DISCUSSION PAPER SERIES

DP15549

SPILLOVER EFFECTS IN EMPIRICAL CORPORATE FINANCE

Tobias Berg, Markus Reisinger and Daniel Streitz

FINANCIAL ECONOMICS

CE PR

SPILLOVER EFFECTS IN EMPIRICAL CORPORATE FINANCE

Tobias Berg, Markus Reisinger and Daniel Streitz

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Financial Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

SPILLOVER EFFECTS IN EMPIRICAL CORPORATE FINANCE

Abstract

Despite their importance, the discussion of spillover effects in empirical research often misses the rigor dedicated to endogeneity concerns. We analyze a broad set of workhorse models of firm interactions and show that spillovers naturally arise in many corporate finance settings. This has important implications for the estimation of treatment effects: i) even with random treatment, spillovers lead to a complicated bias, ii) fixed effects can exacerbate the spillover-induced bias. We propose simple diagnostic tools for empirical researchers and illustrate our guidance in an application.

Tobias Berg - t.berg@fs.de
Frankfurt School of Finance & Management

Markus Reisinger - m.reisinger@fs.de
Frankfurt School of Finance & Management and CEPR

Daniel Streitz - dst.fi@cbs.dk
Copenhagen Business School

1. Introduction

Spillovers arise in many corporate finance settings through firm competition or local interdependencies among firms. Assume, for example, that a firm is subject to a credit supply shock, a natural catastrophe, a cash windfall, a different tax treatment, or a new regulation affecting, for example, the capital or ownership structure. These treatments naturally affect firms in the same industry through competitive interactions, or firms in the same region through local demand and agglomeration effects. In this paper, we explore the impact of spillovers on the estimation of treatment effects in classic corporate finance settings.

The source of spillovers depends on the economics of the problem. We therefore start by providing a theoretical foundation based on a broad set of classic workhorse models of firm interactions. Spillovers in these models arise either from competitive effects (e.g., price or quantity competition between oligopolistic firms) or through spatial interactions (e.g., local demand spillovers or agglomeration effects). A pervasive result of all workhorse models is that the firm-level outcome—such as investments, sales, or employment—depends not only on a firm's own treatment status, but also on the fraction of treated firms in the same industry or region. We highlight key commonalities and differences across models. For example, we show that a shock on the capacity typically results in heterogeneous spillovers on treated and non-treated firms, whereas a shock on the production costs often results in homogeneous spillover effects.

Our theoretical insights have three key econometric implications: First, the existence of spillover effects implies that the stable unit treatment value assumption (SUTVA) is violated. Thus, even with random treatment assignment, spillovers lead to a bias in estimating treatment effects. We further show that the resulting bias is complicated: even in a simple linear model, the bias depends on higher-order moments of the group-level treatment fraction; and estimates that ignore spillovers are not only wrong in magnitude, but can also have the wrong sign. Second, adding (industry or region) fixed effects can worsen the problem. As fixed effects estimators focus on within-group variation, any bias induced by spillovers within an industry or within a region can be amplified with fixed

effects estimators. Third, the econometric strategy to handle spillover effects depends both on the source of spillovers as well as the nature of the shock (e.g., a shock that acts as a capacity constraint requires estimating a model with heterogeneous spillover effects).

Based on our analysis, we provide a three-step guidance for applied researchers how to handle spillovers in empirical corporate finance research using firm-level data. First, researchers should consider the most plausible economic mechanism for spillovers (e.g., competition, demand, or agglomeration effects), the nature of the shock (e.g., a shock on marginal costs or a shock that acts as a binding capacity constraint) as well as the most plausible dimension, g, for spillovers (e.g., industry or region). Second, the relation between outcome and treatment is generally of the form $y_i = f(d_i, \overline{d}_g, d_i\overline{d}_g)$ in all workhorse models of firm interactions, where y_i is the outcome for firm i, d_i is a treatment indicator, and $\overline{d}_g$ is the fraction of treated units in group g. The workhorse models predict that the heterogenous effect $d_i\overline{d}_g$ is particularly important for shocks that act as a binding capacity constraints (such as a financial constraint). This is due to the fact that a treated firm is less sensitive to the overall competitive situation or aggregate demand as it cannot increase its capacity. Third, we provide a simple way to illustrate direct effects, spillover effects, and aggregate effets in a single graph. We hope that this guidance will be useful to academics in future research.

We illustrate the effects of spillovers in an empirical setting based on Huber (2018) who examines the effects of bank lending on the real economy, exploiting an exogenous lending cut by Commerzbank (CB), a large German bank, during the financial crisis. We replicate the analysis using the original methodology and find very similar results: firms fully dependent on CB cut employment by 2-3 percentage points (direct effect). Furthermore, in a county with average CB-dependence, demand spillovers lead to an additional cut in employment for *all* firms by 3 percentage points (indirect effect).

Applying our framework to this setting uncovers two novel insights. First, the direct effect of CB-dependence on firm employment is twice as large as estimated in Huber (2018). For a firm fully dependent on CB, the direct effect amounts to 5.5 percentage points. Second, spillover effects are heterogenous. We show that this heterogeneity is consistent with a local demand channel. The credit supply reduction is the binding con-

straint for treated firms, i.e., they are not exposed to the negative demand spillover. For control group firms the associated drop in demand (and not the credit supply) constitutes the binding constraint, i.e., they experience a negative demand spillover effect.

Our paper contributes to the growing literature on empirical methods in (corporate) finance and economics (Bertrand et al., 2004; Ali et al., 2009; Petersen, 2009; Roberts and Whited, 2012; Gormley and Matsa, 2014; Jiang, 2017; Goodman-Bacon, 2018; Borusyak and Jaravel, 2017; Grieser and Hadlock, 2019). The importance of assuming non-interference when interpreting randomized experiments goes back to Cox (1958) and Rubin (1978), and several papers have discussed either the importance of this assumption in specific settings [see e.g., Angrist et al. (1996), Heckman et al. (1998), and Miguel and Kremer (2004) for earlier work and Boehmer et al. (2020), Bustamante and Frésard (2020), and Duguay et al. (2020) for recent examples], or developed empirical models to estimate spillover effects [see e.g., Manski (1993) for peer effects, Anselin (1988) and Elhorst (2014) for spatial econometrics, Duflo and Saez (2003), Baird et al. (2018), and Vazquez-Bare (2018) for the design of experiments in the presence of spillover effects]. We add to this literature in three ways. First, we document that spillover effects naturally arise in typical papers in corporate finance. Second, we analyze a broad set of workhorse models on firm interactions and derive common insights from these models. In particular, SUTVA is violated in all models, and the relation between outcome and treatment depends on the group-level treatment fraction. Third, we discuss the econometric implications, thereby showing that spillovers in common workhorse models lead to a *complicated* bias in estimating treatment effects.[1]

Our paper is further linked to the literature on structural estimation in (corporate) finance, which relies on an explicit theoretical model to impose identifying restrictions (e.g. Hennessy and Whited, 2005; Hennessy et al., 2007). We document that spillover effects

---

[1]The discussion about spillover effects in empirical research is also related to the literature on estimating general equilibrium effects. While we focus on identification in partial equilibrium, correctly identifying causal effects is informative for general equilibrium models. As argued by Nakamura and Steinsson (2018), "identified moments" (i.e., causal effects estimated using partial equilibrium techniques such as DiD setups) are helpful as target moments that general equilibrium models should match. In this paper, we show that spillover effects have to be taken into account in partial equilibrium estimations to generate meaningful identified moments. Else, estimates are biased and without clear economic interpretations. There is also a nascent literature on using partial equilibrium estimates to make aggregate statements (see e.g. Sraer and Thesmar, 2019; Chodorow-Reich, 2020). Such techniques, however, require an explicit treatment of the aggregation framework, which is beyond the scope of this paper.

in key models of firm interactions are governed by a similar structure. In particular, spillovers depend on the fraction of treated firms in the same industry or region in all models, and shocks to the capacity of a firm result in heterogeneous spillovers.

We also contribute to the literature on the effects of credit supply on the real economy (see, for instance, Chodorow-Reich, 2014; Khwaja and Mian, 2008),[2] documenting that spillovers play a major role in the transmission of credit supply shocks, and can significantly bias direct treatment effect estimates of credit supply shocks.

The rest of the paper proceeds as follows: in Section 2, we provide a survey of papers that estimate treatment effects using difference-in-differences (DiD) models to give an overview if and how spillover effects are discussed in the existing literature. In Section 3, we analyze spillover effects in a broad set of classic workhorse models of firm interactions. This theoretical discussion lays the ground for the conceptual discussion in Section 4 on how to handle spillover effects in empirical research. Section 5 provides guidance for empirical researchers, and Section 6 illustrates the importance of accounting for spillover effects in an empirical application on the effects of credit supply contractions on employment. Section 7 concludes.

2. Survey of papers in major journals

We provide a survey of papers that estimate treatment effects using a difference-in-differences (DiD) approach—a setting with a clear control and treatment group—in Table 1. A DiD approach is used in 103 out of 610 papers published in the main economics and finance journals in 2017. Out of the 103 DiD papers, only 22 papers (21%) contain some discussion of spillovers and only 17 papers (16%) contain a quantitative analysis of spillovers (see Panel A).

[Table 1 here]

---

[2] The literature that examines the real effects of credit supply shocks is large. See for example Acharya et al. (2018), Aiyar (2012), Almeida et al. (2011), Amiti and Weinstein (2011), Chava and Purnanandam (2011), De Haas and Horen (2012), Gan (2007), Iyer et al. (2014), Ivashina and Scharfstein (2010), Paravisini et al. (2014), Peek and Rosengren (1997), Peek and Rosengren (2000), Santos (2011), and Sufi (2009), among others.

Papers that address spillovers mostly follow one of three strategies (see Panel B): 1) control group observations where spillovers are most plausible are dropped from the sample,[3] 2) control group outcomes are regressed on an exposure measure to spillovers,[4] 3) estimates on the individual level are compared with estimates on a more aggregate level.[5] Importantly, all three strategies are applied in a relatively ad-hoc way. We show that all three strategies are viable under specific assumptions, though these assumptions are not identical for the three strategies discussed above (see Section 4.3). We argue that researchers should be aware of these assumptions in order to evaluate the adequacy of each of these three strategies.

Panel C compares papers that discuss with those that do not discuss spillovers. Among papers that discuss spillovers, the fields of labor and education economics are most common (about one-third of all papers). These are areas where indirect (typically peer) effects traditionally receive more attention. Among papers that do not discuss spillovers, about two-thirds of the papers are corporate finance or financial intermediation papers. These are papers that conduct analyses on the firm level, considering variables such as investment, firm value, compensation/earnings, credit supply, asset (or sales) growth, innovation, and employment. Spillover effects may naturally arise in these settings if firms compete with one another or affect one another locally (e.g., through agglomeration or local demand effects). For instance, output choices (such as investment in capacity, employment, or sales growth) or innovation activities may be determined in an industry equilibrium, which implies that shocks to a subset of units will likely have an effect on competitors as well.

The key takeaway from this survey is that spillover effects are often ignored in firm-level analyses, even though they are highly plausible in these settings. In the next section,

---

[3]An example is Hornbeck and Keniston (2017) who analyze the impact of the 1872 Boston Fire on land values. Landplots close to the Boston Fire region are excluded from the control group in order to mitigate the effect of spillovers.

[4]An example is Shue and Townsend (2017) who explain the rise in CEO compensation during the 1990s and 2000s by number-rigidity, i.e., the practice of granting the same number of stock options every year. In one specification, they regress CEO compensation on the proportion of peers with number-rigid pay to analyze spillovers to other CEOs.

[5]An example is Maggio et al. (2017) who exploit variation in the timing of resets of adjustable rate mortgages to analyze the effect of lower interst rates on consumption. Regressions are performed on the borrower level and the region level in order to assess indirect effects from the resulting increase in local demand.

we analyze key economic contexts for firm-level spillovers. We then show that spillover effects can severly bias estimators that are typically applied in empirical research.

3.   Theoretical Foundations

In this section, we introduce "treatments" in classic workhorse models of firm interaction and show that spillover effects naturally arise in all models. This provides the theoretical foundation for our conceptual discussion on how to treat spillover effects in empirical studies with firm-level data. Specifically, we discuss firm interdependencies arising through (imperfect) competition and spatial interaction between firms as plausible sources of spillover effects.[6]

In all models, we introduce two different treatment types that can be considered as simple stylized representations of a variety of treatment effects that are considered in firm-level DiD settings. The first treatment is a shock on the firm's cost function, i.e., the marginal costs of a subset of firms ("treated firms") increase. This may be because of a negative shock to a firm's technology, or because of a regulation that increases the costs of production. The second treatment is a shock on the firm's capacity, i.e., the production capacity is constrained for "treated firms". This may be because of a credit supply shock that cuts off a firm from receiving external finance, or a natural disaster that destroyed production sites.[7] To cleanly isolate spillover and treatment effects, we assume that firms are homogenous before the shock.

In what follows, we focus on the main insights and similarities of the models and provide their intuitions. All mathematical details all relegated to the Online Appendix. To discuss the main results of the models in a concise way, we use the following notation: d_i denotes the treatment indicator of firm (or unit) i; it is equal to 1 if the firm is subject to the treatment, and 0 otherwise. In addition, $\bar{d} \in (0,1)$ represents the fraction of treated firms.

Imperfect competition: We consider two widely-used models of (imperfect) competi-

---

[6]Although imperfect competition and spatial interaction between firms are certainly not the only sources for spillovers on the firm level, both are core economic relationships between firms.

[7]We also allow for a capacity constraint of zero, which implies that treated firms go bankrupt. See Bernstein et al. (2019) for an in-depth analysis of spillovers in this case.

tion between firms. These are, first, classic Cournot (quantity) competition and, second, price competition with differentiated products with firms located on a circle, as developed by Salop (1979). Both are workhorse models in oligopoly theory. In particular, Cournot competition applies to markets in which firms sell commodities, such as oil or wheat, and is heavily used in competition policy, e.g., for merger analysis (see Farrell and Shapiro (1990) and Nocke and Whinston (2010), among many others). Price competition with differentiated products applies to a wide range of markets in which different brands or types of goods are available, ranging from food over electronics to services. Circular competition along the lines of Salop (1979) is a common way to study this form of competition.[8]

In both models, we determine the equilibrium output of treated and non-treated firms. The output variable can subsume a variety of measures, such as firm size or number of sales, and is (positively) correlated with e.g. firm value and employment. Therefore, when empirically estimating potential differences between treated and non-treated firms, several variables can be used.

In a nutshell, the functioning of the models is as follows: A firm's output does not only depend on its own cost/capacity but also on the quantities or prices of its rivals. In case of quantity competition, smaller quantities by rivals increase the equilibrium price, which implies that each firm's profit margin is larger, inducing the firm to increase its output. Similarly, in case of price competition, if rivals charge a higher price and therefore sell less, a firm's output is higher. While the direct shock is relevant only for treated firms, and implies that the output of treated firms is lower compared to non-treated firms, the spillover effect from rivals affects both treated and non-treated firms.

As a result, we obtain that, in both models (i.e., quantity or price competition) and irrespective of the treatment type (i.e., marginal cost shock or capacity constraint), the output of each firm depends on its own treatment status (i.e., d_i), as well as the fraction of treated firms (i.e., $\bar{d}$). The existence of a "$\bar{d}$-effect" implies that SUTVA is violated in both models for both shocks, as the outcome of each firm also depends on the treatment status of other firms in the economy. The intuition is as follows: A firm that is exposed

---

[8]We also confirm our results by considering a third model with a representative consumer, following Bowley (1924) and Singh and Vives (1984), which encompasses both quantity and price competition and allows, for instance, to generalize the Cournot model to account for product differentiation.

to higher costs or reduced capacity (i.e., has a negative shock) chooses a lower output or sets a higher price. If a larger number of firms experiences such a cost shock (i.e., $\bar{d}$ increases), the direct effect on these firms is negative but the spillover effect on *all* firms is positive as competition is reduced.

It is important to note that the "$\bar{d}$-effect" is an *equilibrium property* of both models. The fact that spillovers only depend on the overall treatment intensity ($\bar{d}$) arises naturally in classic oligopoly models in which firms are homogenous without shock, and does not depend on the way we choose to model spillovers.

In addition, $\bar{d}$ may affect treatment and control-group firms *differently*, that is, the outcome may also depend on the interaction between d_i and $\bar{d}$, which we denote by $d_i\bar{d}$. The existence of a heterogenous "$\bar{d}$-effect" for treatment and control-group firms, however, depends on the nature of the treatment (but not the underlying model). In case of a shock on marginal costs, the spillover effect is homogenous, which implies that only $\bar{d}$ but not $d_i\bar{d}$ affects the outcome. The reason is that the reduced quantity (Cournot model) or the higher price (Salop, 1979, model) of a treated firm affects the profit of competitors in the same way, regardless of whether the competitor belongs to the control group or the treatment group. Instead, if a shock occurs on the capacity, treated firms cannot benefit from spillover effects as they produce at the capacity limit, regardless of how many other firms are exposed to the shock. This implies that $\bar{d}$ does not affect the outcome of treated firms. By contrast, firms in the control group benefit from the spillover effect, as some of their competitors sell a lower quantity.

