NBER WORKING PAPER SERIES

ESTIMATING GENERAL EQUILIBRIUM SPILLOVERS OF LARGE-SCALE SHOCKS

Kilian Huber

Working Paper 29908
http://www.nber.org/papers/w29908

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2022, Revised September 2022

I am grateful to Holger Mueller, Tobias Berg, Steven J. Davis, Sarah Griebel, Peter Hull, Erik Hurst, Alan Manning, Steve Pischke, Andrés Sarto, Ludwig Straub, Daniel Streitz, Joseph Vavra, Emil Verner, Iván Werning, and Christian Wolf and numerous seminar audiences for helpful conversations and comments. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Estimating General Equilibrium Spillovers of Large-Scale Shocks
Kilian Huber
NBER Working Paper No. 29908
April 2022, Revised September 2022
JEL No. C13,C2,D5,E0,E51,G0,G21,G30,L2,R11,R23,R51

ABSTRACT

Large-scale shocks directly affect some firms and households and indirectly affect others through general equilibrium spillovers. In this paper, I describe how researchers can directly estimate spillovers using quasi-experimental or experimental variation. I then argue that spillover estimates suffer from distinct sources of mechanical bias that standard empirical tools cannot resolve. These biases are particularly relevant in finance and macroeconomics, where multiple spillover channels and nonlinear effects are common. I offer guidance on how to detect and overcome mechanical biases. An application and several examples highlight that the suggested methods are broadly relevant and can inform policy and multiplier calculations.

Kilian Huber
Booth School of Business
University of Chicago
5807 S. Woodlawn Avenue
Chicago, IL 60637
and NBER
kilianhuber@uchicago.edu

Researchers in finance and macroeconomics are often interested in general equilibrium spillover effects: how shocks to some firms and households affect other parts of the economy. By quantifying spillovers, researchers can evaluate which general equilibrium channels need to be included in economic models and to what extent empirical estimates based on microdata are informative about other levels of aggregation.[1] Spillovers are particularly important when researchers study large-scale financial and macroeconomic shocks because many firms and households are simultaneously affected and spillovers are often large.

To understand why spillover analysis is helpful, consider two concrete examples at the regional level. An empirical literature estimates how regional house price shocks affect regional employment (Mian and Sufi 2014; Giroud and Mueller 2017). But a parameter required to calibrate macro-finance models is the direct effect of a house price change on an individual household (Guren et al. 2020). To convert regional estimates into the direct effect, one needs to know the magnitude of regional spillovers after a housing shock. Another literature shows that firms with an unhealthy bank grow more slowly than other firms in the same region with a healthy bank (Bentolila et al. 2018; Berg 2018). A regional policy maker may wonder how subsidizing the unhealthy bank will affect the entire regional economy and therefore will need to understand regional spillovers after a banking shock. In both cases, estimates of regional spillovers would allow one to convert existing estimates to another level of aggregation, even if direct estimates at the desired level of aggregation are not readily available.

The traditional approach to measuring spillovers in finance and macroeconomics is to calibrate a fully specified, general equilibrium model of the economy. Such models can flexibly quantify spillovers operating among firms and households in the same region, sector, country, or any other group. A weakness of the model-based approach is that results depend on hard-to-verify assumptions about which general equilibrium channels exist. In this paper, I study an alternative empirical approach: direct estimation of spillovers using quasi-experiments or experiments. This approach is

[1] I use the terms "general equilibrium spillover effects" and "spillovers" interchangeably. Both refer to the same concept: namely, the effects of shocks on prices, technology, and other features of the general economic environment. These effects operate not only at the country level but also at lower levels of aggregation, such as regions, sectors, or networks. They imply that shocks propagate beyond directly affected entities.

becoming increasingly popular in finance and macroeconomics. It allows researchers to quantify spillovers operating within groups of firms and households using a regression framework.

This paper offers econometric guidance on how to implement direct estimation of spillovers. I describe the general approach and then point out two sources of mechanical bias that are likely to arise in finance and macroeconomics: the existence of multiple types of spillover and mismeasured treatment status. I suggest practical methods to detect and overcome these biases. I illustrate the relevance of the proposed methods with an application to a real-world credit shock and describe further examples in which the methods would be germane. Finally, I argue that direct estimation of spillovers allows researchers to calculate the impact of policy and multiplier effects.

The method of spillover estimation

The paper begins with an empirical framework for the direct estimation of spillovers. For simplicity, I henceforth use "firms" to describe the unit of direct treatment, but the framework applies equally when households or other entities are the directly treated units. A researcher studies whether a shock to a subset of firms (the "treatment") generates spillovers onto other firms that are in the same "group" as treated firms. Firms that belong to the same group are in some way connected, for example, because they are in the same region, production network, technology space, or any other type of grouping. Direct estimation of spillovers requires identification of a treatment that is exogenous both across individuals and across groups. A group can even constitute an entire country, so that the estimated spillovers operate at the country level, as long as a researcher can identify exogenous variation in treatment at the country level and at least one lower level of aggregation (e.g., regions, sectors, firms, or households).[2]

To directly estimate spillovers, the researcher includes the average treatment status of all other firms in the same group in the regression (the "leave-out mean"). For example, if the researcher is interested in regional spillovers, one regressor is the average treatment status of all other firms in

[2]Recent work has attempted to identify exogenous country-level variation in fiscal and monetary policy plus variation at a lower level of aggregation. Examples include variation due to monetary policy abroad (Jiménez et al. 2012), large political upheavals (Fuchs-Schündeln 2008), or geopolitical developments (Conley et al. 2021). Such settings may be suitable for a spillover analysis at the country level (see Section 7).

the region. Direct estimation of spillovers using leave-out means has several attractive features. It is relatively easy to apply to existing research designs. It allows researchers to directly compare the magnitude of different types of spillovers by including multiple leave-out means in the regression. The method estimates a standard error on the spillover, which enables formal inference on whether spillovers are statistically significant, unlike methods that estimate direct and group-level effects in separate regressions.

Notwithstanding these advantages, I argue that direct estimation of spillovers raises difficult and underappreciated empirical challenges. I focus on two challenges that are common when researchers study large-scale financial and macroeconomic shocks: first, the presence of multiple types of spillovers and second, mismeasured treatment status due to nonlinearity or measurement error. These issues can mechanically bias estimates of both spillover and direct effects, even if the (quasi-)experimental variation defining direct and group treatment status is truly exogenous and if there are no omitted variables correlated with treatment. Standard (quasi-)experimental tools (e.g., testing for sample balance and parallel trends) do not solve these issues. I will discuss the two issues in turn.

Mechanical bias due to multiple spillover types

First, I consider the case of multiple spillover types. Spillovers operate across multiple groups after almost all large-scale shocks. A shock to firms can spill over to other firms through factor markets, product markets, input-output networks, and common lenders, to name a few relevant groups. Despite this empirical complexity, theoretical models typically do not account for all relevant spillover channels. For instance, urban models may include only a regional spillover, whereas industrial organization models may exclusively focus on a sectoral spillover. Motivated by theory, specialized researchers may then empirically test for only one potential spillover, without considering the others.

I explore the consequences of testing for only one spillover in situations where the true model contains multiple spillovers. To simplify the exposition, I consider a shock to firms that simultane-

ously spills over to two groups: to firms in the same region (e.g., through wages, as directly treated firms hire more on local labor markets) and to firms in the same sector (e.g., through output prices, as directly treated firms raise production). Throughout the paper, I use the concrete two-group example of regions and sectors. However, I do not mean to imply that these two groups cover all potential spillovers. The lessons apply more generally when the model contains many other groups.

Using the concrete two-group example, I show that testing for only a sectoral spillover can severely bias estimates if the true model also contains a regional spillover. In fact, the sectoral spillover estimate can have the wrong sign, leading to a complete misinterpretation of general equilibrium forces. The bias is present even if there is zero correlation between the regional and sectoral leave-out means (i.e., even when firms facing many treated firms in their sector are *not* more likely to face many treated firms in their region). In that sense, bias due to multiple spillovers is distinct from standard concerns about omitting correlated variables and it is not typically considered in applied papers. Intuitively, the bias occurs because directly treated firms are disproportionately found in regions and sectors with high average treatment, so omitting one relevant spillover term leaves a correlation between the error term and direct treatment status, biasing all coefficients in the regression.

Mechanical bias due to mismeasurement and nonlinear effects

The second estimation issue I discuss relates to misspecification of direct treatment status. I initially consider a modest degree of classical measurement error, as found in standard data sets (Bound and Krueger 1991). This type of error can generate large spillover estimates, even if true spillovers are zero, because part of the true direct effect erroneously loads onto the spillover estimate. Measurement error can bias spillover estimates in either direction, depending on the underlying data-generation process. This bias is therefore distinct from classical measurement error in settings without spillovers, which always biases coefficients toward zero. I show that mechanical bias due to mismeasurement also applies to the estimation of network effects, which is one

particular type of spillover (see Internet Appendix A).

