate. This entry rate is defined as the weighted number of entrants in a region per year within our sample period divided by the region-specific average working age population in thousands.

Having age information from the BBSR in five year bins, we define the working age to be the interval from 20 to 65.

B.2 Data sources

B.2.1 Creditreform data

To construct the firm sample for our empirical analysis we started out with a random draw of 10,000 East and 12,000 West German firm records from the firm data sources of the leading German credit rating agency, Creditreform, as maintained at the Center for European Economic Research (ZEW), Mannheim. The random sample was drawn according to a stratified sampling rule such that firm groups with high liquidation risk according to Creditreform's encoded information were oversampled approximately twofold. The exact sampling probabilities are known and the sampling scheme is taken into account in all estimation and test procedures used below.

After having eliminated all 5,737 sampled firms that had started market activity before July 1st, 1990, and, thus, before our time period of interest, we applied four further exclusion restrictions. First, we removed 991 records on holding companies, part-time projects, and legally dependent firm units. Second, we eliminated 344 further firms to restrict the sample to those legal forms that are of interest here, leaving us with 14,928 firms: limited liability companies (GmbH), sole proprietorships (Einzelunternehmen, Gewerbebetrieb), commercial partnerships (KG, OHG), and civil law associations (GbR). Third, we eliminated 1,347 firm records without information on the relevant firm characteristics at market entry. Finally, we dropped 2,194 firm records from the main estimation sample because information on the age of the main owner was not available. Table 1 provides the definitions of all variables that we use and descriptive statistics for the main sample of 11,387 firms.

B.2.2 BBSR data

To calculate entry rates we use data from the Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt und Raumforschung, BBSR) partly published by its predecessor organization Bundesforschungsanstalt für Landeskunde und Raumordnung, BfLR (1995). This data set contains for each of the years 1991 to 1993 the district-specific population which is subdivided according to 5 year age bins. Based on this, the district-specific working age population is calculated as the district population aged 20 to 65 averaged over the three sample years. Moreover, for the year 1993 the data set contains district-specific information on the size of the district (in km^2) and the number of

workers in manufacturing sectors.
As in our main data set we exclude Berlin.

B.2.3 BIBB data

In Section 4.1 we use data from the 1992 wave of the "Qualification and Career Survey" provided by the Federal Institute for Vocational Education and Training (Bundesinstitut für Berufsbildung, BIBB) to assess the role migration of self-employed played immediately after reunification. To do so, we start with the full sample of 34,277 individuals. Then current residents of East Germany and West-Berlin, foreigners as well as individuals younger than 18 or older than 65 are eliminated. Finally, we exclude all individuals not working in one of the manufacturing, construction, trade and service industries leaving us with 23,377 observations.
East-West migrants are defined as those individuals that grew up in East Germany and reside in West Germany during the time period of the interview; West-East migrants are defined respectively.

C Additional Empirical Results

Table C1: Incidence of starting a CLL under different CLL-related entry cost (Probit)

Dependent variable:	(1) Probit	(2) Probit	(3) Probit CLL_i	(4) Probit	(5) Probit
E_d	-0.0498***	-0.1325***	-0.1375***	-0.1280***	-0.0285
	(0.0156)	(0.0153)	(0.0176)	(0.0162)	(0.0201)
Firm & Entrep. Controls	No	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
District Controls	No	No	Yes	Yes	Yes
Pseudo Log-Likelihood	-7452.1	-5090.8	-5086.6	-4789.3	-241.7
N	11387	11387	11387	10319	972

Note: The table shows estimated marginal effects (evaluated at the mean) of a probit model for variants of regression model (1). Columns 1 to 3 are based on the main sample of 11,387 entrants. Column 4 focuses on the restricted sample of firms without any affiliation to a parent company, column 5 on the opposite sample of affiliated firms.