Spatial interaction: In case of spatial interaction between firms, we also consider two different types of models: a model of demand spillovers and a model of agglomeration economies. The former is a highly relevant scenario in many local markets and also builds the intuition for our empirical analysis in Section 6. To analyze this case, we consider an adaptation of the framework by Shleifer and Vishny (1988) and Murphy et al. (1989). These papers are the most prominent ones dealing with demand spillovers and develop a tractable model to analyze the phenomenon. To model agglomeration effects, we consider a simple framework that captures the effect that clustered firms benefit more

than isolated firms from factors which, for instance, lead to cheaper production.[9] Our model partly follows the widely-used framework of d'Aspremont and Jacquemin (1988) and considers that firms benefit from a larger number of other firms due to investment spillovers that reduce own costs.

Our main finding is again that, in both models and for both types of shocks, SUTVA is violated, and the outcome for each firm depends on d_i, $\bar{d}$, and in some cases the interaction between the two.

First, a shock on the capacity of treated firms leads to very similar results as in case of imperfect competition, that is, spillover effects are heterogeneous between treated and control group firms. As treated firms produce at the capacity constraint, they are again not affected by spillovers. By contrast, the outcome for firms in the control group depends on the overall fraction of treated firms (i.e., $\bar{d}$). The main difference to the models with imperfect competition is the sign of the spillover effect. The effect is positive with imperfect competition, i.e., a negative shock to competitors strengthens the competitive position of control group firms. In case of spatial interaction, firms in the control group are *negatively* affected by spillovers, as there is either lower employment and therefore reduced demand (local demand model), or fewer investments in aggregate and therefore higher production costs (agglomeration model).

Second, if the shock occurs on marginal costs, spillovers are homogenous with agglomeration economies. This is because the cost-reducing effect of investments by neighboring firms is of the same strength for treated and non-treated firms. By contrast, in local demand models, a marginal cost shock results in heterogeneous spillover effects. The intuition is rooted in the effect that the spillover arises due to a change in employment, which induces a change in the disposable income of consumers. Because treated firms have higher costs, they sell less, and are thus less exposed to a shock on the purchasing power of consumers. Therefore, if the disposable income falls due to the reduction in employment as a consequence of the shock, control group firms experience a stronger negative effect as they sell a larger quantity.

Finally, both for imperfect competition and spatial interaction between firms, we show

---

[9]Agglomeration benefits along these lines are discussed by Marshall (1890), Hoover (1948), and Krugman (1991), among many others.

the results using models with a linear demand function and constant returns to scale (see the Online Appendix for details). These assumptions are common in the literature and allow us to derive closed-form solutions for the equilibrium. If the shock is on marginal costs, we obtain a simple linear relationship between equilibrium output and the variables d_i, $\bar{d}$, and $d_i\bar{d}$. If the shock is on the capacity constraint, $d_i\bar{d}$ affects the equilibrium output non-linearly in some models but d_i and $\bar{d}$ also enter the equilibrium expressions in a linear way in all models. Our theoretical foundation therefore provides some justification for estimating a *linear* model to uncover the strength of the different effects.

Key insights: Overall, several key insights emerge from the theoretical analysis:

1. *SUTVA is violated in common models of firm interaction,* such as competitive interaction and spatial interaction. This holds both for shocks on firms' marginal cost and capacity.

2. *The relation between outcome and treatment is generally of the form* $y_i = f(d_i, \bar{d}, d_i\bar{d})$, which implies that:

    (a) *spillover effects are determined by the overall treatment intensity in the economy*, i.e., $\bar{d}$;

    (b) *spillover effects are often heterogeneous* on units in the treatment group and the control group, i.e., y_i can be a function of $d_i\bar{d}$, even if units are homogenous before the shock.

   Table 2 provides a summary of the results showing which variables determine the outcome for the different contexts and shocks.

[Table 2 here]

4. Econometric framework

In this section, we draw on the theoretical results developed above and introduce an econometric framework to analyze the implications of spillovers on estimating treatment effects.

4.1. Assumptions, notation and framework

We make the following key assumption:

A.1 Treatment status fulfills the conditional independence assumption (CIA).

A.2 Outcomes not only depend on treatment status of an individual firm, but also on the treatment intensity in an industry (in case of competition models) or a region (in case of spatial models).

A.3 Spillovers occur within industries/regions, but not across industries/regions (i.e., we abstract from general equilibrium effects).

A.4 We assume a linear relationship throughout the paper.

A.1 ensures that—in the absence of spillovers—treatment effects can be recovered from a (conditional) comparison of treated and control group firms.[10] A.2 is guided by the workhorse models on firm interaction discussed in Section 3; it implies that SUTVA is violated. A.3 states that we remain in a partial equilibrium framework.[11] A.4 follows standard practice in most empirical papers, e.g., typical DiD models.[12]

These assumptions might not be fulfilled in real-life: treatment might not be random, more sophisticated models can yield non-linear relationships as well as relationships that depend on other moments than the mean treatment intensity in an industry/region, and general equilibrium effects can be important in practice. Our aim here is not to claim that these assumptions are generally fulfilled in every situation. Instead, we want to show in a convincing way that, even with random treatment and in simple but commonly employed models on firm interaction, treatment effects are biased, the resulting bias is complicated, and treatment estimates cannot only be wrong in magnitude, but also have the wrong sign.

---

[10]We make this assumption because methods to tackle endogeneity have been widely discussed in the literature, see Angrist and Pischke (2009) and Roberts and Whited (2012) for an overview.

[11]The two assumptions A.2 and A.3 are also referred to as exchangeability (spillovers do not depend on the specific identity of treated "neighbors") and partial interference (spillovers confined within group), see e.g. Vazquez-Bare (2018).

[12]We discuss the assumption of a linear specification in more detail at the end of Section 5.

Our notations closely follow Roberts and Whited (2012). We denote by d_{ig} a treatment indicator for unit i in group g that is equal to 1 if treatment is received, and 0 otherwise. Group g typically represents an industry or region, and it is known and observable to the researcher.[13] We denote by $\overline{d}_g$ the fraction of units treated in group g. The outcome y_{ig} can thus be written as:

$$\textbf{Spillover model}: \quad y_{ig} \; = \beta_0 + \beta_1 d_{ig} + \beta_T \overline{d}_g d_{ig} + \beta_C \overline{d}_g (1 - d_{ig}) + \epsilon_{ig}. \qquad (1)$$

Eq. (1) contains a direct treatment effect (β_1) as well as spillover effects to treated units (β_T) and to control units (β_C).[14] It is straightforward to see that spillovers—if ignored—induce a bias in estimating β_1 because d_{ig} and $\overline{d}_g$ are correlated $[\rho(d_{ig}, \overline{d}_g) \neq 0]$. In the following, we show that—even in the simple linear model—spillover effects induce a *complicated* bias in estimating treatment effects.

The direct treatment effect (β_1) warrants a more formal discussion. Let $y_{ig} = y_{ig}(d_{ig}, \overline{d}_g)$ be the function linking individual treatment (d_{ig}) and group-level average treatment intensity ($\overline{d}_g$) to the outcome variable y_{ig}. We define the *direct effect*, or equivalently, the *direct treatment effect* as $\lim_{\overline{d}_g \to 0} y(1, \overline{d}_g) - y(0, \overline{d}_g)$. The direct effect is well defined if $y(d_{ig}, \overline{d}_g)$ is continuous in $\overline{d}_g$ and it is meaningful if $\lim_{\overline{d}_g \to 0} y(0, \overline{d}_g) = y(0, 0)$. These two properties are naturally fulfilled by the linear model, Eq. (1), but can also easily be extended to non-linear models. We call this term the direct effect because it measures the treatment versus control group difference when almost all units are control group units. Given the continuity assumption, this term therefore has the natural interpretation of being a direct treatment effect that is not affected by spillover effects.

---

[13]Depending on the setting, defining the most plausible dimension for spillovers can be challenging. Note, however, that any researcher who ignores spillover effects in her setting makes the implicit assumption that either no spillover effects exist or the relevant group g comprises the entire sample population.

[14]Eq. (1) is a special case of both Manski (1993) and of models typically used in the spatial econometrics literature (see, for example, Anselin, 1988; Elhorst, 2014). In our case, spillovers occur via $\overline{d}_g$, while Manski (1993) and spatial spillover models also allow for spillovers via $\overline{y}_g$ (endogenous spillover effects) and via the error term, (see Born and Breitung, 2011, for diagnostic tests for the existence of endogenous spillover effects and spillover effects via the error term). Specific to our setup, the spillover variable $\overline{d}_g$ is bounded between $[0, 1]$, allowing us to derive statements that go beyond the results in the spatial econometrics literature.

4.2. Bias in estimating treatment effects when spillovers are ignored

In the following, we discuss the results of estimating:

$$\textbf{Estimated model}: \quad y_{ig} \quad = \tilde{\beta}_0 + \tilde{\beta}_1 d_{ig} + \tilde{\epsilon}_{ig}, \tag{2}$$

when the true model is Eq. (1). The crux of Eq. (1) is that the spillover terms are mechanically correlated with the treatment indicator. Even if treatment is assigned randomly and thus $Cov(d_{ig}, \epsilon_{ig}) = 0$, the spillover terms are generally still correlated with the treatment indicator $[Cov(d_{ig}, \beta_T \bar{d}_g d_{ig} + \beta_C \bar{d}_g (1 - d_{ig})) \neq 0]$ and thus lead to a bias when estimating β_1 via Eq. (2).

Proposition 1 *Assume y_i follows Eq. (1). Estimating Eq. (2) yields*

$$E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + (\beta_T - \beta_C)\bar{d} + \beta_T \frac{Var(\bar{d}_g)}{\bar{d}} + \beta_C \frac{Var(\bar{d}_g)}{1 - \bar{d}}. \tag{3}$$

Proof: See Appendix A.1.

Proposition 1 has two basic implications. First, estimating Eq. (2) does not generally yield the direct treatment effect β_1, even if treatment is random. The first term, $(\beta_T - \beta_C)\bar{d}$, arises from heterogeneous spillover effects which are attributed to β_1 when estimating Eq. (2). The last two terms arise because treated units are by definition more prevalent in high-treatment groups while control units are disproportionally prevalent in low-treatment groups. This unintentionally leads to an additional bias, $\beta_T \frac{Var(\bar{d}_g)}{\bar{d}} + \beta_C \frac{Var(\bar{d}_g)}{1 - \bar{d}}$, that is increasing in the variance of $\bar{d}_g$.[15] In the empirical application in Section 6, we further show that the resulting bias can have an economically significant effect in real-life settings.

Second, the bias is complicated. The bias depends on higher-order moments of $\bar{d}_g$ that are typically not reported in empirical studies. Estimates that ignore spillovers are not only wrong in magnitude, but can also have the wrong sign. For example, β_1 can be

---

[15]If no pure control ($\bar{d}_g = 0$) and no pure treatment ($\bar{d}_g = 1$) group exists, then the bias associated with $Var(\bar{d}_g)$ in Eq. (3) can be avoided by using sampling weights (see Baird et al., 2018, for example). More specifically, using weights $1/\bar{d}_g$ for treated units and weights $1/(1 - \bar{d}_g)$ for control units offsets the overrepresentation of treated units in high-treatment groups and the overrepresentation of control units in low-treatment groups. Our target in this section is, however, not to provide an estimator that avoids the bias, but rather to show the bias that arises when estimating a model without spillovers in a setting where spillovers are present.

negative, but the positive variance-terms in Eq. (3) can lead to a positive $\hat{\tilde{\beta}}_1$.

Researchers might be inclined to use group fixed effects in order to focus on within-group variation in treatment status. The following proposition shows that group fixed effects do not necessarily decrease the bias in $\tilde{\beta}_1$:

Proposition 2 *Assume y_i follows Eq. (1) and $\bar{d}_g \notin \{0, 1\}$ for at least one group g. Estimating Eq. (2) with group fixed effects yields:*

$$E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + (\beta_T - \beta_C) \left[\bar{d} + \frac{Var(\bar{d}_g)}{\bar{d}} - \underbrace{\frac{\bar{d}E\left(\bar{d}_g^3\right) - \left(E\left(\bar{d}_g^2\right)\right)^2}{\bar{d}\left(\bar{d} - E\left(\bar{d}_g^2\right)\right)}}_{=:\theta}\right], \qquad (4)$$

$$\text{with } 0 \le \theta \le \bar{d} + \frac{Var(\bar{d}_g)}{\bar{d}}.$$

Proof: See Appendix A.1.

Three facts stand out. First, using fixed effects removes any bias in estimating β_1 if spillover effects are homogenous ($\beta_T = \beta_C$). This is intuitive because within-group treatment and control units are subject to the same treatment intensity $\bar{d}_g$.

Second, with heterogenous spillover effects ($\beta_T \ne \beta_C$), the bias is more convoluted than in Proposition 1 and includes the third moment of the distribution of group means.[16] This is likely to be hard to interpret. Specifically, we are not aware of studies that provide higher-order moments of group-level mean treatment intensities. However, these moments would be necessary to understand the bias when using fixed effects in set-ups with spillover effects.

Third, the bias in estimating β_1 can be larger with fixed effects than without. This is because spillovers occur within groups, but not across groups. As the fixed effects estimator focuses on within-group variation, but discards across-group variation, the bias induced by spillover effects can be amplified with fixed effects estimators.[17]

---

[16]Fixed effects regressions are econometrically equal to regressions on de-meaned variables. Given that the mean of the outcome variable y_i is a non-linear function of the treatment intensity $\bar{d}_g$ (see Section 4.3 below), the bias includes the third-order moment of $\bar{d}_g$.

[17]To see this in formulas Eq. (3) and Eq. (4), assume there are some groups with $\bar{d}_g = k$ for fixed $k \in (0, 1)$ and add an arbitrarily large number of pure-control groups ($\bar{d}_g = 0$). In this case, the spillover-induced bias associated with β_C converges to zero in Eq. (3) (because $\bar{d} \to 0$). With fixed effects, the pure control groups are absorbed, and the spillover-induced bias associated with β_C is equal to $\beta_C k$, which is larger than the corresponding bias in Eq. (3).

4.3. Evaluation of existing strategies in the literature

In light of the preceding discussion of biases through spillovers, we can now interpret the strategies applied in the empirical literature to address spillover concerns. In Section 2 (Panel B of Table 1), we discussed three main strategies: first, dropping controls where spillovers are most likely. Second, testing for spillovers within control group units. Third, comparing estimates from a regression on the individual level with estimtates from aggregate-level regressions.

Dropping controls where spillovers are most likely can be interpreted as a regression $y_{ig} = \tilde{\beta}_0 + \tilde{\beta}_1 d_{ig} + \epsilon_{ig}$ using the entire sample of treated units but dropping control units with $\overline{d}_{ig} > 0$. This results in an estimate:[18]

$$E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + \beta_T \overline{d} + \beta_T \frac{Var(\overline{d}_g)}{\overline{d}} \tag{5}$$

Thus, for this strategy to provide an unbiased estimate of β_1, we need to assume $\beta_T = 0$, i.e., there are no spillovers to treated units.

Testing for spillovers within control group units can be interpreted as testing for $\beta_C = 0$ in the regression $y_{ig} = \tilde{\beta}_0 + \tilde{\beta}_C \overline{d}_g + \epsilon_{ig}$ using the sample of control group units only. If $\beta_C = 0$, this results in the same estimate as Eq. (5); hence, as above, testing for $\beta_C = 0$ is a potential avenue to deal with spillover concerns, but requires the assumption $\beta_T = 0$.

In settings where spillovers and aggregate effects are of interest, another common approach in the literature is a two-step procedure: Regressions on a disaggregated level provide estimates of the direct treatment effect. Regressions on an aggregated level are used to inform the researcher about spillovers and aggregated effects:

$$y_{ig} = \tilde{\beta}_0 + \tilde{\beta}_1 d_{ig} + \tilde{\epsilon}_{ig}, \tag{6}$$

$$\overline{y}_g = \gamma_0 + \gamma_1 \overline{d}_g + u_g. \tag{7}$$

This approach faces two challenges: First, $\tilde{\beta}_1$ does not measure the direct effect as

---

[18] $E\left[\hat{\tilde{\beta}}_1\right] = E\left[y_{ig}|d_{ig}=1\right] - E\left[y_{ig}|d_{ig}=0, \overline{d}_g=0\right] = \beta_1 + \beta_T \overline{d} + \beta_T (E\left[\overline{d}_g|d_{ig}=1\right] - \overline{d}) = \beta_1 + \beta_T \overline{d} + \beta_T \frac{Var(\overline{d}_g)}{\overline{d}}$.

discussed in detail above. Second, γ_1 either disproportionately captures spillovers to controls (β_C) or spillovers to treated units (β_T) depending on the distribution of $\overline{d}_g$. To see this, note that the group-level average in Eq. (1) can be written as $E\left[\overline{y}_g\right] = \beta_0 + (\beta_1 + \beta_C)\overline{d}_g + (\beta_T - \beta_C)\overline{d}_g^2$.[19] It follows that

$$\frac{\partial E\left[\overline{y}_g\right]}{\partial \overline{d}_g} = \begin{cases} \beta_1 + \beta_C & \text{for } \overline{d}_g = 0 \\ \beta_1 + \beta_T & \text{for } \overline{d}_g = 0.5 \\ \beta_1 - \beta_C + 2\beta_T & \text{for } \overline{d}_g = 1 \end{cases} \tag{8}$$

Thus, if $\overline{d}_g$ is close to zero, the slope of the regression on the aggregate level is determined by β_1 and β_C. On the other hand, for larger $\overline{d}_g$ the slope of the regression on the aggregate level is mainly determined by β_1 and β_T. This is intuitive: if group level average treatment intensities are low, then control units dominate and spillovers to control units dominate the group level average, and vice versa. This suggests that researchers looking at the same data generating process Eq. (1) will get strikingly different slopes for regressions on the aggregate level depending on the value of $\overline{d}_g$ in the data set at hand.[20] Only for the homogenous case, $\beta_C = \beta_T$, Eq. (8) is equivalent to $\partial E\left[\overline{y}_g\right]/\partial \overline{d}_g = \beta_1 + \beta_C$, so that the slope of the aggregate-level regression is independent of the average treatment intensities $\overline{d}_g$.