A related type of bias arises if true direct effects depend nonlinearly on treatment status. Many financial shocks have this feature; for example, direct effects often exist only when individuals face binding liquidity, borrowing, or capital constraints (Brunnermeier and Sannikov 2014; Giroud and Mueller 2017, 2019; Cloyne et al. 2019). Researchers may not be aware of underlying nonlinearity and instead misspecify treatment using a linear regressor. For concreteness, imagine a model where the true spillover is zero and the true direct effect only occurs for observations where the direct treatment variable is positive. I simulate such models and find that spillover estimates can be much larger than direct estimates, falsely suggesting that group-level effects are primarily driven by spillovers rather than direct treatment.

Detecting and overcoming mechanical bias

I turn to detecting and overcoming the sources of mechanical biases discussed so far (due to multiple spillovers, mismeasurement, and nonlinearities). To detect whether mechanical bias drives results, I argue that researchers can test for heterogeneous effects. Economic theory often predicts which firms should be unaffected by a given type of spillover. For example, tradable firms do not respond to local demand spillovers (Moretti 2010; Mian and Sufi 2014; Giroud and Mueller 2017, 2019). If estimated regional demand spillovers are zero for tradable firms and only exist for nontradable firms, as theory predicts, spillover estimates are unlikely to be mechanically biased. If instead estimated spillovers are of similar magnitude for all types of firms, then mechanical bias is likely an issue.

Solutions to mechanical biases are available. To overcome bias due to multiple spillovers, researchers can include all relevant group-level leave-out means in the regression. However, this may be challenging in practice, as not all relevant connections between firms may be observed in standard data sets, the full set of relevant spillover channels may not be predictable ex ante, and regressions may be underpowered with many regressors. Researchers can also explore flexible functional forms to identify nonlinear direct effects. An instrumental variable (IV) that is correlated

with individual treatment status and uncorrelated with the treatment status of other firms in the group solves all forms of mechanical bias, but may be difficult to find. Taken together, the findings on mechanical biases highlight that researchers should interpret spillover estimates with caution and carefully consider potential solutions.

Application and further examples of spillover estimation

I illustrate the relevance of the biases and solutions using an application. I study an exogenous credit disruption by a large German bank called Commerzbank (Huber 2018). Direct treatment status measures whether a firm had a banking relationship with Commerzbank. Directly treated firms reduced employment when their bank cut credit. Guided by a simple industrial organization model, I initially only test for spillovers among firms in the same product market. I find a significant product market spillover of similar magnitude to the direct effect. However, urban models suggest that local demand and agglomeration forces might also generate spillovers. When I additionally test for a regional spillover, the product market spillover shrinks and becomes insignificant, while the regional spillover is large and significant. This result illustrates that the presence of multiple spillovers can lead to severely misguided conclusions about the nature of spillovers.

These findings leave open the possibility that other, omitted spillover types explain the regional spillover. I investigate this possibility by testing for heterogeneous effects. I identify a subset of sectors that, according to theory, are strongly affected by regional spillovers: nontradable sectors (due to local demand effects) and high-innovation sectors (due to local agglomeration effects). I find that regional spillovers are only significant for firms in such sectors and insignificant for other sectors. Mechanical bias would affect all firms, so this heterogeneity suggests that regional spillovers are not driven by mechanical bias.

Next, I investigate mismeasurement, by introducing measurement error into the direct treatment variable. The direct effect becomes insignificant and close to zero, while the regional spillover more than doubles in size. This falsely suggests that the entire regional effect is driven by spillovers, with directly treated firms not growing any differently to untreated firms. However, the hetero-

geneity test is particularly useful here. I find that spillovers are large and significant for all types of firms. This reveals that the spillover estimates based on mismeasured data are partially driven by mechanical bias and not by the theoretical forces posited in urban models.

I emphasize the usefulness of spillover estimation for policy. Consider a public credit program that extends loans to firms directly affected by a bank lending cut. A naive calculation based only on the direct effect suggests that this program would raise regional employment by 0.4 jobs for US$100,000 of lent funds. In contrast, direct estimation of spillovers implies much larger gains of 1.4 jobs. The calculation requires knowledge of both direct and spillover effects and would not be possible based on regional estimates only. The application results also offer lessons for models, as they imply that realistic general equilibrium models need to include strong regional amplification forces.

In the final section of the paper, I discuss additional examples of shocks hitting firms, households, and regions that lend themselves to spillover estimation. I describe which types of spillovers are relevant in different settings and how researchers can overcome mechanical biases in each case. The examples highlight the broad relevance of spillover estimation and of the issues discussed in this paper.

Checklist for applied researchers

To summarize this paper's lessons, I present a checklist for applied researchers seeking to estimate spillovers.

1. **Set up direct estimation of spillovers**

    (a) Define spillovers of interest that are to be estimated (e.g., spillovers within households in a region, within firms in a sector, or within regions in a country). To do this, identify individual units that can be directly affected by shocks as well as groups of individual units among which spillovers operate. Groups can be any combination of firms or households (e.g., regions, sectors, countries). Individual units can be any smaller level

of aggregation contained within groups (e.g., individual households within regions, firms within sectors, or regions within a country).

(b) Identify an exogenous shock where treatment intensity varies for individual units within groups as well as across groups.

(c) Estimate direct and spillover effects in the same specification by using direct treatment status and group-level leave-out means as regressors.

2. **Detect mechanical bias using heterogeneity tests**

(a) Use theory to understand which mechanisms drive the spillovers of interest (e.g., regional spillovers operate through local demand and agglomeration effects; sectoral spillovers operate through changes in competition on product markets; spillovers across regions operate through trade, migration, capital mobility, and country-level policy).

(b) Identify individual units that, according to mechanisms predicted by theory, should be less affected by spillovers (e.g., tradable firms in low-innovation sectors respond less to local demand and agglomeration spillovers; firms with high market power react less to shocks to other firms in their product market; and autark regions are less exposed to cross-regional spillovers).

(c) Test whether spillover estimates are heterogeneous in line with theory. If spillover estimates are homogeneous, the results may be driven by mechanical bias and should be interpreted with caution.

3. **Overcome potential mechanical bias**

(a) Mechanical bias can be a problem even if the shock is truly exogenous and if there are no omitted variables correlated with treatment. Omitted spillover types, nonlinear direct effects, and measurement error can cause mechanical bias. Address each source of bias in turn.

(b) To overcome omitted spillover types, try to measure leave-out means for other groups where spillovers may operate. Include these additional leave-out means in the regression to test the robustness of spillover estimates. (See Section 7 for which types of spillovers are likely relevant in different types of analyses.)

(c) To overcome nonlinear direct effects, explore flexible functional forms (e.g., use bins for different parts of the treatment distribution as regressors).

(d) All forms of mechanical bias can be overcome by finding an instrument that is correlated with individual treatment status but uncorrelated with the treatment status of other firms in the group (e.g., another mismeasured treatment variable can serve as instrument).

1 Related Literature

This paper relates to the methodological discussion in finance and macroeconomics on how to convert estimates from microdata to higher or lower levels of aggregation. Most of this literature relies on structural, model-based approaches (Browning, Hansen, and Heckman 1999; Acemoglu 2010; Nakamura and Steinsson 2018).[3] Direct estimation of spillovers using (quasi-)experimental variation has traditionally not played a large role. For example, no paper published in the leading economics and finance journals in 2017 jointly analyzes direct and spillover effects using the direct estimation method.[4]

In recent years, however, researchers in finance and macroeconomics have started estimating regional and sectoral spillovers (Dupor and McCrory 2018; Huber 2018; Bernstein et al. 2019;

[3]Recent examples include Li, Whited, and Wu (2016), Auclert, Rognlie, and Straub (2018); Auclert, Dobbie, and Goldsmith-Pinkham (2019), Beraja, Hurst, and Ospina (2019), Guren et al. (2020), Chodorow-Reich, Nenov, and Simsek (2021), and Herreño (2021). Sarto (2018), Adão, Arkolakis, and Esposito (2020), and Wolf (2021) discuss alternative methods that are less reliant on structural assumptions.

[4]The journals published 610 papers in 2017 and are: *American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, Review of Economic Studies, Journal of Finance, Journal of Financial Economics,* and *Review of Financial Studies*. Seven papers in these journals explicitly analyze some form of spillover in a quasi-experimental research design. Three of these seven papers are in the subfield of corporate finance and none is in the other parts of finance and macroeconomics. See Berg, Reisinger, and Streitz (2021) for more discussion on publications using differences-in-differences and spillover estimation.