The dependent variable, CLL_i, is an indicator that takes the value one for entrants starting a firm under the German limited liability company law. Column 1 includes the East indicator, E_d, as the only explanatory variable apart from the intercept. In column 2 firm and entrepreneurial controls as well as industry and year fixed effects are added. Columns 3 to 5 additionally including district-specific controls show the estimates of the model formalized in equation (1). All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table C2: Conditional decomposition of the effect covariates have on main explanatory variables

	(1) OLS	(2) IV
Dependent variable:	$\ln(\mathsf{S}_i)$	
CLL_i		
Firm & Entrep. Controls	0.2497***	0.0909
	(0.0132)	(0.1057)
Industry FE	-0.0475***	-0.5403***
	(0.0121)	(0.1260)
District FE	-0.0090	-0.0483
	(0.0074)	(0.0321)
Year FE	0.0081***	0.0091*
	(0.0023)	(0.0052)
Total Effect	0.2013***	-0.4887**
	(0.0187)	(0.2247)
$\mathsf{E}_d \times \mathsf{CLL}_i$		
Firm & Entrep. Controls	0.0710***	0.0388
	(0.0116)	(0.0243)
Industry FE	0.0996***	0.2252**
	(0.0187)	(0.0969)
District FE	0.0105	0.0420
	(0.0090)	(0.0429)
Year FE	0.0025	-0.0129
	(0.0030)	(0.0095)
Total Effect	0.1835***	0.2931***
	(0.0232)	(0.1103)
N	11387	11387

Note: The table shows the conditional decomposition of the effect covariates have on the coefficients on CLL_i and $\mathsf{E}_d \times \mathsf{CLL}_i$ when adding controls to the base specification in column 1 of table 3 (Gelbach, 2016). Column 1 shows results of the conditional decomposition for the OLS results, column 2 the equivalent for the IV estimation.

The dependent variable, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. The explanatory variable, CLL_i, is an indicator for whether the entrepreneur starts his firm under a legal form with limited liability. The interaction, $\mathsf{E}_d \times \mathsf{CLL}_i$, takes one for all entrants that start a limited liability company in East Germany. The excluded instruments underlying the regression in column 2 are the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ pre\text{-}reunific.}$, and its interaction with the indicator for East Germany as explained in detail in Section 4.2.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table C3: Firm size at market entry using further disaggregated fixed effect sets

	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) IV
Dependent variable:			$\ln(\mathsf{S}_i)$		
CLL_i	0.4990***	0.5009***	1.0130***	0.5113***	0.7487***
	(0.0261)	(0.0257)	(0.2256)	(0.0262)	(0.2630)
$\mathsf{E}_d \times \mathsf{CLL}_i$	0.4242***	0.4317***	0.9124***	0.3999***	1.5214***
	(0.0375)	(0.0373)	(0.1088)	(0.0402)	(0.2879)
Firm & Entrep. Controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	No	No	No	No
3-dig. Industry FE	Yes	No	No	No	No
Industry-Year FE	No	Yes	Yes	No	No
District-Year FE	No	Yes	Yes	No	No
West × Industry FE	No	No	No	Yes	Yes
East × Industry FE	No	No	No	Yes	Yes
District FE, Year FE	No	No	No	Yes	Yes
R^2_{adj}	0.421	0.412	0.319	0.409	0.265
N	11387	11387	11387	11387	11387
Kleibergen-Paap F			30.452		33.841

Note: Columns 1, 2 and 4 provide ordinary least squares estimates for variants of our main regression model formalized in equation (2) for the main sample of 11,387 entrants. In columns 3 and 5, we provide instrumental variable estimates.

The dependent variable, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. Column 1 includes further disaggregated industry and district fixed effects. Columns 2 and 3 use industry and district fixed effects that vary across entry cohorts. Columns 4 and 5 display the estimates of a regression controlling with separate industry fixed effects for East and West Germany. The excluded instruments in columns 3 and 5 are the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ \textit{pre-reunific.}}$, and its interaction with the indicator for East Germany as explained in detail in Section 4.2. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table C4: Heterogenous effects across quantiles of the entry size distribution (quantile regression approach for count data)