Taken together, the strategies applied in the literature are viable under specific assumptions, though these assumptions are not identical for the three strategies discussed above. While the first two strategies require the assumption $\beta_T = 0$ (no spillovers to treated units), the last strategy requires the assumption $\beta_T = \beta_C$. It is clear that both assumptions together can only be fulfilled if $\beta_T = \beta_C = 0$ (i.e., there are no spillovers). It is therefore important to check the plausibility of each assumption in the respective setting before deciding how to deal with spillover effects.

---

[19]The group level average is equal to $E\left[\overline{y}_g\right] = \overline{d}_g y(1, \overline{d}_g) + (1 - \overline{d}_g)y(0, \overline{d}_g) = \overline{d}_g(\beta_0 + \beta_1 + \beta_T\overline{d}_g) + (1 - \overline{d}_g)(\beta_0 + \beta_C\overline{d}_g) = \beta_0 + (\beta_1 + \beta_C)\overline{d}_g + (\beta_T - \beta_C)\overline{d}_g^2$.

[20]One might argue that this is just a matter of estimating the correct functional form: if the data generating model on the individual level is linear, then the model on the aggregate level is quadratic. Given that the linearity assumption on the individual level is ad-hoc, why worry about the functional form on the aggregate level. However, we would argue that regressions on the individual and aggregate level should be internally consistent with each other.

4.4. Example

In this subsection, we provide a stylized example that shows our effects in a simple way. Assume that some coffee shops are subject to a shock that increases marginal costs from $10 to $14 per pound of coffee beans (e.g., due to a rise in the input price of a bean producer delivering half of the shops). Without interaction between coffee shops, the raise in the input price leads to a raise in the final price per cup of coffee for treated shops, and thereby to a reduction in their sales, whereas the price and the sales of control shops remain unchanged. However, in this situation, as well as in many other economic situations, spillovers—e.g., due to competition—are likely to exist. In particular, since a portion of competitors increase their prices, some consumers start searching and may switch coffee shops. For a coffee shop in the control group, this implies an increase in its sales. For a coffee shop in the treatment group, this implies that its sales reduction is lower if more shops in the region are treated as well. Let us assume that spillovers follow a simple homogenous spillover model and that coffee shops compete within but not across cities (g):[21]

$$y_{ig} = 10 - 4d_i + 3.6\bar{d}_g d_i + 3.6\bar{d}_g(1 - d_i). \tag{9}$$

The situation is depicted in case 1 of Fig. 1. The direct treatment effect is equal to $\beta_1 = -4$. Because spillovers are homogenous, the difference between control and treated coffee shops at the mean treatment intensity $\bar{d} = 0.5$ is also equal to -4 (difference between the black and orange square in the figure).

[Fig. 1 here]

Let us further assume that in half of the cities, 90% of coffee shops are subject to the shock, while in the other half of the cities 10% of coffee shops are subject to the shock. Treated coffee shops are therefore overrepresented in high-treatment cities $[E(\bar{d}_g|d_{ig} = 1) = 0.9 \cdot 90\% + 0.1 \cdot 10\% = 0.82]$, and the sales of treated coffee shops are equal to 8.952 $(10 - 4 + 3.6 \cdot 0.82)$. Control group coffee shops are underrepresented in high-treatment

---

[21]The spillover model in Eq. (9) can, for example, be obtained from a Cournot model with an inverse demand of $P = 110 - Q$, a marginal cost (net of shock) equal to 10, and 9 coffee shops. Using these numbers in the Salop model leads to a similar equation.

cities $[E(\bar{d}_g|d_{ig} = 0) = 0.9 \cdot 10\% + 0.1 \cdot 90\% = 0.18]$, and the sales of control group coffee shops are 10.648 $(10 + 3.6 \cdot 0.18)$. Thus, when spillovers are ignored, the treatment effect has the same sign, but a different magnitude (difference between the black and the orange triangle in the figure: $\tilde{\beta}_1 = 8.952 - 10.648 = -1.696$ versus the true direct treatment effect of $\beta_1 = -4$). Specifically, the true direct treatment effect is larger compared to the case when spillovers ignored.

More generally, going beyond the specific numeric example Eq. (9), ignoring spillovers can also lead to wrong conclusions about the *direction* of the treatment effect. First, estimates that ignore spillovers can have the wrong sign even when spillovers are homogenous. This can happen when the spillover coefficients (β_T, β_C) are large in magnitude relative to the direct treatment effect (β_1). An example is depicted in case 2 of Fig. 1: while the direct treatment effect is negative, a researcher who ignores spillovers estimates a positive treatment effect $(\tilde{\beta}_1)$.[22] As an example, assume that firms are subject to a new regulation (such as an increase in reporting requirements, or stricter environmental standards). For an individual firm, the requirement might be burdensome and create a competitive disadvantage. However, outcomes might still be better when all firms are treated than when all firms are not treated.[23] In such a setting, a researcher who ignores spillovers can estimate negative, zero, or positive treatment effects depending on the data at hand, even if the underlying model is exactly the same.

Second, when spillovers are heterogeneous, intuitive statements on the bounds of treatment effects can be invalid. As an example, assume that a negative shock is accompanied by spillovers that widen the difference between treated and control units (see case 3 in Fig. 1). Researchers might argue that ignoring spillover effects results in a lower bound for the direct treatment effect. However, treatment estimates that ignore spillovers might actually be closer to zero than the direct treatment effect (β_1 vs $\tilde{\beta}_1$ in the figure). This is because spillovers are governed by the terms $\beta_T \cdot \bar{d}_g$ and $\beta_C \cdot \bar{d}_g$, and even if $\beta_T < \beta_C$, $\bar{d}_g$ will usually be larger for treated units than for control units because $\bar{d}_g$ is larger in

---

[22]More formally, in the case of homogenous spillovers $(\beta_T = \beta_C)$, Proposition 1 simplifies to $E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + \beta_T \frac{Var(\bar{d}_g)}{\bar{d}(1-\bar{d})}$. Since $\frac{Var(\bar{d}_g)}{\bar{d}(1-\bar{d})} \in [0, 1]$, the sign of the treatment effect estimate can flip if $\beta_T > -\beta_1$.

[23]See e.g., Fudenberg and Tirole (1984), Bulow et al. (1985), and Cabral and Villas-Boas (2005).

high-treatment regions.

5. Guidance for empirical researchers

Our discussion in Sections 3 and 4 can be summarized in the following three-step guidance.

Step 1: Consider the most plausible economic mechanism for spillovers (e.g., competition, demand, or agglomeration spillovers), the nature of the shock (e.g., shock on marginal costs or shock on capacity), as well as the most plausible dimension for spillovers (i.e., the groups g in the preceding discussion, such as industry or region). Guidance ultimately has to come from economic theory and from institutional knowledge of the setting at hand.

Step 2: Estimate the heterogenous spillover model

$$y_{ig} = \beta_0 + \beta_1 d_{ig} + \beta_T \overline{d}_g d_{ig} + \beta_C \overline{d}_g (1 - d_{ig}) + \epsilon_{ig}. \qquad (10)$$

where d_{ig} denotes the treatment indicator and $\overline{d}_g$ denotes the fraction of treated firms in group g.[24] Note that Eq. (10) should be estimated without group fixed effects (unit-level or group-level controls can be added if appropriate). Furthermore, estimating Eq. (10) is based on the assumption that d_{ig} and $\overline{d}_g$ are exogenous. In some models, a homogenous spillover model can occur. However, without a strong prior about the nature of the model and the nature of the shock, the safer route is to allow for heterogenous spillovers.

Step 3: Using the coefficients β_0, β_1, β_T and β_C [for Eq. (10)], plot the outcome variable as a function of the treatment intensity, i.e., $E\left[y_{ig}|\overline{d}_g\right]$, separately for the treatment units, the control units and the group level averages :

$$E\left[y_C|\overline{d}_g\right] = \beta_0 + \beta_C \overline{d}_g \qquad (11)$$

$$E\left[y_T|\overline{d}_g\right] = \beta_0 + \beta_1 + \beta_T \overline{d}_g \qquad (12)$$

$$E\left[\overline{y}_g|\overline{d}_g\right] = \beta_0 + (\beta_1 + \beta_C)\overline{d}_g + (\beta_T - \beta_C)\overline{d}_g^2. \qquad (13)$$

---

[24]Note that the fraction of treated units, $\overline{d}_g$, is not necessarily an equal weighted average of the treatment indicator d_i. Conceptually, $\overline{d}_g$ should reflect the market share of treated firms in competition models, the share of investment/R&D expenses of treated firms in agglomeration models, and the share of employment or the share of the wage bill of treated firms in demand spillover models.

The resulting figure (see Fig. 2 in the next section for an example) provides information about all key treatment effects: The difference between the graphs for the treated and control units at $\bar{d}_g = 0$ provides an estimate of the direct treatment effect β_1. The graph for the group level averages provides the aggregate effect if all groups are treated with a treatment intensity of $\bar{d}_g$. The difference between the graph for treated units and the graph for control units provides an estimate for the treatment-minus-control effect at various levels of $\bar{d}_g$.

In some settings, spillovers may affect firms in a non-linear way. For example, while the relationship $y_{ig} = f(d_{ig}, \bar{d}_g, d_{ig}\bar{d}_g)$ is linear in our basic workhorse models with marginal cost shocks, it can be non-linear in models with a shock to capacity (see Section 3 and the Online Appendix for details).

If the spillovers affect firms in a non-linear way, there are three possible solutions: first, researchers might still use a linear model as an approximation as long as f is continuous – which is always the case in the models discussed in Section 3. Second, if researchers have a strong prior on the exact model, they can estimate the relationship $y_{ig} = f(d_{ig}, \bar{d}_g, d_{ig}\bar{d}_g)$ using the functional form derived from theory. Third, f can be estimated non-parametrically. A non-parametric estimate of f minimizes specification error, but comes at the expense of requiring a larger sample to obtain a precise estimate. A simple scatter plot of (11)-(13) seems like a reasonable exploratory technique to assess whether the pattern appears linear or not (see Fan and Gijbels (1996)).

6. Application: credit supply and employment

We apply our framework to the setting in Huber (2018) who analyzes the causal effect of exposure to a bank lending cut on firms and counties. In particular, he examines the effects of the lending cut by a large German bank, Commerzbank (henceforth CB), during the financial crisis of 2008-2009 and argues that this event represents an exogenous shock to its German borrowers. We use this setting as the research question is of general importance. Further, the careful documentation and execution of the study lends itself to replication and extension.

CB decreased lending primarily as a result of losses suffered on its trading portfolio. Specifically, trading losses were due to investments in asset-backed securities related to the U.S. subprime mortgage market and its exposure to the insolvencies of Lehman Brothers and large Icelandic banks. Huber (2018) provides evidence suggesting that the losses were unrelated to CB's domestic loan portfolio, supporting the conjecture that the lending cut constitutes an exogenous event from the perspective of any given firm. In the presence of credit market frictions, such as switching costs in long-term lending relationships (Sharpe, 1990), lending cuts can negatively affect borrowers and, e.g., result in decreased employment or investment. We refer the reader to Huber (2018) for an in-depth discussion of institutional details, potential endogeneity concerns, and the effect of the lending cut on credit availability for CB dependent firms.

Given our interest in spillover effects, we focus on the regional variation of the employment effects resulting from CB's lending cut. Huber (2018) investigates the existence of "indirect effects" at the county level. Specifically, he tests if the negative effect of the lending cut on employment is increasing in the CB dependence of other firms in the same county, while keeping constant the firms' direct exposure. We re-visit this evidence and apply our framework introduced in the preceding sections.

6.1. Data

We follow Huber (2018) in the data collection and processing as close as possible. Firm level data for German public and private firms is obtained from Bureau van Dijk's AMADEUS database. Information on bank-firm relationships comes from AMADEUS BANKERS. We apply the same filters and restrictions as Huber (2018). We restrict the dataset to German firms with non-missing information on the number of employees in 2007.[25] We further require information on firms' date of incorporation, county, and industry to be available to construct basic firm level control variables. We drop firms in the

---

[25] Huber (2018) uses 2006 as the base year to define control variables. We base our analysis on a 2018 snapshot of AMADEUS data obtained via WRDS. AMADEUS provides at most 10 recent (fiscal) years of data for the same company (Kalemli-Ozcan et al., 2015), i.e., coverage is (reasonably) complete from 2007 onwards.

financial and public sectors,[26] and restrict the sample to firms with available information on their relationship banks.[27] The final sample comprises 23,436 firms.

We follow Huber (2018) and calculate the employment change from 2008 to 2012 (symmetric growth rate) for each firm. We define a variable, $CB\ dep_{ic}$, as the fraction of firm i's lending relationships that are with CB out of the firm's total number of relationship banks. We define $\overline{CB\ dep_{ic}}$ for each firm i as the average CB dependence of all other firms in the same county (c), excluding firm i itself.[28] For robustness, we additionally define an indicator variable, $CB\ dep\ (0/1)_{ic}$, that equals one if $CB\ dep_{ic}$ is greater or equal to 0.5, and zero otherwise.

[Table 3 here]

Table 3 shows summary statistics for the final sample. Firms have an average of two relationship banks. Consistent with Huber (2018), the average CB dependence is about 0.17. The average number of employees is 177 and the average firm age is about 23 years.

6.2.   Baseline results

We start with a baseline estimation of potential indirect effects on firms in counties with a high CB dependence, following Huber (2018). In particular, we estimate the following model:

$$employment\ growth_{ic} = \beta_0\ +\ \beta_1 \text{CB}\ dep_{ic} + \beta_2\overline{\text{CB}\ dep_{ic}} + \gamma'X_{ic} + \epsilon_{ic}, \qquad (14)$$

---

[26]Specifically, we exclude financial services and related industries, including holding companies (NACE codes 65-70), industries that are mainly public sector in Germany, i.e., administrative services, education, healthcare, and arts & culture (NACE codes 81, 82, 84-88, and 90-92), and activities of organizations and private households (NACE codes 94 and 97-99).

[27]Huber (2018) uses proprietary data on historical lending relationships of German firms and fixes lending relationships in 2006. The difference in the base year is because we use a 2018 snapshot of the AMADEUS BANKERS database, which only reports current lending relationships. However, firm-bank relationships are extremely sticky [cf. Giannetti and Ongena (2012) and Kalemli-Ozcan et al. (2018) who compare different vintages of AMADEUS BANKERS]. While using a 2018 snapshot of firm-bank relationships may introduce noise in the estimation, our baseline estimates are very close to those reported in Huber (2018).

[28]We follow Huber (2018) in using an equal weighting to construct $\overline{CB\ dep_{ic}}$. It is plausible that larger firms induce higher (demand) spillovers than smaller firms, which might call for a weighting of $\overline{CB\ dep_{ic}}$ by the number of employees. At the same time, credit supply shocks affect smaller firms more than larger firms [see Berg (2018) for the size-dependency of credit supply shocks in Germany during and after the financial crisis] as larger firms can better substitute to other funding sources.

where *employment growth* is the firm's symmetric employment growth rate over the 2008 to 2012 period, defined as: $2 * (employment_{2012} - employment_{2008})/(employment_{2012} + employment_{2008})$. As noted above, CB dep_{ic} measures the CB dependence of firm i, located in county c. $\overline{\text{CB } dep_{ic}}$ is the average CB dependence of all other firms located in the same county c, excluding firm i itself. X is a set of firm controls, defined following Huber (2018). In particular, we include indicator variables for 4 firm size bins (1-49, 50-249, 250-999, and $> 1,000$ employees as of 2007), the log age as of 2007, and industry indicators (2-digit NACE code). Standard errors are clustered at the county level. Results are shown in Table 4.

[Table 4 here]

Our results are broadly consistent with Huber (2018), whose baseline estimates we report in column 4 for comparison. The coefficient on *CB dep_{ic}* ranges between -0.019 and -0.033 (-0.03 in Huber, 2018) and the coefficient on $\overline{\text{CB } dep_{ic}}$ is -0.155 (-0.166 in Huber, 2018).[29] Using the estimates from column 3, these results imply that full CB dependence reduces employment growth by about 2 percentage points (p.p.) for a firm in a county where no other firm had CB among their relationship banks (the "direct effect"). This effect is amplified by CB dependence of other firms in the region: a one standard deviation (SD) greater CB dependence of other firms (6%, cf. Table 3) reduces employment growth by $6\% \times 0.155 \approx 1$ p.p. more. Given the implicit assumption of homogenous spillover effects for treated and control group firms, -1 p.p. is also the indirect effect on firms not dependent on CB.