Auerbach, Gorodnichenko, and Murphy 2020; Gathmann, Helm, and Schönberg 2020; Helm 2020; Verner and Gyöngyösi 2020; Conley et al. 2021). These papers pay little attention to potential mechanical biases in spillover estimates. While existing papers are well versed in the standard (quasi-)experimental toolkit (e.g., inspecting the IV exclusion restrictions through balancing tests), these tools do not overcome mechanical biases.

The contribution of this paper is to offer econometric advice tailored to estimating spillovers in finance and macroeconomics. I focus on how researchers can design spillover estimation and on mechanical biases arising from multiple spillover types, mismeasurement, and nonlinearities. These biases are particularly relevant to researchers studying large-scale financial and macroeconomic shocks. For one, spillovers after such shocks are inherently complex, operate across multiple overlapping groups, and theory makes no strong predictions about which spillover types are relevant. In addition, nonlinear effects are common in financial settings (e.g., because ofliquidity constraints or regulatory capital thresholds) and treatment is often difficult to measure (e.g., banking relationships). This paper's focus on actionable solutions to underappreciated methodological challenges is inspired by influential earlier work in other areas (Petersen 2009; Gormley and Matsa 2014; Chodorow-Reich 2019; Lerner and Seru 2022).

Berg, Reisinger, and Streitz (2021) provide complementary methodological advice on estimating models with spillovers. Readers may find a brief comparison useful. Berg, Reisinger, and Streitz (2021) focus on how to estimate direct effects in models with a single spillover term, while I emphasize how to quantify (potentially multiple) spillover effects and how to use spillovers to inform policy calculations. In terms of empirical advice, Berg, Reisinger, and Streitz (2021) recommend that researchers investigate whether spillovers are heterogeneous between directly treated and untreated firms, while I show that heterogeneity tests based on theory, multiple spillover terms, flexible functional forms, and instruments can overcome several sources of mechanical bias.[5] While Berg, Reisinger, and Streitz (2021) also show theoretically how spillovers can arise, I focus on empirical advice. Finally, Berg, Reisinger, and Streitz (2021) and I study the same empirical ap-

[5] In my application, I find no evidence for heterogeneous spillovers. That means that regional or sectoral spillovers did not differ between firms directly treated by the Commerzbank credit shock and other firms.

plication, the Commerzbank credit shock. In another methodological paper, Mian, Sarto, and Sufi (2022) show theoretically how regional spillovers arise in a general equilibrium model and discuss identification assumptions in the context of regional credit shocks.[6] Gabaix and Koijen (2022) describe how to estimate spillovers using "granular IV."

Outside of finance and macroeconomics, several applied papers estimate spillovers directly, mainly in education (reviews in Epple and Romano 2011; Sacerdote 2011; List, Momeni, and Zenou 2019), development (RCTs in Miguel and Kremer 2004; Angelucci and De Giorgi 2009; Janssens 2011; Muralidharan, Niehaus, and Sukhtankar 2017; Cunha, De Giorgi, and Jayachandran 2019; Filmer et al. 2021; Egger et al. forthcoming), and public economics (Blundell et al. 2004; Rincke and Traxler 2011; Crépon et al. 2013; Ferracci, Jolivet, and van den Berg 2014; Lalive, Landais, and Zweimüller 2015; Gautier et al. 2018; Boning et al. 2020). Mechanical bias due to multiple spillovers is not discussed in these papers, likely because all relevant spillover types are ex ante defined and observed by the authors, unlike in typical settings in finance and macroeconomics.[7] Similarly, bias due to nonlinearities is not discussed (except in Angrist 2014), likely because sharp nonlinear effects are theoretically and empirically more relevant for financial shocks. Measurement error is a more common problem and has been studied mainly in the context of educational peer effects (Ammermueller and Pischke 2009).

The classic econometrics literature emphasizes different econometric challenges compared to this paper, namely, that spillovers violate the stable unit treatment value assumption (Rubin 1980, 1990) and are difficult to estimate in the absence of exogenous variation (Manski 1993; Moffitt

[6]The estimates in Huber (2018) and Mian, Sarto, and Sufi (2022) lead to remarkably similar conclusions on the magnitude of regional spillover effects, despite the different settings. Huber (2018, table 11) and this paper (Section 6.4) find that local general equilibrium effects account for roughly 60%–70% of the total regional effect, whereas the corresponding number in Mian, Sarto, and Sufi (2022) is 80%. The methods differs slightly, as Mian, Sarto, and Sufi (2022) compare individual-level and region-level regressions, whereas I directly estimate spillovers (and associated standard errors) using a leave-out mean. However, the mechanical issues discussed in this paper are relevant to all these methods.

[7]In education economics, the typical objects of interest are peer effects operating within a classroom or school. In development economics, units are usually self-contained villages (i.e., "largely closed local economies," as put by Egger et al. forthcoming). In public economics, researchers typically focus on a policy that affects well-defined local labor markets or social groups. In all these setting, spillovers chiefly operate through the predefined group of interest rather than through other groups, unlike in macroeconomics and finance where cross-regional sectors, financial markets, input-output chains, etc., play a role.

2001; Glaeser, Sacerdote, and Scheinkman 2003; Bramoullé, Djebbari, and Fortin 2009). Subsequent work develops techniques to optimally estimate spillovers with randomized controlled trials (Duflo and Saez 2003; Hirano and Hahn 2010; Avitabile 2012; Baird et al. 2018; Vazquez-Bare forthcoming) and in the absence of data on group membership (Manresa 2016; Breza et al. 2020).

2 Empirical Framework

2.1 Basic model of direct and spillover effects

Consider an economic shock that affects firms or households with varying intensity, as indicated by their "treatment status." For example, if the shock is a credit supply disruption, treatment status is a firm's dependence on failing banks. If the shock is fiscal stimulus, treatment status measures whether firms or households receive a stimulus check.

Economic theory suggests that the treatment status of a given firm or household can affect the outcomes of other firms and households. For instance, if two firms are located in the same region, they hire on the same local labor market. When one firm is treated, it may change its labor demand, thereby affecting local wages. If two firms sell substitute products, they are competitors. When a competitor is treated, product prices may change, affecting all firms in the product market sector. In general, whenever firms are in some way connected, the treatment status of one firm can generate spillovers onto other firms. Similarly, whenever households are connected, there can be cross-household spillovers. For simplicity, I henceforth use "firms" to describe the unit of observation, but the analysis applies equally when households, regions, sectors, or other entities are directly treated.

While in reality many channels connect firms, to simplify exposition, I assume that there are just two: spillovers may operate among firms in the same region and in the same product market sector. However, the findings on the mechanical biases presented below hold generally for a larger number of groups as well as for settings in which researchers estimate aggregate spillovers among firms in the same country (see Section 7).

The treatment status of an individual firm i in region $r(i)$ and sector $s(i)$ is given by x_i. An outcome, such as firm investment or employment growth, is given by y_i. Assuming linearity, the relationship between outcome and treatment status of various firms is

$$y_i = \beta x_i + \sum_{j \neq i, r(j)=r(i)} \gamma^j x_j + \sum_{k \neq i, s(j)=s(i)} \lambda^k x_k + \alpha + \varepsilon_i. \tag{1}$$

The first coefficient β is the direct effect of individual treatment (x_i) on the outcome. The direct effect represents by how much the outcome would change if firm i alone was treated. In addition, there are spillover effects γ^j and λ^k. Each spillover effect represents by how much outcome y_i of firm i would change if another firm in the same region (firm j with treatment status x_j) or in the same sector (firm k with treatment status x_k) was treated.[8] Throughout the paper, I assume that treatment of all firms is exogenous, such that $E(x_i \varepsilon_i) = 0 \; \forall i$ (see below for a detailed discussion).

The superscripts on the coefficients γ^j and λ^k indicate that spillover effects are firm-specific, so that spillovers arising from two different firms may not be identical. In practice, however, it is difficult to estimate one spillover coefficient per firm. To facilitate estimation, researchers commonly assume that spillovers are homogeneous for firms in the same sector or region:

$$\gamma^j = \frac{\gamma}{N_{r(j)} - 1} \; \forall j, \tag{2}$$

$$\lambda^k = \frac{\lambda}{N_{s(j)} - 1} \; \forall k. \tag{3}$$

The number of firms in a region and sector is $N_{r(j)}$ and $N_{s(j)}$, respectively. Intuitively, the assumptions imply that the greater the number of firms in a region or sector, the less important the region- or sector-level spillovers generated by an individual firm.