	(1) $Q_z(0.30\|x)$	(2) $Q_z(0.40\|x)$	(3) $Q_z(0.50\|x)$	(4) $Q_z(0.60\|x)$	(5) $Q_z(0.70\|x)$	(6) $Q_z(0.80\|x)$	(7) $Q_z(0.90\|x)$
Panel A:							
CLL_i	0.2919***	0.3630***	0.4416***	0.5371***	0.6558***	0.7827***	0.9415***
	(0.0190)	(0.0214)	(0.0235)	(0.0290)	(0.0338)	(0.0411)	(0.0663)
$\mathsf{East}_d \times \mathsf{CLL}_i$	0.5327***	0.5512***	0.5315***	0.4716***	0.3921***	0.2869***	0.1403
	(0.0326)	(0.0334)	(0.0374)	(0.0428)	(0.0488)	(0.0676)	(0.1179)
Panel B:							
$ME_z : \mathsf{LLC}_i$	0.598***	0.845***	1.182***	1.700***	2.583***	4.106***	7.815***
	(0.041)	(0.054)	(0.069)	(0.104)	(0.158)	(0.265)	(0.698)
$ME_z : \mathsf{East}_d \times \mathsf{LLC}_i$	1.273***	1.495***	1.631***	1.655***	1.638***	1.511***	1.085
	(0.095)	(0.109)	(0.138)	(0.175)	(0.227)	(0.390)	(0.959)
$\hat{Q}_{\mathsf{z}}(\cdot\|\bar{x})$	2.290	2.644	3.057	3.589	4.359	5.586	8.288
$ME_{\mathsf{S}} : \mathsf{LLC}_i$	0	1	1	2	2	4	8
$ME_{\mathsf{S}} : \mathsf{East}_d \times \mathsf{LLC}_i$	1	2	1	2	1	2	1
N	11387	11387	11387	11387	11387	11387	11387

Note: The table reports estimates of a quantile regression approach for count data (Machado and Silva, 2005) for the full sample of 11,387 entrants.
The dependent variables in columns 1 to 7 are the 30%-, 40%-, 50%-, 60%-, 70%-, 80%-, and 90% quantiles of the jittered entry size, z. The estimates in Panel A result from the model formalized in equation (2). Panel B depicts the corresponding marginal effects (ME) on the jittered entry size, z, and the discrete entry size, S, respectively, as well as the predicted quantiles, $\hat{Q}_z(\cdot|x)$. To calculate marginal effects all continuous variables are set to their mean and all dummy variables are set to zero.
The estimates are based on 1,000 jittered samples. Standard errors are reported in parentheses. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.

Table C5: Firm size at market entry excluding large and small start-ups

	(1) OLS	(2) OLS	(3) IV	(4) IV
Dependent variable		$\ln(\mathsf{S}_i)$		
CLL_i	0.4584***	0.3745***	0.9155***	0.6727**
	(0.0215)	(0.0260)	(0.2007)	(0.2802)
$\mathsf{East}_d \times \mathsf{CLL}_i$	0.2466***	0.3100***	0.7742***	0.6587***
	(0.0290)	(0.0361)	(0.1037)	(0.1064)
Firm & Entrep. Controls	Yes	Yes	Yes	Yes
District FE, Industry FE, Year FE	Yes	Yes	Yes	Yes
R^2_{adj}	0.335	0.311	0.227	0.267
N	10840	7882	10840	7882
Kleibergen-Paap F			29.849	19.286

Note: Columns 1 and 2 provide ordinary least squares estimates of our main regression model formalized in equation (2). In columns 3 and 4, we provide instrumental variable estimates. While columns 1 and 3 use the restricted sample of firms with up to 20 employees, columns 3 and 4 report the estimates for a sample excluding 1-person firms.

The dependent variable, $\ln(\mathsf{S}_i)$, is defined as the logarithm of the number of employees including working entrepreneurs. The explanatory variable, CLL_i, is an indicator for whether the entrepreneur starts his firm under a legal form with limited liability. The interaction, $\mathsf{E}_d \times \mathsf{CLL}_i$, takes one for all entrants that start a limited liability company in East Germany. The excluded instruments in column 3 and 4 are the predicted industry- and region-specific incidence of CLLs before reunification, $\mathsf{I}_{jd,\ pre\text{-}reunific.}$, and its interaction with the indicator for East Germany as explained in detail in Section 4.2. All other explanatory variables are defined as in Table 1.

Sampling weights are used as explained in detail in Appendix B.2. Standard errors in parentheses are robust and clustered to allow for unrestricted correlation between observations within the same industry in the same district. Statistical significance at the 10%, 5% and 1% level is indicated by *, ** and ***.