---

[29]The fact that the coefficient on the average CB dependence ($\overline{\text{CB } dep_{ic}}$) is very close to the estimate in Huber (2018) while the coefficient estimate on the individual CB dependence (*CB dep_{ic}*) is closer to zero is consistent with our measure of CB dependence being somewhat noisy given that we rely on more recent bank-firm relationship data.

6.3. Full spillover model

Next, we amend the model and allow for heterogenous effects for treated and control group firms, i.e., we estimate the flexible model introduced in Section 4:

$$
\begin{aligned}
employment\ growth_{ic} = {} & \beta_0 + \beta_1 \text{CB}\ dep_{ic} + \beta_T \overline{\text{CB}\ dep_{ic}} \times \text{CB}\ dep_{ic} \\
& + \beta_C \overline{\text{CB}\ dep_{ic}} \times (1 - \text{CB}\ dep_{ic}) + \gamma' X_{ic} + \epsilon_{ic}.
\end{aligned}
\tag{15}
$$

This model relies on the four key assumptions A.1-A.4 discussed in Section 4.1: conditional independence assumption (CIA), outcomes depend on individual treatment status and treatment intensity in the same region (exchangeability), no spillovers across regions (partial interference), and linearity. The CIA is discussed in-depth in Huber (2018). The exchangeability assumption relies on spillovers not depending on the specific identity of the treated neighbors. This assumption is plausible for demand spillovers if the composition of local demand is similar for employees at different firms. Partial interference is also plausible for local demand spillovers, particularly for non-tradable sectors. We return to this point in the analyses below. The relationship at hand is plausibly continuous (no cliff effects) and we use a linear model as a first-order approximation.

As explained in Section 3, our theoretical foundation can provide guidance on the potential channels through which credit supply shocks can create externalities. A particularly plausible spillover channel, also emphasized in Huber (2018), is the existence of local demand effects. That is, a negative credit supply shock for (some) firms in a region can lead to a demand contraction with negative externalities.

We provide a simple model with local demand effects in the Online Appendix (see also the discussion in Section 3).[30] The model makes the following key predictions: i) a negative shock on firms' marginal cost or production capacity (here: due to a credit supply contraction) will result in a negative direct effect on employment for treated (here: CB dependent) firms. ii) A drop in employment leads to a demand contraction, which

---

[30]We use a simple local demand model (based on Murphy et al., 1989) with linear demand and a 1:1 relationship between output and employment. However, we also checked richer set-ups with e.g. more complex production functions and obtained very similar results. In particular, in all set-ups, the spillover effects are negative and heterogenous, i.e., firms in the control group are affected more strongly than firms in the treatment group.

puts additional pressure on employment in the economy. iii) This effect is heterogenous. If treatment is a binding shock on production capacity, this is the main constraint for treated firms. In particular, treated firms produce at the capacity constraint, irrespective of local demand conditions. By contrast, control group firms (here: firms not dependent on CB) are affected only through a drop in local demand, i.e., control group firms experience a negative indirect effect.

In summary, theory predicts heterogenous spillover effects in the form of $\beta_C < \beta_T \leq 0$ in the presence of local demand effects. We test this prediction by estimating Eq. (15). The results are shown in Table 5.

[Table 5 here]

The results indicate that spillover effects are indeed driven by control group firms. Consistent with the theoretical predictions of the local demand model, β_C is negative and highly statistically significant, while β_T is close to zero and insignificant (column 3). The coefficient β_C indicates that a one SD greater CB dependence of other firms would result in an employment growth reduction by about $6\% \times 0.184 \approx 1.1$ p.p. for firms with no relationship to CB.

The effect for a firm fully dependent on CB is -5.5 p.p. irrespective of the CB dependence of other firms in the county (column 3). This "direct effect" is more than twice as large compared to the estimate that does not account for spillover effects (column 1), i.e., estimating Eq. (2). The difference in effect size is, if anything, even more pronounced relative to a model that only allows for heterogenous spillover effects (column 2). Given that $\beta_C < 0$, the fact that estimating Eq. (14) results in an underestimation of β_1 follows directly from Proposition 1. This highlights that not accounting for spillover effects, or assuming that effects are homogenous, can result in a severe over- or underestimation of the direct treatment effect. Results are similar when we use an indicator variable that equals one if CB $dep_{ic} \geq 0.5$, and zero otherwise (columns 4 to 6).

[Fig. 2 here]

Fig. 2 depicts the county level spillover effects using Eq. (11) - (13) and the estimates from column 3 of Table 5. The figure illustrates several points. First, one can easily

read off the direct effect, i.e., the implied employment growth at a treatment fraction of zero (−5.5 p.p.). Second, the difference between treatment and control units diminishes quickly with increasing county level treatment as result of heterogenous spillovers. This visualizes why not accounting for heterogenous effects leads to an underestimation of the direct effect. Third, the employment decline is similar for treatment and control units in counties with a treatment fraction of ∼0.3.[31] Fourth, the relationship between average county employment growth and average county CB dependence is non-linear as implied by Eq. (8).

Highlighting that credit supply-induced spillover effects can be different for treated and control group firms is of economic importance. First, this result fosters our understanding of the propagation of credit supply shocks through the economy. Heterogenous spillover effects are consistent with credit supply-induced local demand effects. We provide further empirical evidence in support of this channel in the next subsection. Second, the estimates can also be used to assess the relative contribution of direct and indirect effects. The original results in Huber (2018) imply a ratio of indirect-to-direct effect of 5.5 to 1 in the average county, i.e., for each employee who loses her job as a result of her firm's CB dependence, 5.5 employees lose their jobs due to (demand) spillovers. Applying our framework reduces this ratio significantly: each direct job loss results in only additional 3 job losses due to (demand) spillovers in the average county.[32] This difference in results primarily stems from the incorporation of heterogenous spillover effects. Third, assessing the relative contribution of direct and indirect effects is important in order to formulate an effective policy response. For example, while direct effects can be addressed on the bank level (e.g., through a recapitalization), indirect effects are not internalized even by a well-capitalized bank.

---

[31] At a higher treatment fraction the implied effect for firms without CB dependence would even be below the effect for firms dependent on CB. This result, however, should be treated with caution given that the 95th percentile of the county CB dependence is 0.26 (cf. Table 3), i.e., there are very few observations above this value and the confidence intervals are large for treatment fractions beyond 0.3.

[32] The employment reduction resulting from direct effects is −0.48 p.p. in Huber (2018) for the average county (average CB dependence of 0.16×-0.03, cf. Table 4, column 4). The indirect effect is −2.66 p.p. (0.16×-0.166, cf. Table 4, column 4), giving a ratio of 5.53 to 1. The corresponding calculation using the estimates from Table 5, column 3, is: direct effect = −0.94 p.p. (0.17×-0.055) and indirect effect = −2.60 p.p. ($(1 - 0.17) \times 0.17 \times -0.184$), implying a ratio of 2.77 to 1. Note that an average CB dependence of 0.16 (0.17) is used in the former (latter) calculation to reflect the (marginal) differences in summary statistics, see Table 1 in Huber (2018) vs. Table 3.

Overall, this discussion highlights that ignoring spillover effects or assuming that spillover effects are homogenous for treated and control group firms can lead to a biased (i.e., over- or under-) estimation of the direct treatment effect, as shown formally in Proposition 1. In this specific setting, our results suggest that the direct effect of CB's lending cut on affected firms reported by Huber (2018) likely underestimates the true effect. This is because spillovers are heterogenous: while control group firms are indirectly affected, the employment growth at treatment firms does not vary with county level CB dependence.

6.4. Cross-sectional variation

In this section, we seek to provide further evidence for the conjecture that indirect effects arise due to a local demand channel. We exploit that firms in tradeable sectors are less susceptible to local demand conditions than firms in non-tradeable sectors (see, e.g., Mian et al., 2020). We split the sample into tradeable and non-tradeable sectors based on NACE codes (Bertinelli et al., 2016).

[Table 6 here]

Table 6, column 1, reports baseline estimates removing sectors that cannot be classified as tradeable or non-tradeable. Results are virtually identical to Table 5, column 3. The subsample analysis is reported in columns 2 and 3. While standard errors increase in the subsamples, the coefficient on CB *dep*, i.e., the direct effect, is very similar for firms in tradeable and non-tradeable sectors (-5.3 vs. -6.5 p.p.). This is expected as also firms in the tradeable sector should be negatively affected if their own credit supply conditions deteriorate. More interestingly, the results indicate that the indirect effect for firms not dependent on CB is significantly more pronounced in the non-tradeable sector (coefficient of -0.245 vs. -0.122). This is consistent with the indirect effect working through a local demand channel that predominantly affects firms more susceptible to local conditions.

Results are even stronger when using an indicator variable that equals one if CB $dep_{ic} \geq$ 0.5, and zero otherwise (columns 4 to 6). The direct effect is almost the same and highly statistically significant in both subsamples (-5 vs. -5.7 p.p.). The indirect effect is

27

large and significant in the non-tradeable subsample (-0.165) and relatively small and not statistically significant in the tradeable subsample (-0.064).[33]

7. Conclusion

Spillover effects arise naturally in many corporate finance settings. Yet, despite their importance, the discussion of spillover effects in empirical research misses the rigor dedicated to endogeneity concerns. In this paper, we have analyzed of broad set of workhorse models of firm interactions—via competitive effects, demand spillovers, and agglomeration effects— and have discussed the implications on estimating treatment effects in empirical corporate finance.

Conceptually, we highlight two key results. First, even with random treatment, spillovers lead to a complicated bias in estimating treatment effects. The bias is convoluted and depends on second- or, in the case of fixed effects regressions, third-order moments of group-level treatment intensities. Simple rules (such as "divide the treatment effect by two" in case that control group units benefit from a negative shock to treated units) are insufficient to describe the resulting bias. Second, we document that including fixed effects, a common approach to strengthen identification in the presence of endogeneity concerns, can exacerbate the bias arising from spillovers.

We develop a simple guidance for empirical researchers, apply it to an empirical credit supply shock setting, and highlight differences in the results compared to current empirical practice. For example, we demonstrate that direct effects of a credit supply shock are underestimated by a factor of 2-3 using current practice. We hope that this guidance will be useful to academics in future research.

---

[33]While economically large, this difference is marginally not statistically significant. We also examined to what extent ex ante financial conditions of firms affect their exposure to credit supply shocks. As argued, among others, by Berg (2018) and Beck et al. (forthcoming), liquid firms may be able to (temporarily) counterbalance credit supply contractions by drawing on existing cash buffers. Consistent with this conjecture, we find evidence that both direct and indirect effects are stronger for firms that have ex ante lower cash to assets ratios (untabulated).

A.   Proof of proposition 1 and 2

Proof of Proposition 1: For the following proofs, note that for a dummy variable d_{ig} the following equations hold:

$$Var(d_{ig}) = \bar{d}(1 - \bar{d}) \tag{16}$$

$$E(d_{ig}\bar{d}_g) = E(\bar{d}_g^2) \tag{17}$$

$$Cov(d_{ig}, \bar{d}_g) = E(\bar{d}_g) - \bar{d}^2 = Var(\bar{d}_g) \tag{18}$$

$$Cov(d_{ig}, d_{ig}\bar{d}_g) = E(d_{ig}\bar{d}_g) - E(d_{ig}) \cdot E(d_{ig}\bar{d}_g) = E(\bar{d}_g^2)(1 - \bar{d}) \tag{19}$$

$$= \left[Var(\bar{d}_g) + \bar{d}^2)\right](1 - \bar{d}) \tag{20}$$

Using Eq. (16) - (20) and the standard omitted-variable bias formula yields:

$$E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + \beta_T \frac{Cov(d_{ig}, d_{ig}\bar{d}_g)}{Var(d_{ig})} + \beta_C \frac{Cov(d_{ig}, (1 - d_{ig})\bar{d}_g)}{Var(d_{ig})} \tag{21}$$

$$= \beta_1 + \beta_T \cdot \left[\frac{Var(\bar{d}_g)}{\bar{d}} + \bar{d}\right] + \beta_C \cdot \left[\frac{Var(\bar{d}_g)}{\bar{d}(1 - \bar{d})} - \frac{Var(\bar{d}_g)}{\bar{d}} - \bar{d}\right] \tag{22}$$

$$= \beta_1 + (\beta_T - \beta_C)\bar{d} + \beta_T \frac{Var(\bar{d}_g)}{\bar{d}} + \beta_C \frac{Var(\bar{d}_g)}{1 - \bar{d}} \tag{23}$$

Proof of Proposition 2: The proof proceeds in two steps:

Step 1: We show that the following relation holds:

$$E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + (\beta_T - \beta_C)\frac{E(\bar{d}_g^2) - E(\bar{d}_g^3)}{Var(d_{ig}) - Var(\bar{d}_g)} \tag{24}$$

$$= \beta_1 + (\beta_T - \beta_C)\left[\bar{d} + \frac{Var(\bar{d}_g)}{\bar{d}} - \underbrace{\frac{\bar{d}E\left(\bar{d}_g^3\right) - \left(E\left(\bar{d}_g^2\right)\right)^2}{\bar{d}\left(\bar{d} - E\left(\bar{d}_g^2\right)\right)}}_{=:\theta}\right] \tag{25}$$

To see Eq. (24), note that a regression with group fixed effects $y_i = \tilde{\beta}_1 d_{ig} + \gamma_g + \epsilon_i$ is equivalent to the de-meaned regression $y_{ig} - \bar{y}_g = \tilde{\beta}_1(d_{ig} - \bar{d}_g) + \epsilon_i$. For the following

29

steps, it helps to recognize that

$$y_{ig} - \overline{y}_g = \beta_1(d_{ig} - \overline{d}_g) + (\beta_T - \beta_C)(d_{ig}\overline{d}_g - \overline{d}_g^2) \tag{26}$$

Using $E(d_{ig} - \overline{d}_g) = 0$, Eq. (16) - (20), $E(d_{ig}\overline{d}_g^2) = E(\overline{d}_g^3)$, $Var\left(d_{ig} - \overline{d}_g\right) = Var(d_{ig}) - Var(\overline{d}_g)$ and the standard omitted-variable bias formula yields:

$$E\left[\hat{\hat{\beta}}_1\right] = \frac{Cov\left(y_{ig} - \overline{y}_g, d_{ig} - \overline{d}_g\right)}{Var\left(d_{ig} - \overline{d}_g\right)} \tag{27}$$

$$= \frac{Cov\left(\beta_1(d_{ig} - \overline{d}_g) + (\beta_T - \beta_C)(d_{ig}\overline{d}_g - \overline{d}_g^2), d_{ig} - \overline{d}_g\right)}{Var\left(d_{ig} - \overline{d}_g\right)} \tag{28}$$

$$= \beta_1 + (\beta_T - \beta_C)\frac{E\left(d_{ig}\overline{d}_g - 2d_{ig}\overline{d}_g^2 + \overline{d}_g^3\right)}{Var\left(d_{ig} - \overline{d}_g\right)} \tag{29}$$

$$= \beta_1 + (\beta_T - \beta_C)\frac{E(\overline{d}_g^2) - E(\overline{d}_g^3)}{Var(d_{ig}) - Var(\overline{d}_g)} \tag{30}$$

To see Eq. (24), note that $Var(d_{ig}) - Var(\overline{d}_g) = \overline{d} - E\left(\overline{d}_g^2\right)$, implying that Eq. (24) can also be written as

$$E\left[\hat{\hat{\beta}}_1\right] = \beta_1 + (\beta_T - \beta_C)\frac{E(\overline{d}_g^2) - E(\overline{d}_g^3)}{\overline{d} - E\left(\overline{d}_g^2\right)} \tag{31}$$

Adding $0 = \overline{d} + \frac{Var(\overline{d}_g)}{\overline{d}} - \overline{d} - \frac{Var(\overline{d}_g)}{\overline{d}}$ yields:

$$E\left[\hat{\hat{\beta}}_1\right] = \beta_1 + (\beta_T - \beta_C)\left[\overline{d} + \frac{Var(\overline{d}_g)}{\overline{d}} + \frac{\left(-\overline{d} - \frac{Var(\overline{d}_g)}{\overline{d}}\right)d\left(d - E\left(\overline{d}_g^2\right)\right) + E(\overline{d}_g^2) - E(\overline{d}_g^3)}{\overline{d}\left(\overline{d} - E\left(\overline{d}_g^2\right)\right)}\right] \tag{32}$$

$$= \beta_1 + (\beta_T - \beta_C)\left[\overline{d} + \frac{Var(\overline{d}_g)}{\overline{d}} - \frac{\overline{d}E\left(\overline{d}_g^3\right) - \left(E\left(\overline{d}_g^2\right)\right)^2}{\overline{d}\left(\overline{d} - E\left(\overline{d}_g^2\right)\right)}\right] \tag{33}$$

$$= \beta_1 + (\beta_T - \beta_C)\left[\overline{d} + \frac{Var(\overline{d}_g)}{\overline{d}} - \theta\right] \tag{34}$$

Step 2: $0 \leq \theta \leq \overline{d} + \frac{Var(\overline{d}_g)}{\overline{d}}$

$\theta \leq \bar{d} + \frac{Var(\bar{d}_g)}{\bar{d}}$ follows directly from Eq. (24) and $\frac{E(\bar{d}_g^2) - E(\bar{d}_g^3)}{Var(d_{ig}) - Var(\bar{d}_g)} > 0$. To see that $\theta \geq 0$, first note that the denominator of θ is larger than zero. It remains to be shown that the nominator is larger than zero. To see this, using n as the total number of units and $\bar{d}_{ig}$ for the group-level average of d_{ig}, write

$$\bar{d} E\left(\bar{d}_g^3\right) - \left(E\left(\bar{d}_g^2\right)\right)^2 = \frac{1}{n^2}\left(\sum_i \bar{d}_{ig} \sum_i \bar{d}_{ig}^3 - \left(\sum_i \left[\bar{d}_{ig}^2\right]\right)^2\right) \tag{35}$$

$$= \frac{1}{n^2}\sum_{i,j} \bar{d}_{ig}\,\bar{d}_{jg}^3 - \bar{d}_{ig}^2 \bar{d}_{jg}^2 \tag{36}$$

$$= \frac{1}{n^2}\sum_i \sum_{j>i} \bar{d}_{ig}\,\bar{d}_{jg}\left(\bar{d}_{ig} - \bar{d}_{jg}\right)^2 > 0 \tag{37}$$

References

Acharya, V.V., Eisert, T., Eufinger, C., Hirsch, C., 2018. Real effects of the sovereign debt crisis in europe: Evidence from syndicated loans. Review of Financial Studies 31, 2855–2896.