Under these assumptions, the outcome depends on only three coefficients: individual treatment

[8] This equation assumes that spillover effects affect firms equally, i.e., γ^j and λ^k do not depend on firm i. The framework can be generalized to account for heterogeneous spillovers by characteristics of firm i, as shown in the application in Section 6.

status and two "leave-out means":

$$y_i = \beta\, x_i + \gamma \overline{x_{r(i)}} + \lambda\, \overline{x_{s(i)}} + \alpha + \varepsilon_i. \tag{4}$$

The leave-out mean $\overline{x_{r(i)}}$ is the average treatment status of all other firms in region $r(i)$ apart from firm i:

$$\overline{x_{r(i)}} = \frac{\sum_{j \neq i,\, r(j)=r(i)} x_j}{N_{r(i)} - 1} \tag{5}$$

and $\overline{x_{s(i)}}$ is defined analogously. The coefficients γ and λ are the region- and sector-level spillovers, whereas β is the direct effect.

The assumptions in Equations (2) and (3) imply that spillovers are identical across firms in a group. Alternatively, researchers may prefer assuming that spillovers are activity-weighted (e.g., that firms with more workers generate larger spillovers). The framework above can be easily adapted to incorporate this assumption. Treatment can be defined at the level of individual workers, so that x_{pi} measures the treatment status of worker p employed at firm i. The direct treatment status of a firm is then given by the average treatment status of all workers at firm i:

$$\widetilde{x_i} = \frac{\sum_{p \in i} x_{pi}}{\widetilde{N_i}},$$

where the number of workers at firm i is $\widetilde{N_i}$. (If treatment is determined at the firm level, x_{pi} is identical for all workers at firm i, so that $\widetilde{x_i} = x_i$, where x_i is simply the treatment status of the firm from Equation (4).) If we then assume that spillovers are identical for workers in the same sector or region, the model becomes

$$y_i = \widetilde{\beta}\, \widetilde{x_i} + \widetilde{\gamma} \widetilde{x_{r(i)}} + \widetilde{\lambda}\, \widetilde{x_{s(i)}} + \widetilde{\alpha} + \widetilde{\varepsilon_i}.$$

Here, the regional leave-out mean is the average treatment of all workers employed by other firms in a region:

$$\widetilde{x_{r(i)}} = \frac{\sum_{j \neq i, r(j)=r(i)} x_{pj}}{\widetilde{N_{r(i)}} - \widetilde{N}_i}, \tag{6}$$

where the total number of workers in the region is $\widetilde{N_{r(i)}}$. The sectoral leave-out mean $\widetilde{x_{s(i)}}$ can be defined analogously.

2.2 Variation in treatment is exogenous

I assume that individual treatment status as well as treatment status of firms in the same region and sector is exogenous to all other determinants of firm outcomes:

$$E\left(x_i \varepsilon_i\right) = 0 \; \forall i.$$

In practice, exogenous variation means that researchers have either experimentally randomized treatment status or identified quasi-random variation. As a result of this assumption, all estimation issues described below are not driven by the usual endogeneity concerns about correlations between treatment and unobserved errors (as in Manski 1993; Moffitt 2001). As shown below, the issues I discuss are more subtle and depend on the distribution of treatment across regions and sectors.

Exogenous variation is a high bar in practice. In many studies, variation in direct treatment may be exogenous within region and sector, but variation in treatment of firms in the same region and sector is not. For instance, exposure to failing banks may be exogenous when comparing firms within regions, but the distribution of failing banks across regions may be correlated with other shocks to firm growth. In such cases, the group definition fails the exogeneity criterion and cannot be used to estimate region-level spillover effects.

2.3 Treatment may vary systematically across regions and sectors

I assume that direct treatment status depends on several random variables:

$$x_i = u_{r(i)} + u_{s(i)} + z_i + v_i, \tag{7}$$

where $u_{r(i)}$ is a common factor for all firms in region $r(i)$ and $u_{s(i)}$ is a common factor for all firms in sector $s(i)$. The other components vary at the individual level: z_i is an observed variable, which is uncorrelated within regions and sectors and can serve as instrument for x_i, and v_i is an unobserved random error. The variables $u_{r(i)}$, $u_{s(i)}$, z_i, and v_i are uncorrelated with each other and with the error ε_i in Equation (4).

If $u_{r(i)}$ is identical across regions and $u_{s(i)}$ is identical across sectors, treatment status does not vary systematically across regions and sectors. However, variation across regions and sectors is systematic in most research designs. Variation is always systematic in experiments in which researchers intentionally treat some groups more than others. In most naturally occurring settings, variation is also systematic. For instance, exposure to the 2008 credit crisis varied systematically across regions and sectors because banks tend to specialize in certain regions and sectors, rather than picking borrowers at random (Chodorow-Reich 2014; Bentolila et al. 2018; Huber 2018). As a result, certain areas and sectors were systematically more exposed to failing banks. Similarly, fiscal stimulus tends to be concentrated in specific regions (Chodorow-Reich 2019).

On the positive side, systematic variation guarantees a large degree of variation across regions and sectors when the number of firms per region and sector is large, making it easier to estimate spillovers. In contrast, when variation across regions and sectors is not systematic, there will be little variation when groups are large, making it difficult to precisely estimate spillovers.

The challenge is that naturally occurring systematic variation is often not exogenous. The factors generating systematic variation may also drive differences in firm outcomes across groups. For example, failing banks might be more likely to operate in regions with low growth potential. This would generate a correlation between the leave-out mean and other shocks to firm growth

(correlation between $u_{r(i)}$ and the error term in Equation (4)). For the purpose of this paper, I leave aside concerns about exogeneity and focus on other issues.

2.4 Setup for the simulations

I investigate the properties of spillover estimates by running simulations. In each simulation, I randomly sort 5,000 observations (indexed by i) into 500 equally sized regions and 500 equally sized sectors. In the baseline simulations, I assume that the region and sector terms $u_{r(i)}$ and $u_{s(i)}$ are both independently and lognormally distributed with a mean of zero and a standard deviation of one. This implies that variation is systematic across regions and sectors in the baseline simulations. In additional simulations, I assume that variation is not systematic, in which case $u_{r(i)}$ and $u_{s(i)}$ are zero. ε_i, z_i, and ν_i are normally distributed with a mean of zero and a standard deviation of one. Throughout the paper, I report coefficients and standard errors averaged over 100 simulations.

3 Interpreting the Magnitude of Spillover Coefficients

Spillover estimates allow researchers to calculate the share of a group-level outcome that is due to spillover effects versus direct effects. For instance, to calculate the share of a regional outcome due to spillover effects, researchers can take region-level averages of Equation (4):

$$\overline{y}^{r(i)} = (\beta + \gamma)\, \overline{x}^{r(i)} + \lambda\, \overline{x_{s(i)}}^{r(i)} + \alpha + \overline{\varepsilon}^{r(i)}, \tag{8}$$

where the average outcome in region $r(i)$ is $\overline{y}^{r(i)}$, average treatment status in region $r(i)$ is $\overline{x}^{r(i)}$, and the average sectoral leave-out mean in region $r(i)$ is $\overline{x_{s(i)}}^{r(i)}$.[9] The total effect is the change in the average regional outcome relative to the change in average regional treatment:

$$Total\ Effect = \frac{d\overline{y}^{r(i)}}{d\overline{x}^{r(i)}} = \beta + \gamma,$$

[9] To derive this equation, note that the average regional leave-out mean in region $r(i)$ is equivalent to the average treatment status in region $r(i)$. In the notation here, $\overline{x_{r(i)}}^{r(i)} = \overline{x}^{r(i)}$.

whereas the direct effect is change in the average regional outcome, assuming no spillovers:

$$Direct\ Effect = \frac{d\overline{y}^{r(i)}}{d\overline{x}^{r(i)}} \mid (\gamma = 0) = \beta.$$

The share of the total regional effect due to direct effects is $\frac{\beta}{\beta+\gamma}$ and the share due to spillovers is $\frac{\gamma}{\beta+\gamma}$. In general, ratios are a useful way to report coefficient magnitudes. Spillover coefficients on their own often do not have a natural interpretation because they typically do not capture absolute effects, but rather, effects relative to a control firm for whom direct treatment and spillover effects are all zero.[10]

Researchers may find it useful to present results in form of a dollar multiplier, which is the total effect ($\beta + \gamma$) relative to the average dollar spending by policy makers required to achieve the direct effect. For instance, imagine that treatment is a policy that spends US$$M$ on the average firm. Then the regional multiplier is:

$$Multiplier = \frac{\beta + \gamma}{M}.$$

If treatment is already measured in dollar values, then M is simply equal to $\overline{x}^{r(i)}$, the average treatment amount received by firms. However, if treatment is not in dollar values, M may need to be estimated. For instance, if treatment is an indicator for firms exposed to an exogenous decrease in product demand or credit supply, then researchers may not immediately know the dollar value needed to offset the initial shock. In such cases, researchers first need to estimate a direct effect on the decisive dollar outcome (such as firm revenue or credit) and then convert that direct effect into an average dollar treatment to get M (see the application in Section 6.4).

In some settings, researchers may want to estimate spillover coefficients using an outcome

[10]To be clear, in the example using regions and sectors, γ and λ measure direct and spillover effects relative to a firm that was not directly exposed to the shock ($x_i = 0$) and in whose region and sector no other firm was directly exposed to the shock ($\overline{x_{r(i)}} = \overline{x_{s(i)}} = 0$). This means that the coefficients do not capture the total difference in firm outcomes relative to a world where the shock did not happen. Instead, they capture the effect of treatment relative to firms that were treated neither directly nor through spillovers. See Chodorow-Reich (2020) for a formal discussion of relative versus absolute effects, which is not the focus of this paper.

measured in percentage changes (e.g., to account for outliers) and then report effects measured in absolute units (e.g., to relate results to a model or other estimates). For instance, the regression outcome may be the percentage change in firm employment (which is approximately equal to the log change or the symmetric growth rate), but researchers may want to know the effect on total jobs. This conversion is trivial if average firm size is roughly constant. In these cases, researchers can approximate the total effect on jobs by multiplying the percentage effects with the average pretreatment size (given by $\bar{y}$). The total effect then becomes $(\beta + \gamma) \times \bar{y}$ and the direct effect $\beta \times \bar{y}$, with the share of the total effect due to spillovers remaining unchanged.

The conversion from percentage change to absolute effect is more involved if firm size varies with treatment. I describe a practical conversion method in detail in Internet Appendix B. In short, the method splits the sample into a number of size bins k and calculates the effect on absolute values in each bin separately. The absolute total effect on the region is then the weighted average of absolute effects in each size bin:

$$Total\ Effect \approx \sum_k \left[\left(\beta\, d\bar{x}^k + \gamma d\overline{x_{r(i)}}^k \right) \times \bar{y}^k \times \omega^k \right], \tag{9}$$

where $d\bar{x}^k$ is the average direct treatment status in k, $d\overline{x_{r(i)}}^k$ is the average leave-out mean in k, $\bar{y}^k$ is the average pretreatment outcome of firms in size bin k, and ω^k is the fraction of firms in k. Correspondingly, the absolute direct effect is:

$$Direct\ Effect \approx \sum_k \left[\beta\, d\bar{x}^k \times \bar{y}^k \times \omega^k \right]. \tag{10}$$

Choosing the number of size bins involves a trade-off. A greater number of bins increases the accuracy of the method, but many small bins lead to measurement error in the averages. Ultimately, researchers should choose as many bins as possible until averages become statistically indistinguishable.

4 Bias due to Multiple Spillovers

Having laid out the empirical framework, I highlight practical difficulties that arise when estimating spillovers. In this section, I show that spillover estimates can be biased if there are multiple potential spillover types and I suggest ways to detect bias.

4.1 Testing the wrong type of spillover

I assume that no true spillover occurs within sectors, but a spillover occurs within regions with a coefficient of one. The true data-generating process is thus:

$$y_i = x_i + \overline{x_{r(i)}} + \varepsilon_i. \tag{11}$$

Treatment varies systematically across regions and sectors (that is, $u_{r(i)}$ and $u_{s(i)}$ in Equation (7) are not identical across regions and sectors).

Researchers may not include all relevant spillovers in their specification. Theoretical models often focus on one type of spillover mechanism. For instance, industrial economists focus on competition, so their research question might only consider spillovers within product markets. Financial economists study credit reallocation, so they might be interested in spillovers among borrowers of the same bank only. Based on theory, researchers may be drawn to empirically investigating only one type of spillover, even if that spillover does not appear in the true model.

Measurement difficulties are another reason researchers may overlook relevant spillover types. Some economic connections between firms are not recorded in standard data sets. For example, the default of one firm might generate capital constraints for a lender, but lender identities are often not observed in the data. In practice, the range of possible spillover channels is large. As a result, researchers may not be able to include all relevant spillover forces in their specifications.

4.2 Bias due to testing the wrong type of spillover

Testing for a spillover with zero coefficient while omitting a true spillover biases estimates. The bias arises even if the included and omitted leave-out means are not correlated. The lack of correlation makes this form of biased spillover estimate less salient and detectable relative to standard forms of omitted variable bias. In Equation (11), the regional and sectoral leave-out means are uncorrelated by construction, so it is not obvious that the coefficient for the regional leave-out mean should be biased.

To illustrate the effects of ignoring relevant spillovers, I use the simulated data based on Equation (11) and run regressions that only contain direct treatment status x_i and the sectoral leave-out mean $\overline{x_{s(i)}}$. The true ratio of regional spillover to direct effect is a positive 100%. In contrast, the estimated sectoral spillover coefficient for $\overline{x_{s(i)}}$ is negative and significant (Table 1, column 1). The ratio of estimated spillover to direct effect is large at -33%. The focus on the wrong spillover therefore changes the sign of estimated spillovers and leads to a severe misinterpretation of economic forces.

The reason for the bias is the presence of systematic variation across regions and sectors. When x_i and $\overline{x_{s(i)}}$ are the only regressors, the omitted $\overline{x_{r(i)}}$ enters the error term. Both x_i and the omitted $\overline{x_{r(i)}}$ are functions of the regional factor $u_{r(i)}$ (Equation (7)). As a result, x_i and the error term are positively correlated, which biases the estimated coefficient for x_i. The spillover estimate is then also biased because x_i and $\overline{x_{r(i)}}$ are positively correlated (because of the common factor $u_{r(i)}$).

4.3 Overcoming bias due to testing the wrong type of spillover

How can researchers detect mechanical bias due to multiple spillover types? Note that the bias appears mechanically in all subgroups of firms and households because of the way leave-out means are constructed. But in many settings, economic theory predicts that spillovers should not exist among a certain subgroup. For instance, firms selling nontradable goods are affected by regional demand shocks, but firms selling tradables are not (Mian and Sufi 2014). The spillover coefficient for tradable producers should be zero if regional demand drives spillovers. If the estimated

spillover on tradable producers is indeed zero, researchers can conclude that the nonzero spillover estimate on other firms is not a mechanical bias due to including the wrong type of spillover. If instead significant spillovers show up for all firm types, the spillover estimates are likely biased.

An obvious solution to the bias is to control for other potential spillover types by including additional leave-out means. For instance, including the regional leave-out mean overcomes the bias (Table 1, column 2).[11] Controlling for other spillover types is only possible if data defining groups are available and if the number of groups is large enough to generate statistical power.

Instrumental variables can also solve the bias, although instruments are difficult to find in practice. Ideally, researchers identify an instrument at the individual level, such as z_i (as in Equation (7)). The instrument needs to be correlated with individual treatment status x_i, but uncorrelated with the treatment status of other firms in the group (i.e., uncorrelated with $u_{r(i)}$). The sectoral leave-out mean $\overline{z_{s(i)}}$ can then serve as instrument for $\overline{x_{s(i)}}$. Coefficients based on instrumenting for x_i and $\overline{x_{s(i)}}$ using z_i and $\overline{z_{s(i)}}$ are consistent (Table 1, column 3). In the absence of an instrument at the individual level, researchers can still estimate the spillover coefficient consistently by using an instrument for only the leave-out mean $\overline{z_{s(i)}}$ and controlling for x_i (see also the use of a group-level instrument in Huber 2018).

Finally, note that multiple spillover types do not lead to bias if there is no systematic group-level variation (column 4). This requires that both u_r and u_s are identical across regions and sectors, respectively (Equation (7)). However, this condition is often not met in large-scale shocks, as outlined in Section 2.3.

4.4 Bias due to multiple nonzero spillovers

The bias is not limited to the case in which a spillover with zero coefficient is included in the model. In an additional simulation, I assume that spillovers operate within regions and sectors

[11] If researchers are only interested in estimating a direct effect in a single spillover model, they can weight observations in the manner suggested by Baird et al. (2018, footnote 23). However, this approach does not estimate spillover effects and does not easily translate to models with multiple spillovers.

with a coefficient of one. The true data-generating process is thus:

$$y_i = x_i + \overline{x_{r(i)}} + \overline{x_{s(i)}} + \varepsilon_i. \tag{12}$$

However, as above, researchers include only the sectoral leave-out mean, possibly because they follow a model focused on product market competition or because they are unable to observe firm region. The direct effect is biased upward and sectoral spillover downward (Table 2, column 1). While the true ratio is 100%, the estimated ratio is 29%. Instrumenting (column 2) and controlling for all relevant spillover types (column 3) overcome the bias.