Aiyar, S., 2012. From financial crisis to great recession: The role of globalized banks. American Economic Review 102, 225–230.

Ali, A., Klasa, S., Yeung, E., 2009. The limitations of industry concentration measures constructed with compustat data: Implications for finance research. Review of Financial Studies 22, 3839–3871.

Almeida, H., Campello, M., Laranjeira, B., Weisbenner, S., 2011. Corporate debt maturity and the real effects of the 2007 credit crisis. Critical Finance Review 1, 3–58.

Amiti, M., Weinstein, D., 2011. Exports and financial shocks. Quarterly Journal of Economics 126, 1841–1877.

Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 91, 444–472.

Angrist, J.D., Pischke, J.S., 2009. Mostly harmless econometrics. Princeton University Press .

Anselin, L., 1988. Spatial econometrics, methods and models. Springer Netherlands .

Baird, S., Bohren, A., McIntosh, C., Özler, B., 2018. Optimal design of experiments in the presence of interference. Review of Economics and Statistics 100, 844–860.

Beck, T., Da-Rocha-Lopes, S., Silva, A.F., forthcoming. Sharing the pain? Credit supply and real effects of bank bail-ins. Review of Financial Studies .

Berg, T., 2018. Got rejected? real effects of not getting a loan. Review of Financial Studies 31, 4912–4957.

Bernstein, S., Colonnelli, E., Giroud, X., Iverson, B., 2019. Bankruptcy spillovers. Journal of Financial Economics 133, 608–633.

Bertinelli, L., Cardi, O., Restout, R., 2016. Technical change biased toward the traded sector and labor market frictions. Working Paper, University of Luxembourg .

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust in difference-in-difference estimates? Quarterly Journal of Economics 119, 249–275.

Boehmer, E., Jones, C.M., Zhang, X., 2020. Potential pilot problems: Treatment spillovers in financial regulatory experiments. Journal of Financial Economics 135, 68–87.

Born, B., Breitung, J., 2011. Simple regression based tests for spatial dependence. The Econometrics Journal 14, 330–342.

Borusyak, K., Jaravel, X., 2017. Revisiting event study designs. Working Paper, London School of Economics .

Bowley, A.L., 1924. The Mathematical Groundwork of Economics. Oxford University Press, Oxford.

Bulow, J.I., Geanakoplos, J.D., Klemperer, P.D., 1985. Multimarket oligopoly: Strategic substitutes and complements. Journal of Political Economy 93, 488–511.

Bustamante, C., Frésard, L., 2020. Does firm investment respond to peers' investment? Management Science, forthcoming .

Cabral, L.M., Villas-Boas, M., 2005. Bertrand supertraps. Management Science 51, 1599–1613.

Chava, S., Purnanandam, A., 2011. The effect of banking crisis on bank-dependent borrowers. Journal of Financial Economics 99, 116–135.

Chodorow-Reich, G., 2014. The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis. Quarterly Journal of Economics 129, 1–59.

Chodorow-Reich, G., 2020. Regional data in macroeconomics: Some advice for practioners. Journal of Economic Dynamics and Control 115.

Cox, D.R., 1958. Planning of experiments. Wiley, New York .

d'Aspremont, C., Jacquemin, A., 1988. Cooperative and noncooperative r&d in duopoly with spillover. American Economic Review 78, 1133–1137.

De Haas, R., Horen, N.V., 2012. International shock transmission after the lehman brothers collapse: Evidence from syndicated lending. American Economic Review Papers & Proceedings 102, 231–237.

Duflo, E., Saez, E., 2003. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. Quarterly Journal of Economics 118, 815–842.

Duguay, R., Minis, M., Sutherland, A., 2020. Regulatory spillovers in common audit markets. Management Science 66, 3389–3411.

Elhorst, J.P., 2014. Spatial econometrics: From cross-sectional data to spatial panels. Springer, Berlin, Heidelberg .

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and its Applications. London: Chapman and Hall.

Farrell, J., Shapiro, C., 1990. Horizontal mergers: An equilibrium analysis. American Economic Review 80, 107–126.

Fudenberg, D., Tirole, J., 1984. The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. American Economic Review Papers & Proceedings 74, 361–366.

Gan, J., 2007. The real effects of asset market bubbles: Loan- and firm-level evidence of a lending channel. Review of Financial Studies 20, 1941–1973.

Giannetti, M., Ongena, S., 2012. "lending by example": Direct and indirect effects of foreign banks in emerging markets. Journal of International Economics 86, 167–180.

Goodman-Bacon, A., 2018. Difference-in-difference with variation in treatment timing. Working Paper, Vanderbilt University .

Gormley, T.A., Matsa, D.A., 2014. Common errors: How to (and not to) control for unobserved heterogeneity. Review of Financial Studies 27, 617–661.

Grieser, W.D., Hadlock, C.J., 2019. Panel-data estimation in finance: Testable assumptions and parameter (in)consistency. Journal of Financial and Quantitative Analysis 54, 1–29.

Heckman, J., Locner, L., Taber, C., 1998. General equilibrium treatment effects: A study of tuition policy. NBER Working Paper No. 6426 .

Hennessy, C., Levy, A., Whited, T.M., 2007. Testing q theory with financing frictions. Journal of Financial Economics 83, 691–717.

Hennessy, C.A., Whited, T.M., 2005. Debt dynamics. Journal of Finance 60, 1129–1165.

Hoover, E.M., 1948. The Location of Economic Activity. New York: McGraw-Hill.

Hornbeck, R., Keniston, D., 2017. Creative destruction: Barriers to urban growth and the boston fire of 1872. American Economic Review 107, 1365–1398.

Huber, K., 2018. Disentangling the effects of a banking crisis: Evidence from german firms and counties. American Economic Review 108, 868–898.

Ivashina, V., Scharfstein, D., 2010. Bank lending during the financial crisis of 2008. Journal of Financial Economics 97, 319–338.

Iyer, R., Peydró, J.L., da Rocha-Lopes, S., Schoar, A., 2014. Interbank liquidity crunch and the firm credit crunch: Evidence from the 2007-2009 crisis. Review of Financial Studies 27, 347–372.

Jiang, W., 2017. Have instrumental variables brought us closer to the truth. The Review of Corporate Finance Studies 6, 127–140.

Kalemli-Ozcan, S., Laeven, L., Moreno, D., 2018. Debt overhang, rollover risk, and corporate investment: Evidence from the european crisis. NBER Working Paper No. 24555 .

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., Yesiltas, S., 2015. How to construct nationally representative firm level data from the orbis global database. NBER Working Paper No. 21558 .

Khwaja, A.I., Mian, A., 2008. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. American Economic Review 98, 1413–1442.

Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99, 481–499.

Maggio, M.D., Kermani, A., Keys, B., Piskorski, T., Ramcharan, R., Seru, A., Yao, V., 2017. Interest rate pass-throug: Mortgage rates, household consumption, and voluntary deleveraging. American Economic Review 107, 3550–3588.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. Review of Economic Studies 60, 531–542.

Marshall, A., 1890. Principles of Economics. London: McMillan.

Mian, A.R., Sufi, A., Verner, E., 2020. How does credit supply expansion affect the real economy? The productive capacity and household demand channels. Journal of Finance 75, 949–994.

Miguel, E., Kremer, M., 2004. Worms: Identifying impacts on education and health in the presence of treatment externalities. Econometrica 72, 159–217.

Murphy, K.M., Shleifer, A., Vishny, R.W., 1989. Industrialization and the big push. Journal of Political Economy 97, 1003–1026.

Nakamura, E., Steinsson, J., 2018. Identification in macroeconomics. Journal of Economic Perspectives 32, 59–86.

Nocke, V., Whinston, M.D., 2010. Dynamic merger review. Journal of Political Economy 96, 1200–1251.

Paravisini, D., Rappoport, V., Schnabl, P., Wolfenzon, D., 2014. Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. Review of Economic Studies 81, 333–359.

Peek, J., Rosengren, E., 1997. The international transmission of financial shocks. American Economic Review 85, 495–505.

Peek, J., Rosengren, E., 2000. Collateral damage: Effects of the japanese bank crisis on real activity in the united states. American Economic Review 90, 30–45.

Petersen, M., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Review of Financial Studies 22, 435–480.

Roberts, M.R., Whited, T.M., 2012. Endogeneity in empirical corporate finance, in: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier. volume 2.

Rubin, D.B., 1978. Bayesian inference for causal effects: The role of randomization. Annals of Statistics 6, 34–58.

Salop, S.C., 1979. Monopolistic competition with outside goods. Bell Journal of Economics and Management Science 10, 141–156.

Santos, J., 2011. Bank corporate loan pricing following the subprime crisis. Review of Financial Studies 24, 1916–1943.

Sharpe, S.A., 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. Journal of Finance 45, 1069–1087.

Shleifer, A., Vishny, R.W., 1988. The efficiency of investment in the presence of aggregate demand spillovers. Journal of Political Economy 96, 1221–1231.

Shue, K., Townsend, R.R., 2017. Growth through rigidity: An explanation for the rise in ceo pay. Journal of Financial Economics 123, 1–21.

Singh, N., Vives, X., 1984. Price and quantity competition in a differentiated duopoly. RAND Journal of Economics 15, 546–554.

Sraer, D.A., Thesmar, D., 2019. A sufficient statistics approach for aggregating firm-level experiments. Working Paper, University of California, Berkeley .

Sufi, A., 2009. The real effects of debt certification: Evidence from the introduction of bank loan ratings. Review of Financial Studies 22, 1659–1691.

Vazquez-Bare, G., 2018. Identification and estimation of spillover effects in randomized experiments. University of California, Santa Barbara Working Paper .

Fig. 1: Illustration of the bias arising from spillover effects

This figure illustrates the bias that arises when spillover effects exist, but are ignored in the estimation of treatment effects. All graphs are based on the linear spillover model, Eq. (1), and we assume that half of the groups have a treatment intensity of $\bar{d}_g = 90\%$ and half of $\bar{d}_g = 10\%$. Case 1 provides a scenario with homogenous spillovers ($\beta_0 = 10, \beta_1 = -4, \beta_T = \beta_C = 3.6$). Case 2 provides a scenario with larger, but still homogenous, spillovers ($\beta_0 = 10, \beta_1 = -4, \beta_T = \beta_C = 8$). Case 3 provides a scenario with heterogenous spillovers ($\beta_0 = 10, \beta_1 = -4, \beta_T = 2, \beta_C = 4$).
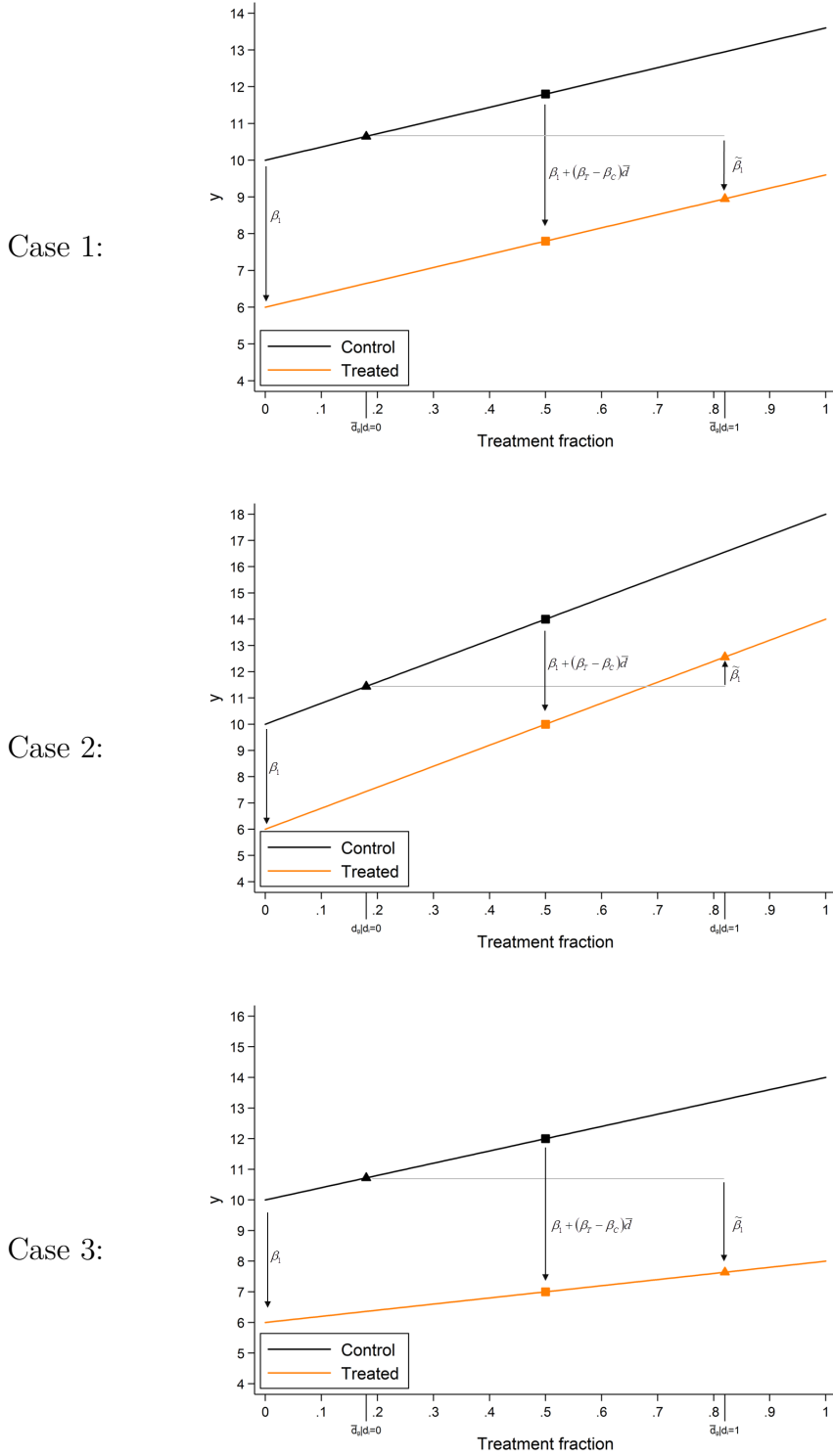
Fig. 2: Commerzbank's lending cut and spillover effects at the county level

This figure illustrates the county level spillover effects of Commerzbank's lending cut on firms with and without Commerzbank dependence. In particular, the figure plots employment growth from 2008 to 2012 as a function of the average Commerzbank dependence of a county using equations Eq. (11) - (12) and the estimated coefficients from Table 5, column 3. Further shown are 90% confidence intervals and the county level average employment growth (*y avg.*; cf. Eq. (13)).
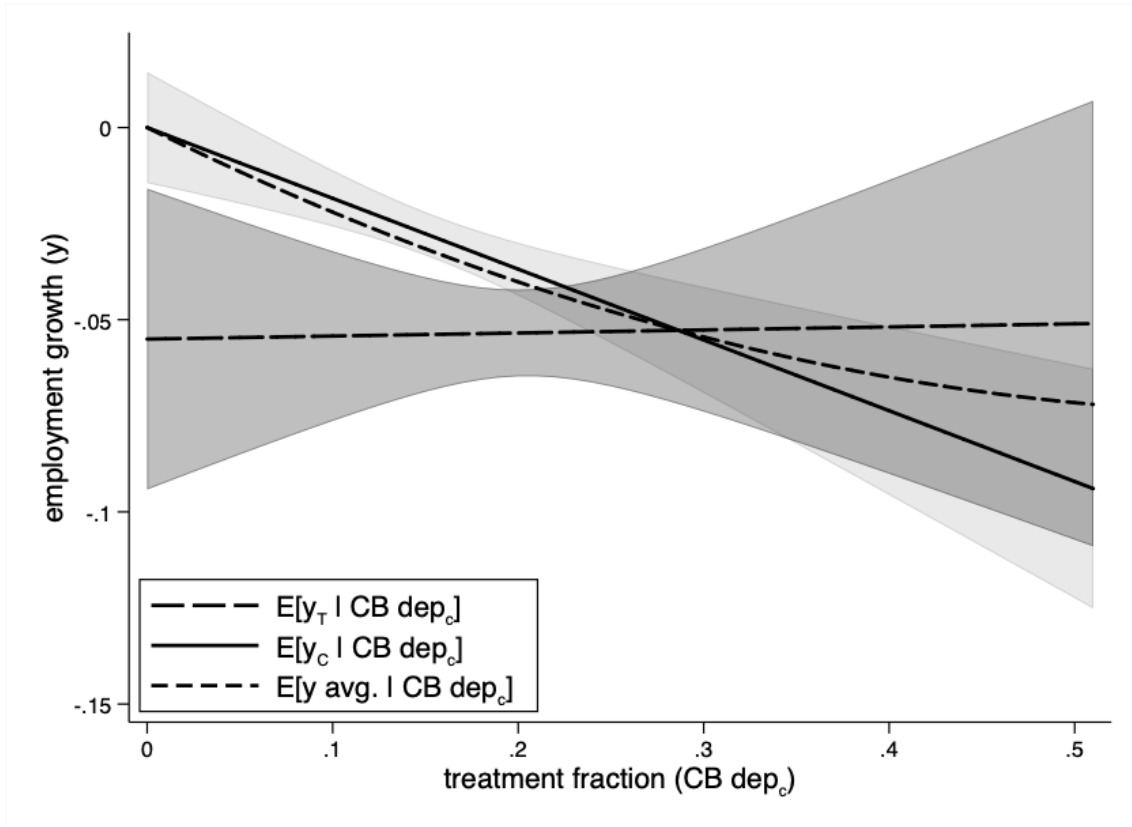
Table 1: Survey of papers in Economics and Finance journals

This table provides a survey of difference-in-differences papers published in the main economics and finance journals in 2017 (*American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, *Review of Economic Studies*, *Journal of Finance*, *Journal of Financial Economics*, *Review of Financial Studies*). In step 1, we automatically search for different versions of the term "difference-in-difference". In step 2, we manually check all papers and exclude those that were marked by the algorithm but did not contain a difference-in-difference analysis. In step 3, we automatically check for terms related to spillovers ("spillover", "spillovers", "indirect effect", "general equilibrium", "aggregation", and "aggregate"). In step 4, we manually check whether spillovers were indeed discussed and/or analyzed in these papers.

Panel A: Overview

Total number of papers	610	
Number of difference-in-differences papers	103	17%
No disussion of spillovers	81	79%
Some discussion of spillovers	22	21%
Discussion only	5	5%
Discussion and analysis	17	16%

Panel B: Informal method used to analyze spillovers

1. As potential concern in direct effects estimation	8	47%
Drop controls where spillovers most likely	6	35%
Other	2	12%
2. As supplementary evidence	9	53%
Within control group estimation of spillovers	7	41%
Regression on individual and aggregate level	2	12%

Panel C: Papers with versus without spillover discussion

Spillover discussed (N=22)			Spillovers not discussed (N=81)		
Research area					
Labor/Education	7	32%	Corporate finance	39	48%
Corporate finance	6	27%	Financial interm.	14	17%
Financial interm.	3	14%	Asset pricing/mgmt	10	12%
Other	6	27%	Labor/Education	4	5%
			Other	14	17%
*Key outcome variable(s) (categories)**					
Compensation, earnings	4	18%	Investment	9	11%
Schooling outcomes	3	14%	Firm value	5	6%
Cash holdings	2	9%	Compensation, earnings	5	6%
Employment	2	9%	Credit supply	5	6%
Consumption	2	9%	Cap. structure, ext. fin.	4	5%
			Asset or sales growth	4	5%
			Innovation	3	4%
			Employment	3	4%
Unit of observation					
Individual	10	45%	Firms	42	52%
Firm	7	32%	Individual	12	15%
Region	3	14%	Region	7	9%
Other	2	10%	Industry	2	2%
			Other (e.g. loan, stock)	18	22%

*Papers may use multiple key outcome variables. All outcome variables that are used in at least 2 (3) papers that do (not) discuss spillovers are listed.

Table 2: Outcome dependency on spillovers

This table provides an overview of the functional relationship of spillovers in classic workhorse models of firm interactions. See the Online Appendix for detailed formulas and the derivation of the relationships.

Competition models		
	Shock on cost	Shock on capacity
Cournot	$y_i = f(d_i, \bar{d})$	$y_i = f(d_i, \bar{d}, d_i\bar{d})$
Salop	$y_i = f(d_i, \bar{d})$	$y_i = f(d_i, \bar{d}, d_i\bar{d})$
Spatial Interaction		
	Shock on cost	Shock on capacity
Demand spillover	$y_i = f(d_i, \bar{d}, d_i\bar{d})$	$y_i = f(d_i, \bar{d}, d_i\bar{d})$
Agglomeration	$y_i = f(d_i, \bar{d})$	$y_i = f(d_i, \bar{d}, d_i\bar{d})$

Table 3: Summary statistics

This table shows summary statistics for the firm employment cross section.

	Mean	SD	p5	p50	p95	N
CB dep_{ic}	0.17	0.29	0.00	0.00	1.00	23,436
CB $dep\ (0/1)_{ic}$	0.18	0.38	0.00	0.00	1.00	23,436
$\overline{\text{CB } dep_{ic}}$	0.17	0.06	0.07	0.17	0.26	23,436
Number of relationship banks$_{ic}$	2.02	1.15	1.00	2.00	4.00	23,436
Employment (fiscal year 2007)$_{ic}$	176.78	2,645.54	2.00	49.00	455.00	23,436
Age (fiscal year 2007)$_{ic}$	22.67	21.31	4.00	17.00	62.00	23,436

Table 4: Baseline results

The unit of observation is the firm level i. The dependent variable is the symmetric growth rate of firm employment from 2008 to 2012. $CB\ dep_{ic}$ is the fraction of the firm's relationship banks that are Commerzbank branches. $\overline{CB\ dep_{ic}}$ is the average Commerzbank dependence of all other firms in the same county (c), excluding firm i itself. The following control variables are included when indicated: indicator variables for 4 firm size bins (1-49, 50-249, 250-999, and over 1,000 employees as of 2007), the log of firm age (as of 2007), and industry fixed effects (2-digit NACE codes). Robust standard errors, clustered at the county level, are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively. Column 4 shows the estimates from Huber (2018) Table 10, column 1 for comparison.

				Huber (2018) estimates
	(1)	(2)	(3)	(4)
CB dep_{ic}	-0.033***	-0.024***	-0.019**	-0.030***
	(0.008)	(0.009)	(0.009)	(0.009)
$\overline{\text{CB } dep_{ic}}$			-0.155***	-0.166**
			(0.044)	(0.076)
Industry fixed effects	No	Yes	Yes	Yes
Size bin fixed effects	No	Yes	Yes	Yes
ln age	No	Yes	Yes	Yes
Observations	23,436	23,436	23,436	48,101

Table 5: Full spillover model

The unit of observation is the firm level i. The dependent variable is the symmetric growth rate of firm employment from 2008 to 2012. $CB\ dep_{ic}$ is the fraction of the firm's relationship banks that are Commerzbank branches. $CB\ dep\ (0/1)_{ic}$ is a dummy variable that equals one if the fraction of the firm's relationship banks that are Commerzbank branches is $>= 0.5$, and zero otherwise. $\overline{CB\ dep}_{ic}$ ($\overline{CB\ dep\ (0/1)_{ic}}$) is the average Commerzbank dependence, calculated based on $CB\ dep_{ic}$ ($CB\ dep\ (0/1)_{ic}$), of all other firms in the same county (c), excluding firm i itself. The following control variables are included when indicated: indicator variables for 4 firm size bins (1-49, 50-249, 250-999, and over 1,000 employees as of 2007), the ln of firm age (as of 2007), and industry fixed effects (2-digit NACE codes). Robust standard errors, clustered at the county level, are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

				CB dependence as indicator		
	(1)	(2)	(3)	(4)	(5)	(6)
CB dep_{ic}	-0.024***	-0.019**	-0.055**			
	(0.009)	(0.009)	(0.027)			
$\overline{\text{CB } dep_{ic}}$		-0.155***				
		(0.044)				
$\overline{\text{CB } dep_{ic}} \times CBdep_{ic}$			0.008			
			(0.112)			
$\overline{\text{CB } dep_{ic}} \times (1 - CBdep_{ic})$			-0.184***			
			(0.052)			
CB $dep\ (0/1)_{ic}$				-0.028***	-0.025***	-0.053***
				(0.006)	(0.007)	(0.017)
$\overline{\text{CB } dep\ (0/1)_{ic}}$					-0.092***	
					(0.033)	
$\overline{\text{CB } dep\ (0/1)_{ic}} \times CBdep\ (0/1)_{ic}$						0.025
						(0.068)
$\overline{\text{CB } dep\ (0/1)_{ic}} \times (1 - CBdep\ (0/1)_{ic})$						-0.115***
						(0.038)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Size bin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
ln age	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,436	23,436	23,436	23,436	23,436	23,436

Table 6: Tradeable vs non-tradeable sectors

The unit of observation is the firm level i. The dependent variable is the symmetric growth rate of firm employment from 2008 to 2012. $CB\ dep_{ic}$ is the fraction of the firm's relationship banks that are Commerzbank branches. $CB\ dep\ (0/1)_{ic}$ is a dummy variable that equals one if the fraction of the firm's relationship banks that are Commerzbank branches is $>= 0.5$, and zero otherwise. $\overline{CB\ dep}_{ic}$ ($\overline{CB\ dep\ (0/1)}_{ic}$) is the average Commerzbank dependence, calculated based on $CB\ dep_{ic}$ ($CB\ dep\ (0/1)_{ic}$), of all other firms in the same county (c), excluding firm i itself. The following control variables are included when indicated: indicator variables for 4 firm size bins (1-49, 50-249, 250-999, and over 1,000 employees as of 2007), the ln of firm age (as of 2007), and industry fixed effects (2-digit NACE codes). The sample is restricted to firms in tradeable or non-tradeable sectors, when indicated. *Base* refers to the full sample, excluding firms in sectors that cannot be classified as tradeable or non-tradeable. Robust standard errors, clustered at the county level, are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Sample:				CB dependence as indicator		
	Base	Non-tradeable	Tradeable	Base	Non-tradeable	Tradeable
	(1)	(2)	(3)	(4)	(5)	(6)
CB dep_{ic}	-0.062**	-0.065	-0.053			
	(0.030)	(0.041)	(0.042)			
$\overline{CB\ dep}_{ic} \times CBdep_{ic}$	0.044	-0.008	0.079			
	(0.131)	(0.169)	(0.193)			
$\overline{CB\ dep}_{ic} \times (1 - CBdep_{ic})$	-0.190***	-0.245***	-0.122*			
	(0.052)	(0.066)	(0.074)			
CB $dep\ (0/1)_{ic}$				-0.055***	-0.057**	-0.050**
				(0.017)	(0.027)	(0.022)
$\overline{CB\ dep\ (0/1)}_{ic} \times CBdep\ (0/1)_{ic}$				0.033	-0.001	0.062
				(0.076)	(0.114)	(0.095)
$\overline{CB\ dep\ (0/1)}_{ic} \times (1 - CBdep\ (0/1)_{ic})$				-0.118***	-0.165***	-0.064
				(0.037)	(0.051)	(0.046)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Size bin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
ln age	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,423	11,346	10,077	21,423	11,346	10,077

Handling Spillover Effects in Empirical Research

Online Appendix (Not for Publication)

The Online Appendix consists of two parts. In Section 1, we study the structure of spillover effects in workhorse models of imperfect competition (i.e., oligopoly models), and in Section 2, we analyze spillover effects in models of spatial interaction.

1 Imperfect Competition

We analyze three different settings of imperfect competition. We first consider classic Cournot competition, that is, firms compete in quantities (Section 1.1). We then study price competition between firms selling differentiated products, using the model of Salop (1979) (Section 1.2). Finally, we analyze a model with a representative consumer, thereby allowing for both quantity and price competition, and for product differentiation in both settings (Section 1.3). For each setting, we consider two different types of shocks: a shock on the marginal costs of treated firms and a shock on the capacity of treated firms.

1.1 Cournot Competition

Consider a simple model of Cournot competition (i.e., quantity competition) with n firms. The firms face a linear inverse demand function of $p = 1 - \sum_{i=1}^{n} y_i = 1 - Y$, where p denotes the price, y_i the quantity produced by firm i, and Y the aggregate quantity produced by all firms. Following the notation of Section 3, we denote by d_i the treatment indicator. A proportion $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$ of firms is treated, while firms in the control group represent a proportion $1 - \bar{d}$ of all firms. We first consider the case in which treated firms face a shock on their marginal costs and then analyze the case in which treated firms face a shock on their capacity.

Shock on marginal costs

Firms' marginal costs are constant but differ between firms in the control and the treatment group. Firms in the control group have marginal costs of $c^C = c$, whereas firms

in the treatment group have marginal costs of $c^T = c + \gamma$, with $\gamma > 0$ (i.e., the treatment is a negative cost shock, for instance, as a result of higher funding costs or a shock on the input price).

The profit function of firm i, denoted by π_i and the resulting first-order conditions are:

$$\pi_i = y_i \left(p - c_i\right) = y_i \left[1 - \left(\sum_{j \neq i}^{n} y_j + y_i\right) - c_i\right], \quad i = 1, ..., n,$$

and:

$$\frac{\partial \pi_i}{\partial y_i} = 1 - Y - c_i - y_i = 0, \quad i = 1, ..., n,$$

respectively. Deducting the first-order condition of a firm in the control group from the first-order condition of a firm in the treatment group provides the relationship between the quantities of the firms in the different groups:

$$y_i^T - y_i^C = - \left(c_i^T - c_i^C\right) = -\gamma \quad \leftrightarrow \quad y_i^T = y_i^C - \gamma. \tag{1}$$

Using (1), we can write $Y = n \left(1 - \bar{d}\right) y_i^C + n \bar{d} y_i^T = n \left(1 - \bar{d}\right) y_i^C + n \bar{d} \left(y_i^C - \gamma\right)$. Plugging this into the first-order condition of a firm in the control group yields:

$$1 - n \left(1 - \bar{d}\right) y_i^C - n \bar{d} \left(y_i^C - \gamma\right) - c - y_i^C \quad \leftrightarrow \quad y_i^C = \frac{1 - c + \gamma n \bar{d}}{n + 1}. \tag{2}$$

Combining (1) and (2) yields a simple linear equation for the equilibrium quantities denoted by y_i^*:

$$y_i^* = \underbrace{\frac{1 - c}{n + 1}}_{f_1} - \underbrace{\gamma d_i}_{f_2(d_i)} + \underbrace{\gamma \frac{n}{n + 1} \bar{d}}_{f_3(\bar{d})}. \tag{3}$$

It is evident from (3) that firms suffering from a negative marginal cost shock produce an output that is below the output of firms in the control group by an amount γ. This is reflected in the second term of (3), which is denoted by $f_2\left(d_i\right)$. This result is intuitive and is based on the effect that the margin of treated firms is below the one of firms in the control group due to the higher marginal costs of treated firms. However, the spillover effect is the same regardless of whether a firm belongs to the treatment or the control group, and is represented by the third term in (3), denoted by $f_3\left(\bar{d}\right)$. The intuition

for the homogeneous spillover effect is as follows: spillovers occur because a change in the quantity of competitors affects each firm through the change in the market price. If the portion of treated firms increases, these firms sell less output, which leads to an increase in the market price. As the same price prevails for all firms, the effect of this price increase is the same for both types of firms. The increase in each firm's margin is therefore homogeneous among firms, which implies that treated and non-treated firms increase their quantities by the same extent.[1]

Shock on the capacity constraint

Consider now a (negative) shock on the capacity instead of a (negative) shock on marginal costs. Specifically, firms in the treatment group are only able to produce an amount of K, whereas firms in the control group are not capacity constrained. As above, a proportion $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ of firms is treated. Apart from the negative shock on the capacity constraint, firms in the treatment and the control group are homogeneous (i.e., marginal costs of all firms are c). For the shock of the capacity constraint to have economic consequences, we assume that:

$$K < \frac{1-c}{1+n}. \tag{4}$$

The equilibrium quantity of each firm in case without a shock is $(1-c)/(1+n)$; hence, if (4) does not hold, the capacity constraint does not bind, which implies that all firms produce a quantity of $(1-c)/(1+n)$. In what follows, we denote by Δ the difference between the equilibrium quantity in a model without a shock on the capacity constraint and the capacity constraint, that is:

$$\Delta \equiv \frac{1-c}{1+n} - K = \frac{1-c-K(n+1)}{n+1}.$$

The parameter Δ serves a similar role as γ in the previous case, as it represents the extent of the shock.