5 Bias due to Mismeasurement and Nonlinear Direct Effects

In this section, I outline how misspecification of treatment status biases spillover estimates. I describe two cases: classical measurement error and mismeasurement due to nonlinear effects.

5.1 Definition of measurement error

To illustrate the role of classical error, I assume that there is only a regional spillover, given by γ, so that the true data-generating process is:

$$y_i = \beta x_i + \gamma \overline{x_{r(i)}} + \varepsilon_i. \tag{13}$$

Imagine that direct treatment status x_i is measured with error. The observed treatment status is

$$x_i^* = x_i + \eta_i.$$

Measurement error η_i is normally distributed with a mean of zero and a standard deviation of σ. It is uncorrelated with ε_i, $u_{r(i)}$, $u_{s(i)}$, z_i, and ν_i. The leave-out mean is constructed from individual-

level data, so measurement error affects the observed leave-out mean too:

$$\overline{x_{r(i)}}^{*} = \overline{x_{r(i)}} + \overline{\eta_{r(i)}}.$$

The distortion caused by measurement error can be measured using the signal-to-total variance ratio, which is:

$$STV = \frac{V\left[x_i\right]}{V\left[x_i^*\right]}.$$

The greater the standard deviation of the measurement error, the lower the information content of the observed variable.

5.2 Bias due to measurement error

Using simulated data, I illustrate how estimates of spillovers depend on classical error. I generate data based on Equation (13), assuming no spillover effect ($\gamma = 0$). In the absence of measurement error ($STV = 1$), the regression results are consistent. The estimated direct effect (coefficient for x_i^*) is close to one and significant, whereas the estimated spillover (coefficient for $\overline{x_{r(i)}}^{*}$) is small and insignificant (Table 3, panel A, column 1).

With low measurement error ($STV = 0.95$), the spillover becomes statistically significant. The ratio of spillover to direct effect rises to 15% (column 2). The greater the measurement error, the greater the spillover estimate. Bound and Krueger (1991) document that measurement error in earnings growth in the Current Population Survey leads to $STV = 0.7$. The ratio of spillover to direct effect is 101% with $STV = 0.7$ (column 4). Hence, with an empirically plausible degree of measurement error, the estimated spillover is more than twice as large as the estimated direct effect, even though the true spillover is zero.

The intuitive reason for the overestimated spillover is the presence of systematic group-level variation (i.e., the variation in $u_{r(i)}$). The individual measurement error is partially averaged out when calculating $\overline{x_{r(i)}}^{*}$. As a result, $\overline{x_{r(i)}}^{*}$ contains relatively less measurement error than x_i^* and relatively more information about the group-level component $u_{r(i)}$. That means some of the true

direct effect (the part caused by high $u_{r(i)}$) shows up in the spillover estimate.

5.3 The direction of bias due to measurement error

The examples so far have shown that measurement error can inflate a spillover estimate when the true spillover is zero. In general, measurement error can cause bias in either direction (Ammermueller and Pischke 2009). Algebraically, the spillover estimate from specification 13 converges to:

$$plim\ \widehat{\gamma} = \beta C_1 + \gamma C_2,$$

where $0 \leq C_1$; $0 \leq C_2 \leq 1$; and $C_1 = 0$ if $u_{r(i)}$ is identical across regions.[12]

This equation shows that the spillover estimate is always attenuated if variation is not systematic (i.e., $u_{r(i)}$ is identical across regions). If there is systematic variation (i.e., $u_{r(i)}$ varies across regions), the relative magnitude of direct and spillover effects determines the bias. If the true direct effect is nonzero and the true spillover is zero ($\beta \neq 0$ and $\gamma = 0$), the direction of bias of the spillover estimate has the sign of the direct effect. If the true direct effect is zero and the true spillover is nonzero ($\beta = 0$ and $\gamma \neq 0$), the spillover estimate is attenuated.

To illustrate this result, I generate data where the true direct and spillover effects are both one ($\beta = \gamma = 1$). Under systematic variation, the spillover is overestimated (Table 3, panel B, column 1). Under random group-level variation ($u_{r(i)} = 0$), the spillover is attenuated (column 2).

5.4 Bias due to nonlinear direct effects

Nonlinear responses to shocks are common in financial settings. For instance, liquidity-constrained households extract housing equity when house prices go up, but do not inject equity when house prices fall (Cloyne et al. 2019). Similarly, large losses in bank capital have disproportionate effects on lending and real outcomes, relative to small losses (Brunnermeier and Sannikov 2014).

[12]The full derivation and definitions of C_1 and C_2 are in Internet Appendix C.

Researchers may not be aware of the underlying data-generating process, however, and mismeasure direct treatment status. Standard practice is to use linear regressors. This introduces a similar bias as classical measurement error. I illustrate this bias by specifying the true data-generating equation as

$$y_i = w_i + \varepsilon_i,$$

where w_i is a nonlinear variable based on an observed x_i:

$$w_i = \begin{cases} x_i & if\ x_i > 0, \\ 0 & otherwise. \end{cases}$$

The true spillover effects in the model are zero.

If researchers correctly account for the nonlinear relationship between y_i and x_i, the regression produces consistent estimates. The estimated direct coefficient for w_i is close to one and significant, whereas the regional spillover coefficient for $\overline{w_{r(i)}}$ is small and insignificant (Table 4, column 1).

If researchers incorrectly use a linear regressor, the estimated spillover on the linear leave-out mean is positive and significant (column 2). The ratio of estimated spillover to direct effect is 19%. This result falsely suggests that spillover effects played an important role in amplifying the effects of the shock. The ratio of estimated spillover to direct effect rises with the degree of nonlinearity. For instance, I redefine:

$$w_i = \begin{cases} x_i^2 & if\ x_i > 0, \\ 0 & otherwise. \end{cases}$$

The correctly specified regressors are still consistently estimated (column 3). However, using the linear regressors leads to an estimated ratio of 166% (column 4). This estimated ratio incorrectly implies that the spillover is quantitatively more important than the direct effect.

The reason for the overestimated spillover is that the specification with linear regressors fits the same coefficient for observations with $x_i > 0$ and for observations with $x_i \leq 0$. As a result, the

direct estimate is too low for observations with $x_i > 0$ (relative to the true effect). With systematic variation (i.e., $u_{r(i)}$ differs across groups), some of the true direct effect for observations with $x_i > 0$ (the part that is caused by high $u_{r(i)}$) loads on the coefficient for the leave-out mean and generates bias. The bias gets worse with the degree of nonlinearity, as the wedge between true and estimated direct effect rises. With random variation (i.e., $u_{r(i)}$ identical across groups), using linear regressors does not produce a biased spillover estimate because no common component in direct exposure could load onto the leave-out mean (columns 5 and 6).

5.5 Overcoming bias due to mismeasurement and nonlinear effects

Testing for heterogeneous spillovers, based on theory, is a useful tool. If spillovers are only significant for a subset of firms, for which theory predicts they should be, generic bias due to mismeasurement across all firm types cannot explain the spillover results.

A solution to nonlinearity is to relax the linearity assumption. For instance, plotting direct effects by bins of x_i should reveal which parts of the distribution of x_i are treated. As with multiple spillovers, instrumenting also overcomes the bias from mismeasurement. Using z_i as individual-level instrument (Equation (7)), the IV estimates are consistent if there is measurement error (Table 3, panel A, column 5) or if direct effects are nonlinear (Table 4, column 7).

6 Application: Estimating Spillovers following a Banking Shock

In this section, I illustrate that mechanical biases (due to multiple spillover types and mismeasurement) can be large in practice, by studying spillovers after a bank credit shock. I then use the estimated spillovers to inform a policy calculation.

6.1 Empirical setting and data

I analyze a lending cut by Commerzbank, the second-largest German bank in 2008. Commerzbank primarily lent to German firms and households. It suffered severe losses on its international financial investments during the financial crisis 2008, having held positions in U.S. mortgage markets

and failing Icelandic banks. Importantly, the losses were not caused by Commerzbank's lending to the German economy. German firms borrowing from Commerzbank were of comparable credit quality and on similar growth paths compared to firms borrowing from other banks.

Nonetheless, Commerzbank's crisis affected its German borrowers. As Commerzbank became financially constrained in 2008, it cut lending to German firms. Finding another lender is difficult for firms, especially in a time of crisis, as documented by a large literature on relationship banking (Sharpe 1990; Boot 2000). As a result, firms borrowing from Commerzbank faced a reduction in their credit supply and grew more slowly after the lending cut. In contrast, aggregate lending by other German banks actually increased slightly during the crisis.