In the equilibrium with shock, each firm in the treatment group produces at the capacity constraint and sells a quantity of K. The profit function of firm i in the control

---

[1]The profit increase is nevertheless larger for firms in the control group as they sell a larger output than treated firms.

group can therefore be written as:

$$\pi_i = y_i \left[1 - n\bar{d}K - \left(\sum_{j \neq i}^{n} (1 - d_j) y_j + y_i \right) - c \right].$$

Taking the first-order condition and solving for the symmetric equilibrium (i.e., all firms in the control group sell the same quantity) yields:

$$y^C = \frac{1 - c - n\bar{d}K}{1 + n(1 - \bar{\bar{d}})}.$$

Therefore, the equilibrium quantity y_i^* of each firm i is:

$$y_i^* = \underbrace{\frac{1 - c}{1 + n}}_{f_1} - \underbrace{\Delta d_i}_{f_2(d_i)} + \underbrace{\Delta \frac{n\bar{d}}{1 + n(1 - \bar{d})} (1 - d_i)}_{f_3(\bar{d}, d_i \bar{d})}. \tag{5}$$

In contrast to the case of a shock on the marginal costs, the spillover effect is now only relevant for the control group but not for the treatment group. This can be seen in (5) because the last term, denoted by $f_3\left(\bar{d}, d_i \bar{d}\right)$, depends on $1 - d_i$. This result is intuitive: as a firm in the treatment group produces at the capacity constraint, is is not affected by spillover effects. In addition, it is easy to check that $f_3\left(\bar{d}, d_i \bar{d}\right)$ is increasing in $\bar{d}$, that is, the more firms face a capacity constraint, the higher the quantity produced by firms in the control group. The intuition is again that the reduction in output of the treated firms leads to a higher market price, and thereby to a higher margin. If more firms are treated, this margin increase is higher, which implies that non-treated firms sell more.

1.2 Circular Competition (Salop, 1979)

Consider next a model of circular competition between firms, which was developed by Salop (1979). There are n firms producing differentiated products. The firms are evenly distributed on a circle with circumference 1. Consumers are uniformly distributed on this circle and incur a transport cost of t per unit of distance. That is, if a consumer located at point x on the circle purchases from firm i located on point x_i, her net utility is $v - p_i - t|x - x_i|$, where v is the benefit from the product and p_i is firm i's price.

Consumers wish to buy one unit of the good and buy from the firm that offers them the highest net utility. To simplify the exposition, we assume that v is sufficiently large, so that all consumers buy one unit of the product. Firms simultaneously set prices.

Shock on marginal costs

We first consider a marginal-cost shock as in the previous section. Firms have constant marginal costs, and firms in the treatment group face a negative cost shock—i.e., firms in the control group have marginal costs of $c^C = c$ and firms in the treatment group have marginal costs of $c^T = c + \gamma$. As above, a proportion $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$ of firms is treated, while firms in the control group represent a proportion $1 - \bar{d}$ of all firms. Following e.g. Raith (2003) and Aghion and Schankerman (2004), we assume that firms do not know the cost characteristics of neighboring firms, and thus base their pricing decisions on the 'average' costs of their neighbors.

We first derive the demand for each firm i. Denoting the prices by firm i's neighboring firms by p_{i-1} and p_{i+1}, the marginal consumer between i and $i-1$ (i.e., the consumer indifferent between buying from firm i and from firm $i-1$) is given by:

$$x_{i,i-1} = \frac{1}{2n} + \frac{p_{i-1} - p_i}{2t}$$

and the marginal consumer between i and $i+1$ is given by:

$$x_{i,i-1} = \frac{1}{2n} + \frac{p_{i+1} - p_i}{2t}.$$

Therefore, firm i's demand is:

$$D_i(p_i, p_{i-1}, p_{i+1}) = \frac{1}{n} + \frac{p_{i-1} + p_{i+1} - 2p_i}{2t}.$$

When setting its price p_i, firm i does not observe prices p_{i-1} and p_{i+1} charged by its neighbors. However, it anticipates that all treated (non-treated) firms will charge the same price p^T (p^C) in equilibrium. Given that a proportion $\bar{d}$ of firms is treated, a firm

in the control group faces an expected price of its neighbors given by:

$$\frac{n\bar{d}p^T + \left(n(1-\bar{d})-1\right)p^C}{n-1},$$

whereas a firm in the treatment group faces an expected price of its neighbors given by:

$$\frac{\left(n\bar{d}-1\right)p^T + n(1-\bar{d})p^C}{n-1}.$$

The maximization problem of firm j belonging to the control group is therefore:

$$\max_{p_j} \quad \pi_j = (p_j - c)\left(\frac{1}{n} + \frac{\frac{n\bar{d}p^T + \left(n(1-\bar{d})-1\right)p^C}{n-1} - p_j}{t}\right)$$

and the maximization problem of firm k belonging to the treatment group is:

$$\max_{p_k} \quad \pi_k = (p_k - c - \gamma)\left(\frac{1}{n} + \frac{\frac{\left(n\bar{d}-1\right)p^T + n(1-\bar{d})p^C}{n-1} - p_k}{t}\right)$$

Taking the first-order conditions and solving for the symmetric Nash equilibrium ($p_j = p^C$ and $p_k = p^T$), we obtain that the equilibrium prices of the two types of firms are:

$$p^C = c + \frac{t}{n} + \frac{\gamma n\bar{d}}{2n-1} \quad \text{and} \quad p^T = c + \frac{t}{n} + \frac{\gamma(n-1)}{2n-1} + \frac{\gamma n\bar{d}}{2n-1}.$$

This implies that $p^T = p^C + \gamma(n-1)/(2n-1)$, that is, the price difference between the two types of firms is smaller than the cost difference. In other words, treated firms do not shift the cost shock γ fully into the consumer price. The intuition is that, due to the fact that firms in the control group face lower costs and therefore set a lower price, shifting the cost increase fully onto consumers would reduce the demand of a treated firm by a very large amount. Solving for the equilibrium quantities y_i^*, we obtain:

$$y_i^* = \underbrace{\frac{1}{n}}_{f_1} - \underbrace{\gamma\frac{n}{t(2n-1)}d_i}_{f_2(d_i)} + \underbrace{\gamma\frac{n}{t(2n-1)}\bar{d}}_{f_3(\bar{d})}. \tag{6}$$

VI

It follows from (6) that firms in the treatment group produce an output that is $\gamma n/(t(2n-1))$ below the one of firms in the control group. This is reflected in the term denoted by $f_2(d_i)$. However, as in the Cournot model above, the spillover effect is the same regardless of whether a firm belongs to the treatment or the control group, and is represented by the spillover effect $f_3(\bar{d})$. The homogeneous spillover effect occurs because each firm benefits from the expected price increase resulting from a larger portion of treated firms due to the higher expected price charged by its neighboring firms. As both treated and non-treated firms face the same expectation with respect to the type of its neighboring firms, the spillover effect is homogeneous. We also note that the relation between shock and quantity is a linear one, as in the Cournot model.

Shock on the capacity constraint

We consider next a shock on the capacity constraint, that is, firms in the treatment group are only able to produce a quantity of K. As above, a proportion $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ of firms is treated. To ensure that K is a real constraint, we assume:

$$K < \frac{1}{n},$$

which implies that the capacity constraint is binding in equilibrium (i.e., the constraint is below the quantity that a firm would have produced in the equilibrium without shock). As above, we denote by $\Delta \equiv 1/n - K$ the difference between the equilibrium in a model without shock and the capacity constraint.

Anticipating that all treated (non-treated) firms will charge the same price p^T (p^C), the maximization problem of firm j in the control group is:

$$\max_{p_j} \quad \pi_j = (p_j - c)\left(\frac{1}{n} + \frac{\frac{n\bar{d}p^T + \left(n(1-\bar{d})-1\right)p^C}{n-1} - p_j}{t}\right),$$

yielding a first-order condition of:

$$\frac{1}{n} + \frac{\frac{n\bar{d}p^T + \left(n(1-\bar{d})-1\right)p^C}{n-1} - p_j}{t} - \frac{p_j - c}{t} = 0.$$

In equilibrium, $p_j = p^C$. Inserting this into the the first-order condition and solving for p^C, we obtain:

$$p^C = \frac{(t + cn)(n - 1) + p^T n^2 \bar{d}}{n\left(n(1 + \bar{d}) - 1\right)}. \tag{7}$$

Instead, a firm belonging to the treatment group will set its price in such a way that it will sell its capacity K, regardless of the type of its neighboring firms. Specifically, if its two neighboring firms are firms which are not treated—and therefore face no capacity constraint—the firm will set its equilibrium price so that it sells to $K/2$ consumers on each side. This implies that the relationship between p^C and p^T is such that:

$$K = \frac{1}{n} + \frac{p^C - p^T}{t}. \tag{8}$$

Setting its price according to (8) also ensures that the firm sells its entire capacity if its neighbors do not only consist of non-treated firms: as firms in the treatment group sell a lower quantity, the firm can then also serve $K/2$ consumers on each side. However, some consumers do not get served in equilibrium if two firms of the treatment group are neighbors.

Solving (7) and (8) for the equilibrium prices yields:

$$p^C = c + \frac{t}{n} + \frac{nt\bar{d}(1 - Kn)}{n(n - 1)} \quad \text{and} \quad p^T = c + \frac{2t}{n} - Kt + \frac{nt\bar{d}(1 - Kn)}{n(n - 1)}.$$

Inserting these prices into the expressions for the quantities yields that the equilibrium quantity for each firm i is:

$$y_i^* = \underbrace{\frac{1}{n}}_{f_1} - \underbrace{\Delta d_i}_{f_2(d_i)} + \underbrace{\Delta \frac{n}{n - 1}\bar{d}(1 - d_i)}_{f_3(\bar{d}, d_i \bar{d})}. \tag{9}$$

As a consequence, we obtain a linear relationship between the shock and the quantity of each firm. As firms in the treatment group produce at the capacity constraint, the spillover effect is only relevant for firms in the control group, which can be seen in the last term of (9), as this term involves $(1 - d_i)$.

VIII

1.3 Representative-Consumer Model

A third widely-used demand model is one with a representative consumer. Its linear version was first developed by Bowley (1924) and popularized in a highly-cited paper by Singh and Vives (1984). We first describe how to derive the linear demand function from the representative consumer's utility function and then solve for the equilibrium, both with quantity and price competition.

The utility function of the representative consumer is:

$$U(y_1, ..., y_n) = \alpha \sum_{i=1}^{n} y_i - \frac{1}{2} \left(\beta \sum_{i=1}^{n} y_i^2 - \delta \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} y_i y_j \right),$$

The consumer maximizes her utility subject to the budget constraint, which is given by $\sum_{i=1}^{n} p_i y_i \leq M$, where M is the consumer's income. Because the consumer optimally spends her entire income on the goods, the maximization problem is $U(y_1, ..., y_n) - \sum_{i=1}^{n} p_i y_i$. This leads to an inverse demand function of product i given by:

$$p_i = \alpha - \beta y_i - \delta \sum_{j=1, j \neq i}^{n} y_j, \quad i = 1, ..., n. \tag{10}$$

Therefore, β measures the effect of firm i's quantity on its own price—i.e., it determines firm i's price elasticity of its own demand—whereas δ measures the effect of another firm's quantity on firm i's price—i.e., it determines the cross elasticity of demand. The inverse of δ represents the degree of differentiation between firms' products. Specifically, if $\delta = 0$, products are independent and each firm is a monopolist. Instead, if $\delta \to \beta$, products become perfect substitutes.

Shock on the marginal costs

As in the examples above, firms have constant marginal costs. Firms in the treatment group face a negative cost shock and have marginal costs of $c^T = c + \gamma$ whereas firms in the control group have marginal costs of $c^C = c$. A proportion $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$ of firms is treated, while firms in the control group represent a proportion $1 - \bar{d}$ of all firms.

(i) Quantity competition

If firms compete in quantities, each firm i's profit function is:

$$\pi_i = y_i \left[\alpha - \left(\beta y_i + \delta \sum_{j \neq i}^{n} y_j \right) - c_i \right].$$

It is evident that this is a generalized version of the first setting (i.e., Cournot competition). In the latter case $\alpha = \beta = \delta = 1$, which implies that firms' products are homogeneous.

Following the same procedure as above, we can solve for the equilibrium quantities to get:

$$y_i^* = \underbrace{\frac{\alpha - c}{2\beta + \delta(n-1)}}_{f_1} - \gamma \underbrace{\frac{1}{2\beta - \delta} d_i}_{f_2(d_i)} + \gamma \underbrace{\frac{n\delta}{(2\beta - \delta)(2\beta + \delta(n-1))} \bar{d}}_{f_3(\bar{d})}.$$

It follows that the output of the treatment group is lower by an amount of $\gamma/(2\beta - \delta)$ compared to the control group but the spillover effect is again homogeneous among firms in the treatment and firms in the control group. The intuition is the same as the one given in Section 1.1.

(ii) Price competition

To analyze price competition, we first need to invert the demand system given by (10) to obtain the quantity as a function of the prices. This can be done following the technique introduced by Hackner (2000). Specifically, summing (10) over all n firms yields:

$$n\alpha - \beta \sum_{j=1}^{n} y_j - \delta(n-1) \sum_{j=1}^{n} y_j - \sum_{j=1}^{n} p_j = 0.$$

Using that $\sum_{j=1}^{n} y_j = y_i + \sum_{j=1, j \neq i}^{n} y_j$ and solving (10) for $\sum_{j=1, j \neq i}^{n} y_j$, allows us to derive firm i's quantity as a function of prices:

$$y_i(p_i, p_{-i}) = \frac{\alpha}{\beta + \delta(n-1)} - \frac{p_i}{\beta - \delta} + \frac{\delta \sum_{j=1, j \neq i}^{n} p_j}{(\beta - \delta)(\beta + \delta(n-1))}, \tag{11}$$

where, following standard notation, p_{-i} denotes the set of prices of all firms but firm i.

X

The respective profit function of firm i is:

$$(p_i - c_i) \, y_i(p_i, p_{-i}). \tag{12}$$

We can solve for the equilibrium prices in the same way as above. Plugging the equilibrium prices into the quantities yields that the equilibrium quantities are given by:

$$y_i^* = \underbrace{\frac{\alpha(\beta - \delta) - \beta c}{(\beta - \delta)(\beta + \delta(n-1))}}_{f_1} - \gamma \underbrace{\frac{\beta + n\delta}{(\beta - \delta)(2\beta + \delta(2n-1))} d_i}_{f_2(d_i)}$$

$$+ \gamma \underbrace{\frac{n\delta(\beta + \delta(n-1))}{(\beta - \delta)(2\beta + \delta(2n-1))(\beta + \delta(2n-1))} \bar{d}}_{f_3(\bar{d})}.$$

It is evident from the term $f_2(d_i)$ that a treated firm sells a lower quantity than a firm in the control group. However, the spillover effect is the same on both types of firms, as can be seen from the term $f_3(\bar{d})$, which does not depend in d_i. The intuition is that a price increase by competitors (positively) affects the residual demand of each firm in the same way, regardless of whether the firm is treated or not (as can be seen from (11)). The equation determining y_i^* is also linear, both for quantity and price competition.

Shock on the capacity constraint

As above, we consider next a shock on the capacity constraint. Firms in the treatment group, of which there is a proportion $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$, are then only able to produce an amount of K, whereas firms in the control group are not capacity constrained.

(i) Quantity competition

We again start with quantity competition. To ensure that K is a real constraint, we assume:

$$K < \frac{\alpha - c}{2\beta + \delta(n-1)},$$

and denote by Δ the difference between the equilibrium quantity in a model without a

shock on the capacity constraint and the capacity constraint, that is:

$$\Delta \equiv \frac{\alpha - c}{2\beta + \delta(n-1)} - K = \frac{\alpha - c + K\left(2\beta + \delta(n-1)\right)}{2\beta + \delta(n-1)}.$$

Solving for the equilibrium quantities in the same way as in the Cournot model yields:

$$y_i^* = \underbrace{\frac{\alpha - c}{2\beta + \delta(n-1)}}_{f_1} - \underbrace{\Delta d_i}_{f_2(d_i)} + \underbrace{\Delta \frac{n\bar{d}}{2\beta + \delta\left(n(1-\bar{d})-1\right)}\left(1 - d_i\right)}_{f_3\left(\bar{d}, d_i \bar{d}\right)}.$$

It is evident that the same result as in the Cournot model is obtained, that is, the spillover effect only affects firms in the control group but not those in the treatment group.

(ii) Price competition

Turning to price competition, the assumption that guarantees that K is below the equilibrium quantity that a firm in the treatment group produces without shock is:

$$K < \frac{\alpha}{2\beta - \delta(n-1)} - \frac{c\beta}{(\beta - \delta)\left(2\beta - \delta(n-1)\right)}.$$

The maximization problem of a firm in the treatment group is, as above, given by (12), with $y_i(p_i, p_{-i})$ given by (11), whereas a firm in the control group sets its price such that is sells its capacity K. Solving for the equilibrium then yields that the equilibrium quantity of each firm i is:

$$y_i^* = \underbrace{\frac{\alpha}{2\beta - \delta(n-1)} - \frac{c\beta}{(\beta - \delta)\left(2\beta - \delta(n-1)\right)}}_{f_1} - \underbrace{\Delta d_i}_{f_2(d_i)}$$

$$+ \underbrace{\Delta \frac{\delta n\left(\beta + \delta(n-1)\right)\bar{d}}{2\beta^2 - \delta^2 n(n-1)(1-\bar{d}) + \beta\delta\left(n(3-\bar{d})-1\right)}\left(1 - d_i\right)}_{f_3\left(\bar{d}, d_i \bar{d}\right)}.$$

Therefore, the same result as above is obtained: the spillover effect only affects the control group but not the treatment group, as the last term, which involves $\bar{d}$, is multiplied by $1 - d_i$.