Recent papers analyze the effect of Commerzbank's lending cut on firms (Huber 2018; Berg, Reisinger, and Streitz 2021; Biermann and Huber 2021). The evidence suggests that Commerzbank's lending cut was exogenous to the German economy, so that firms, product markets, and regions with greater dependence on Commerzbank would have grown at the same rates as other firms, had the lending cut not happened. Firms with a Commerzbank relationship became financially constrained and grew employment more slowly after the lending cut, compared to firms borrowing from other banks. In addition, firms grew more slowly when a large share of other firms in the region had a Commerzbank relationship.

I construct a firm-level data set following Huber (2018). Direct treatment status x_i is a binary indicator for whether a firm had a relationship to Commerzbank in 2006, measured using a confidential record of German firms' relationship banks by the credit rating agency Creditreform.[13] The outcome is the symmetric growth rate of firm employment between 2008 and 2012, calculated using the database Dafne by Bureau van Dijk.

I calculate leave-out means to test for spillovers at the level of two groups: product markets and regions. The share of other firms with a Commerzbank relationship (leave-out mean) in the product

[13]Bank relationships are available for 112,344 firms. German firms and banks usually form long-lasting relationships, as only 1.7% of firms add a new bank per year (Dwenger, Fossen, and Simmler 2015). This system of relationship banking facilitates credit provision during good times, but makes it more difficult to access credit when the bank cuts lending. Employment growth and the full set of controls (age, export and import shares, and industry) is available for 45,252 firms, 26% of which had a relationship to Commerzbank in 2006.

market is $\overline{x_{s(i)}}$ and the share in the region is $\overline{x_{r(i)}}$. Regions are defined as administrative counties *(Kreise)* where firms are located. Product markets are defined as industry cells (at the level of two-digit industries in the German WZ classification) for tradable firms and industry-region cells for nontradable firms (since they sell locally).[14]

For the purpose of this paper, I take as starting point that firms with a relationship to Commerzbank experienced an exogenous shock after Commerzbank's lending cut. I therefore take as given the identification assumption, which is that direct treatment status as well as product market and region leave-out means are uncorrelated with other shocks hitting firms. Detailed arguments in favor of this assumption are presented in the above-cited papers.

6.2 Bias due to multiple potential spillovers

I begin with an analysis that an economist interested in product markets might conduct. Theory suggests that firms may benefit from increased market share when firms in the same product market are treated, but may also suffer from lower technological spillovers (Greenstone, Hornbeck, and Moretti 2010; Bloom, Schankerman, and Van Reenen 2013; Giroud et al. 2021). To test the net effect of these opposing channels, I regress firm employment growth between 2008 and 2012 on direct treatment status and the product market leave-out mean. The coefficients on both direct treatment and market leave-out mean are statistically significant, negative, and of equal magnitude (Table 5, column 1). This suggests that the spillover is as large as the direct effect in a market in which all firms are treated. Taken at face value, the finding supports theoretical models where reduced technological spillovers play an important role in amplifying crises.

Economic theory suggests that there may be other spillovers, however. At the regional level, the sign of the spillover is also theoretically ambiguous. Firms may suffer from reductions in local demand and agglomeration forces when firms in the same region are treated, but may benefit from lower local wages (Ellison, Glaeser, and Kerr 2010; Moretti 2010; Mian and Sufi 2014; Giroud and

[14]Following Mian and Sufi (2014), I classify an industry as tradable if it exports at least US$10,000 per worker, US$500 million in total, or if the industry's regional Herfindahl index is in the top quartile (using U.S. industry data). The Herfindahl criterion uses the fact that tradable industries are geographically concentrated because they do not need to produce where they sell.

Mueller 2017, 2019). Including the regional leave-out mean in the specification strongly changes the conclusions. The estimated market spillover shrinks toward zero and becomes statistically insignificant (column 2). The estimated regional spillover is large and significant, consistent with models that include strong local demand and agglomeration effects, but inconsistent with large spillovers through product markets.

These findings highlight that a specification testing only for the market spillover leads researchers to misinterpret spillover forces. Consistent with the earlier conceptual discussion, spillover estimates are misleading if a relevant spillover is not included in the specification. Unlike in the case of standard omitted variable bias, such bias can arise even if the different leave-out means are uncorrelated. This implies that researchers should include all potential spillover forces in their specification, even when they are orthogonal to the leave-out mean of interest.

However, this poses practical difficulties. Many group connections are not reported. For instance, the data used here do not include information on whether firms use common inputs. Directly treated firms may generate spillovers onto other firms that use common inputs. The regional spillover estimate may be biased because the specification does not consider spillovers among common input users.

To get around this difficulty, researchers can test for heterogeneous spillover effects based on theory. Regional models predict that nontradable producers and innovative firms with high R&D are strongly affected by local shocks, whereas other firms are not (Jaffe, Trajtenberg, and Henderson 1993; Henderson 2003). If regional spillovers are present in equal measure among all types of firms, it is likely that the estimates are mechanically biased. However, if regional spillovers are zero for firms in tradable and low-R&D sectors, as theory predicts, spillover estimates are not driven by a generic mechanical bias. Splitting the sample, I find that the regional spillover is significant and large for nontradable and high R&D sectors (Table 5, column 3), but it is small and insignificant for tradable and low-R&D sectors (column 4). This suggest that the regional spillover is not an artifact of mechanical bias. In general, identifying a placebo category of firms, where spillovers should be zero, is a useful way for researchers to ensure that spillovers are not

mechanically biased.

6.3 Bias due to measurement error

Next, I explore the impact of measurement error. Both spillover and direct effects are significant in a specification without measurement error. The ratio of spillover to direct effect is 4.6 (Table 6, column 1). Direct treatment status is a binary variable in this application, so I add measurement error by misclassifying a random subset of the sample: 5% of observations are misclassified with low measurement error; 10% with medium; and 30% with high. I calculate the regional leave-out mean based on the mismeasured direct treatment status, as researchers in practice would.

The estimated ratio of spillover to direct effect rises with the magnitude of measurement error, from 6.7 with low error to 28.4 with high error (columns 2-4). These findings show that the intuition derived from the simulations has practical relevance. Measurement error attenuates the direct effect, so part of the direct effect falsely loads onto the spillover coefficient. In fact, with high error, the direct coefficient becomes insignificant and close to zero (column 4). Researchers using mismeasured data would erroneously conclude that local general equilibrium forces account for essentially all of the impact of a shock on a region.

Researchers can explore heterogeneous effects based on theory to test whether mechanical bias drives the spillover estimate. As above, I split the sample by the degree to which firms should be affected by local spillovers. With high measurement error, I find that spillovers are large and significant for both types of firms (columns 5 and 6). This finding should raise concern among researchers testing for regional spillovers. It suggests that mechanical bias plagues the estimates and that results are not driven by the theoretical forces described in urban models. Finding an appropriate instrument is one potential avenue to solving the issue. In the case of classical measurement error, any other variable that measures the same treatment would be an option, even if this instrument is also measured with error. In the absence of an instrument and heterogeneity tests, researchers should interpret spillover estimates with caution.

6.4 Magnitude and policy implications of spillover estimates

The results suggest that an untreated firm in a median region (with 24% of other firms treated) grew by 2.7 percentage points less solely because of regional spillovers (Table 5, column 2). The spillover effect in the median region is of equal magnitude to the direct effect, which is also estimated at 2.7 percentage points. Spillovers thus played a first-order role in the regional impact of the lending cut.

I also calculate how spillovers affected the number of jobs. To convert the estimated percentage changes into job growth, I split the sample into four firm size bins (1–49, 50–249, 250–999, and over 1,000 employees) and record the average direct treatment and leave-out means in each size bin.[15] I then plug these average numbers into Equations (9) and (10), following the method described in Section 3. The calculation shows that the total regional effect reduced job growth by roughly 10 jobs at the average firm, whereas the direct effect reduced job growth by roughly 3 jobs at the average firm. Hence, 30% of the total effect on jobs was due to direct effects and 70% due to spillovers.

Spillover estimates can be useful in the analysis of government policy. I calculate the dollar multiplier as the total effect on jobs relative to the dollar spending required to offset the direct effects of the lending cut (which is equal to M in the notation used in Section 3). The average firm lost 0.47 million euro in bank debt due to the direct effects of the lending cut.[16] The total regional effect on the average firm was 10 jobs. This implies that a government policy that provides 0.47 million euro in debt to the average firm increases average job growth in the region by 10 jobs. Converting these figures suggests that 100,000 euro in debt increases average regional job growth

[15] Average direct treatment increases with firm size, as it is 0.21 in the first bin (average firm size of 15 employees), 0.37 in the second (104 employees), 0.47 in the third (482 employees), and 0.57 in the largest (7,950 employees). Much of this correlation is driven by the fact that large firms are mechanically more likely to have more banks. In contrast and unsurprisingly, the average leave-out mean was relatively constant, lying between 0.23 and 0.26 for all groups.