2 Spatial Interaction

We next turn to firm interdependencies via spatial interaction. As described in Section 3 of the main text, we consider two settings. First, we study a model of demand spillovers (Section 2.1); second, we analyze agglomeration effects between firms (Section 2.2).

2.1 Demand Spillovers

In this section, we consider a simple model of demand spillovers and determine how such spillovers affect treated and control firms. In contrast to the oligopoly models, there is no classic model of demand spillovers. However, the most influential papers considering demand spillovers are Shleifer and Vishny (1988) and Murphy, Shleifer, and Vishny (1989). Our model is an adaptation of these models, allowing us to analyze the different effects on output and employment of treated and control firms. Specifically, these papers analyze demand spillovers in a general equilibrium model in which demand in each market is a function of labor income and profits. To achieve this, they make some simplifying assumptions, such as inelastic demand and the existence of a competitive fringe, which puts a cap on prices. In contrast to these papers, we focus on the case in which local demand only depends on labor income but we allow for elastic demand and put no constraint on prices.

The description of the model is as follows: There are n goods, where n is considered to be a very large number (in the limit, a continuum). Each good is produced by one firm, which is a local monopolist for the good it produces. The local demand in each market i is denoted by q_i and is given by $q_i = \alpha(L)(1 - p_i)$, where p_i is the price charged by firm i and L is the sum of the labor income over all n markets. The function $\alpha(L)$ is strictly increasing and is assumed to take the form $\alpha(L) = \sqrt{L}$. This represents the demand spillover effect. The economic reasoning behind this function is that the employees of each firm consume the local goods and spend a portion of their labor income to buy these goods. The concave shape of $\alpha(L)$ is due to decreasing marginal utility of each good.

For simplicity, we assume that the relationship between employment and the aggregate output sold by the n firms (i.e., $\sum_{i=1}^{n} q_i$) is one-to-one, and that the wage is normalized

to 1. This implies that $L = \sum_{i=1}^{n} q_i$. Inserting this into $\alpha(L)$ yields

$$\alpha(L) = \sqrt{L} = \sqrt{\sum_{i=1}^{n} q_i} = \sqrt{\sum_{i=1}^{n} \alpha(1 - p_i)}.$$

Shock on marginal costs

As in the previous section, each firm has constant marginal costs. They are equal to c for non-treated firms and $c+\gamma$ for treated firms (i.e., a negative cost shock). A proportion $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ of firms is treated.

First, consider the maximization problem of a firm in the control group. As n is very large, the firm cannot influence α and takes it as given. The profit function of a firm j in this group can therefore be written as:

$$\pi_j = \alpha \left(p_j - c \right) \left(1 - p_j \right).$$

Maximizing with respect to p_j leads to an optimal price of $(1+c)/2$, which implies that the optimal quantity is $\alpha(1 - c)/2$. Similarly, the profit function of a firm k in the treatment group is:

$$\pi_k = \alpha \left(p_j - c - \gamma \right) \left(1 - p_k \right),$$

leading to an optimal price of $(1 + c + \gamma)/2$ and an optimal quantity of $\alpha(1 - c - \gamma)/2$.

Because the wage is normalized to 1 and the aggregate output equals the aggregate employment, we can write:

$$L = n(1 - \bar{d})\frac{\alpha(1 - c)}{2} + n\bar{d}\frac{\alpha(1 - c - \gamma)}{2}.$$

Inserting this into $\alpha = \sqrt{L}$ and solving for α yields:

$$\alpha = \frac{n(1 - c - \gamma\bar{d})}{2}.$$

As a consequence, the equilibrium quantity of firm i can be written as:

$$y_i^* = \frac{n}{4}\left(\underbrace{(1-c)^2}_{f_1} - \underbrace{\gamma(1-c)d_i}_{f_2(d_i)} - \underbrace{\gamma(1-c)\bar{d}}_{f_3(\bar{d})} + \underbrace{\gamma^2\bar{d}d_i}_{f_4(d_i\bar{d})}\right) \tag{13}$$

As the assumption of the model is that the relation between employment and quantity in a sector is one-to-one, quantities also reflect employment.

The quantity (and therefore the employment) of a treated firm is smaller than the one of a non-treated firm, which can be seen from the term $f_2(d_i)$ of the right-hand side in (13). This result is the same as in the models with imperfect competition. It is again driven by the fact that a higher marginal cost induce a treated firm to reduce its output. In contrast to the case of imperfect competition, however, the spillover effect is negative in case of demand spillovers. This can be seen from the term $f_3(\bar{d})$. The reason is that the shock not only lowers quantities but, through the reduced employment, also the disposable income of consumers, which implies that demand falls. As a consequence, due to the spillover, employment of each firm is lower than in the absence of a shock. In addition, the spillover effect now affects treated and non-treated firms heterogeneously. This can be seen from the term $f_4(d_i\bar{d})$, which enters the expression with a positive sign. Therefore, the spillover effect is less dramatic for treated firms as compared to firms in the control group.[2] The intuition is that treated firms produce a lower quantity and are therefore less affected by the demand reduction. It follows that a lower employment level, and thus a reduction in disposable income, affects larger firms more negatively than smaller firms.

We also note that the relationship between output (or employment) and the marginal cost shock/spillover effect is again linear in this simple model.

Shock on the capacity constraint

We next consider a shock on the capacity of the treated firms. Suppose that the proportion $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ of firms can only sell an amount K (whereas firms in the control group are not capacity constrained). Solving the model without such a shock yields that

---

[2] The overall spillover effect on treated firms is, however, always negative, as $\gamma < 1 - c$, due to the fact that all firms produce a strictly positive quantity.

each firm sells a quantity of $n(1-c)^2/4$ in equilibrium. Hence, to allow the capacity constraint to have bite, we assume:

$$K < \frac{n(1-c)^2}{4}.$$

We again denote by Δ the difference between $n(1-c)^2/4$ and K.

The maximization problem of a firm j in the control group is the same as in case of a shock on marginal costs—i.e., it is given by $\alpha(p_j - c)(1 - p_j)$—which implies that the optimal quantity is again $\alpha(1 - c - \gamma)/2$. Instead, a firm in the treatment group sells its entire capacity K. As a consequence, the aggregate output is:

$$n(1-\bar{d})\frac{\alpha(1-c)}{2} + n\bar{d}K. \tag{14}$$

Due to the fact that the wage is normalized to 1 and the relationship between output and employment is one-to-one, labor income is given by (14). Using this in $\alpha = \sqrt{L}$ and solving for α yields:

$$\alpha = \frac{n(1-c)(1-\bar{d}) + \sqrt{n\left(n(1-c)^2(1-\bar{d})^2 + 16K\bar{d}\right)}}{4}.$$

As a consequence, the equilibrium quantity for each firm i can be written in a concise form as:

$$y_i^* = \underbrace{\frac{n(1-c)^2}{4}}_{f_1} - \underbrace{\Delta d_i}_{f_2(d_i)} - \underbrace{\frac{(1-c)\left(n(1+\bar{d})(1-c) - \sqrt{n\left(n(1-\bar{d})^2(1-c)^2 - 16\bar{d}\Delta\right)}\right)}{8}(1-d_i)}_{f_3(\bar{d},d_i\bar{d})}, \tag{15}$$

and the equilibrium employment in sector i is also given by this expression. It follows from the last term of (15) that there is no spillover effect for firms in the treatment group, as these firms produce at the capacity constraint. Instead, for firms in the control group, the term $f_3\left(\bar{d}, d_i\bar{d}\right)$ shows that their quantity is negatively affected by Δ, which implies that the spillover effect is negative for them. As $\bar{d}$ is present in the square root of the term, the effect is non-linear.

An important case is the one in which the shock is so severe that firms in the treatment group go bankrupt, and therefore need to exit the market. This is equivalent to a capacity constraint of $K = 0$. In this case, $\Delta = n(1-c)^2/4$; using this in (15), we obtain that a firm in the control group sells an output of:

$$y_C^* = \underbrace{\frac{n(1-c)^2}{4}}_{f_1} - \underbrace{\Delta \bar{d}(1-d_i)}_{f_3(\bar{d}, d_i \bar{d})} .$$

As a consequence, the spillover effect is linear in this case.

2.2 Agglomeration

Finally, we analyze a simple model in which spillovers occur due to agglomeration effects. Dating back to Marshall (1890), agglomeration economies are broadly considered as factors that allow clustered firms, or firms that are present in the same location, to obtain higher profits than isolated firms. Among the various reasons for this, Marshall (1890), Hoover (1948), and Krugman (1991), among many others, emphasize that agglomerated firms, first, benefit from information spillovers, allowing them to operate with a cheaper cost function than segmented firms, and, second, offer a pooled market for workers with industry-specific skills, leading to a higher probability to get skilled labor, which again allows for cheaper production. Taking these reasons into account, we formulate a simple model in which firms benefit from spillovers of other firms via a reduction in their marginal costs—i.e., marginal costs are the lower, the more firms are present and the larger is each firm's investment in cost reduction. The model follows d'Aspremont and Jacquemin (1988) in the way spillovers between firms are modeled.[3]

The description of the model is as follows: there are n firms, each one operating in a separate market, that is, there are no competition externalities between firms. The inverse demand function in each market i is $p_i(y_i) = \alpha - y_i$. The marginal cost of firm i depends on its own investment and of the investment of all other firms. To capture this

---

[3]To simplify the exposition, we abstract from competition between firms, which is also considered by d'Aspremont and Jacquemin (1988). However, we extend their model by allowing for n firms instead of only 2 and, naturally, consider firm heterogeneity.

in a simple way, marginal costs of firm i are:

$$c_i - x_i - \beta \sum_{j=1, j \neq i}^{n} x_j, \tag{16}$$

where x_j is the investment level of firm $j \neq i$. Therefore, own investment reduces marginal costs at a one-to-one relation, whereas agglomeration effects due to spillovers are measured by the parameter $\beta \in [0, 1]$. If $\beta = 0$, agglomeration effects are absent; instead, if $\beta = 1$, there are full spillovers, which implies that each firm benefits to same extent from investment of other firms in the cluster as from own investment. Investment cost is quadratic and given by $\kappa x_i^2 / 2$, which reflects diminishing returns from investment. To ensure an interior solution, we assume $\kappa > (1 + \beta(n-1))/2$, that is, investment costs are sufficiently convex.

The profit function of each firm i is therefore given by:

$$\pi_i(y_i, y_{-i}, x_i, x_{-i}) = (\alpha - y_i) y_i - \left(c_i - x_i - \beta \sum_{j=1, j \neq i}^{n} x_j \right) y_i - \kappa \frac{x_i^2}{2}.$$

The maximization variables are y_i and x_i.

Shock on the marginal costs

As in the previous examples, we first consider a shock on the firms' marginal costs. Following the above examples, a firm in the control group is characterized by $c_i = c$, whereas a firm in the treatment group is characterized by marginal costs $c_i = c + \gamma$, with $\gamma > 0$ (i.e., the treatment is again a negative cost shock). A proportion $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$ of firms is treated, while the remaining proportion $1 - \bar{d}$ are firms is in the control group.

The first-order conditions for profit maximization are:[4]

$$\alpha - c_i + x_i + \beta \sum_{j=1, j \neq i}^{n} x_j - 2y_i = 0 \quad \text{and} \quad y_i - \kappa x_i = 0, \quad \forall i = 1, ..., n.$$

In a symmetric equilibrium, all firms of the same type set the same variables in equilibrium, that is, all firms in the control group set $y = y^C$ and $x = x^C$ in equilibrium, whereas

---

[4]Due to the assumption $\kappa > (1 + \beta(n-1))/2$, the Hessian is strictly positive definite, which implies that all maximization problems are strictly concave.

all firms in the treatment group set $y = y^T$ and $x = x^T$. The respective first-order conditions for firms in the control and firms in the treatment group can then be written as follows:

$$\alpha - c + x^C + \beta\left((n-1-k)x^C + kx^T\right) - 2y^C = 0, \quad y^C - \kappa x^C = 0,$$

and:

$$\alpha - c - \gamma + x^T + \beta\left((n-k)x^C + (k-1)x^T\right) - 2y^T = 0, \quad y^T - \kappa x^T = 0.$$

Solving these four equations for the equilibrium investment levels and quantities of treated firms and control firms yields:

$$y_i^* = \frac{1}{2\kappa - 1 - \beta(n-1)}\left(\underbrace{\kappa(\alpha - c)}_{f_1} - \underbrace{\kappa\gamma d_i}_{f_2(d_i)} - \underbrace{\frac{n\kappa\gamma\beta}{2\gamma - 1 + \beta}\bar{d}}_{f_3(\bar{d})}\right) \quad \text{and} \quad x_i^* = \frac{y_i^*}{\kappa}. \tag{17}$$

As can be seen from (17), both investment level and quantity are affected in a similar way from the shock. The term $f_2(d_i)$ shows that firms in the treatment group invest less and sell a smaller quantity than firms in the control group. The term $f_3(\bar{d})$ shows that firms are negatively affected from spillover effects and that this effect is homogeneous for treated and control firms. The intuition for these results is as follows: first, as treated firms have higher marginal costs, they produce less, which in turn renders investment in cost reduction less profitable for them. Therefore, a larger portion of treated firms leads to less aggregate investment, which implies that the spillover effect is negative. Second, investments by other firms reduce marginal costs of treated and non-treated firms in the same (as can be seen from (16)). As a consequence, the spillover lowers the output and investment of both types of firms to the same extent.[5]

Shock on the capacity constraint

Finally, we consider a shock on the capacity of firms in the treatment group. Suppose

---

[5]We note, however, that the spillover affects the profit of firms in the control group more than firms in the treatment group, as the former sell a larger quantity.

that the proportion $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ of firms can only sell an amount K. In a model without such a shock, each firm sells a quantity of $\kappa\left(\alpha - c\right)/\left(2\kappa - 1 - \beta(n-1)\right)$ in equilibrium. Hence, to allow the capacity constraint to have bite, we assume:

$$K < \frac{\kappa\left(\alpha - c\right)}{2\kappa - 1 - \beta(n-1)},$$

and denote by Δ the difference between the equilibrium quantity without shock and the capacity constraint.

Solving the model in the same way as with a shock on the marginal costs but setting $\gamma = 0$ and instead inserting $y^T = K$ yields that the equilibrium quantities and investment levels are:

$$y_i^* = \underbrace{\frac{\kappa\left(\alpha - c\right)}{2\kappa - 1 - \beta(n-1)}}_{f_1} - \underbrace{\Delta d_i}_{f_2(d_i)} - \underbrace{\Delta\frac{\beta n\bar{d}}{2\kappa - 1 - \beta\left(n(1-\bar{d}) - 1\right)}\left(1 - d_i\right)}_{f_3\left(\bar{d},d_i\bar{d}\right)} \quad \text{and} \quad x_i^* = \frac{y_i^*}{\kappa}.$$

$$(18)$$

From (18), it is easy to see that the spillover effect resulting from a shock on the capacity is again only relevant for the control group but not for the treatment group. For a firm in the control group, the effect is negative (and non-linear). In addition, as the term denoted by $f_3\left(\bar{d}, d_i\bar{d}\right)$ is increasing in $\bar{d}$, the more firms face a capacity constraint, the lower the quantity and the investment level of a firm in the control group.

References

AGHION, P., AND M. SCHANKERMAN (2004): "On the Welfare Effects and Political Economy of Competition-Enhancing Policies," *Economic Journal*, 114(498), 800–824.

BOWLEY, A. L. (1924): *The Mathematical Groundwork of Economics.* Oxford University Press, Oxford.

D'ASPREMONT, C., AND A. JACQUEMIN (1988): "Cooperative and Noncooperative R&D in Duopoly with Spillover," *American Economic Review*, 78(5), 1133–1137.

HACKNER, J. (2000): "A Note on Price and Quantity Competition in Differentiated Oligopolies," *Journal of Economic Theory*, 93(2), 233–239.

HOOVER, E. M. (1948): *The Location of Economic Activity.* New York: McGraw-Hill.

KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 481–499.

MARSHALL, A. (1890): *Principles of Economics.* London: McMillan.

MURPHY, K. M., A. SHLEIFER, AND R. W. VISHNY (1989): "Industrialization and the Big Push," *Journal of Political Economy*, 97(5), 1003–1026.

RAITH, M. (2003): "Competition, Risk, and Managerial Incentives," *American Economic Review*, 93(4), 1425–1436.

SALOP, S. C. (1979): "Monopolistic Competition with Outside Goods," *Bell Journal of Economics and Management Science*, 10, 141–156.

SHLEIFER, A., AND R. W. VISHNY (1988): "The Efficiency of Investment in the Presence of Aggregate Demand Spillovers," *Journal of Political Economy*, 96(6), 1221–1231.

SINGH, N., AND X. VIVES (1984): "Price and Quantity Competition in a Differentiated Duopoly," *RAND Journal of Economics*, 15(4), 546–554.