[16] To get this figure, note that table 4, column 3 of Huber (2018) shows that the bank debt effect on the average treated firm is 10.3 percentage points. Average bank debt is 3,798,000 euro in the first bin, 6,883,000 euro in the second, 27,915,000 euro in the third, and 228,838,000 euro in the largest. Multiplying average bank debt in each bin with the fraction treated in each bin and taking a weighted average over bins gives the average euro effect on bank debt.

by 2.1 jobs and US$100,000 by 1.4 jobs (at the average 2008 exchange rate). In contrast, the direct effect on its own would imply that US$100,000 in debt generates only 0.4 jobs.

Note that separate estimates of direct and spillover effects are key for this type of policy analysis. We need an estimate of the direct effect on bank debt to measure the funds required to offset the initial shock to treated firms. And then we need to know the sum of direct and spillover effects on employment to get the regional impact. If we only knew the total impact on bank debt, we could not evaluate a policy targeted at directly treated firms. And if we only knew the direct employment effect, we would significantly understate the benefit of the policy by ignoring the large spillovers.

It is instructive to compare these estimates to the literature on fiscal stimulus, which often analyzes the effect of fiscal spending on job-years. Commerzbank's lending cut hit firms hardest in 2009 and the estimates in this paper are growth rates from 2008 to 2012. Depending on when jobs were lost, this suggests that the effect on job growth needs to be scaled up by a factor of roughly two or three to match job-years, resulting in job-year estimates of roughly 2.8 to 4.2 per US$100,000. In comparison, the regional impact of the 2009 American Recovery and Reinvestment Act, averaged across studies, was 2.1 jobs per US$100,000 of stimulus (Chodorow-Reich 2019). However, in contrast to fiscal stimulus, the government would recoup most funds lent to Commerzbank's borrowers in subsequent years, because firm delinquencies typically remain below 20% even in severe recessions. This makes such a lending policy relatively efficient from a net present value perspective.

7 Further Examples of Spillover Estimation

I describe examples for how researchers can directly estimate spillovers and overcome mechanical bias, using shocks commonly studied in the literature.

7.1 Spillovers among firms

7.1.1 Relevant spillovers after financing shocks

Many papers investigate how shocks to the supply of financial capital affect firms. These shocks may be lending cuts by banks, which arise unexpectedly and for reasons exogenous to borrowers' growth, as in the application above. They also include quasi-random access to government grants (e.g., Kerr and Nanda 2015; Howell 2017) and changes in firms' collateral values (e.g., Gan 2007; Chaney, Sraer, and Thesmar 2012). Researchers studying such shocks often estimate the direct effect on individual firms; that is, how firm performance changes following a direct financing shock relative to similar firms whose financing did not change.

To shed light on the propagation of financing shocks and to understand effects at higher levels of aggregation, researchers can directly estimate spillovers. Researchers' aims and motivations determine which types of spillovers are to be estimated. Some researchers may want to inform practical considerations of policy makers and firm managers. For instance, regional policy makers may like to know how entire regional economies evolve after financing shocks, while antitrust authorities and firm managers may wonder how product market competitors are affected. The leave-out means to estimate such spillovers are relatively easy to construct because region and product sector are often directly observed.[17]

Some researchers may be interested in testing general equilibrium models, for example, predictions about the strength of knowledge spillovers or about Keynesian demand spillovers on product markets (i.e., employees of affected firms earn and consume more). To test knowledge spillovers, researchers need to form groups of firms that operate in the same technology space and may benefit from knowledge spillovers (e.g., see the method in Bloom, Schankerman, and Van Reenen 2013). To test demand effects, researchers need to observe where employees of directly affected firms purchase products and then construct leave-out means for this group of "demand-dependent" firms. Ideally, this approach requires observing where employees of different firms shop (as in An-

[17]Of course, researchers need to ensure that variation across regions and product markets is credibly exogenous. Shift-share instruments (as in Greenstone, Mas, and Nguyen 2020) or historical variation (as in Huber 2018) may be helpful avenues to test group-level exogeneity.

dersen et al. 2022). If such data are not available, researchers can identify a subgroup of demand-dependent firms by focusing on producers of nontradable goods operating in the vicinity of directly affected firms (Mian and Sufi 2014; Huber 2018; Giroud and Mueller 2019).

Of particular interest in the context of financing shocks are spillovers operating through the banking system. After a positive financing shock, directly affected firms may be more likely to repay existing loans, thereby strengthening the balance sheet of their banks. This, in turn, can increase lending by banks with directly affected borrowers. Researchers can estimate this spillover channel by constructing a leave-out mean at the firm level, namely, the share of directly affected borrowers at the firms' relationship banks. If data on banking relationships are not available, researchers can rely on the fact that banking relationships of small firms are mostly local. A proxy leave-out mean to capture local banking spillovers could be the weighted average treatment status of small firms in the area, where weights are the total bank debt of a firm (relative to other small firms in the region).

7.1.2 Potential mechanical biases after financing shocks.

Mechanical biases likely play a role when researchers estimate spillovers after financing shocks. First, there exist multiple spillover types, as the discussion above shows. Second, measurement error may be a problem. Banking relationships and collateral value are often imperfectly measured, as it is unclear which preexisting lending relationships and which pieces of collateral actually influence firms' credit supply. Third, nonlinear direct effects may be an issue. Small changes in credit are easy to compensate, whereas large changes can have large effects (e.g., Huber 2018). Similarly, collateral may only matter once firms hit a binding constraint.

To investigate these issues, one may find heterogeneity tests inspired by theoretical models useful. To be concrete, consider the following heterogeneity tests for the spillover types discussed above. Theory predicts that regional spillovers are larger among nontradable and high-innovation firms; product market spillovers among firms with less market power; household demand spillovers among firms with larger dependence on consumer-facing sales (as opposed to firm-to-firm sales);

technology spillovers among firms in high-innovation sectors; and banking spillovers among firms with high leverage and dependence on external credit. To overcome potential bias from multiple spillover types, researchers should construct as many leave-out means (or proxies) as possible to test which spillover really drives the estimates.

7.1.3 Spillovers and potential biases after managerial shocks.

Of course, any firm-level shock that generates a direct effect may also generate spillovers. As additional example, consider shocks to firm managers, such as managerial turnover or compensation shocks (Jenter, Matveyev, and Roth 2018; Edmans, Gabaix, and Jenter 2017; Huber, Lindenthal, and Waldinger 2021). Apart from the spillover channels discussed so far, spillovers onto firms in the same labor market are more relevant after managerial shocks, whereas spillovers through lenders are likely less important than for financing shocks. Researchers can construct sector-by-country groups to proxy for managerial labor markets. Measurement error and nonlinearities are less of a problem in these cases (at least for large firms in which manager identities and compensation are public information), but multiple spillovers remain a concern.

7.2 Spillovers among households

7.2.1 Spillovers after borrowing and consumption shocks.

Several studies find that shocks to housing wealth affect household borrowing and consumption. Variation at the household level comes from idiosyncratic fluctuations in collateral values and house prices, either induced by price regulation or predetermined mortgage choices (e.g., Leth-Petersen 2010; DeFusco 2018; Cloyne et al. 2019). Similarly, financial education programs that facilitate household access to loans can raise borrowing and consumption (e.g., education about student loan applications, as in Mueller and Yannelis 2022).

Household borrowing may generate spillovers onto other households through the health of lenders. On the one hand, greater lending may raise lender profits, thereby strengthen lenders' balance sheets, and improve loan conditions to other household borrowers. On the other hand,

lenders may suffer more delinquencies when highly leveraged borrowers enter the market, which can worsen loan conditions. To estimate this potential spillover, researchers can construct a leave-out mean based on the share of borrowers at a household's set of potential lenders that are directly affected by the borrowing shock. If the set of potential lenders is not directly observed, researchers can construct proxies using lenders with local branches and with remote/online services provided in the region.

Household consumption generates spillovers onto customer-facing firms and their employees. This type of spillover is generated by the same economic mechanism as the firm-level spillover on "demand-dependent" firms described in Section 7.1. Researchers can construct similar proxies as in that section to identify other households that may benefit from an increase in consumption by directly affected households (e.g., employees of local nontradable firms). Relatedly, there may be spillovers in the local labor market, as wealth effects due to house prices reduce household labor supply, raising wages to the benefit of local workers. These labor market spillovers can be estimated using a leave-out mean for all employees employed in the same labor market.

7.2.2 Potential mechanical biases after borrowing and consumption shocks.

Access to financial education can in principle be well measured, but house prices and collateral values are difficult to observe accurately, making measurement error a challenge. There are also nonlinear effects in household responses to house prices, as increases in house prices raise borrowing, but decreases do not reduce borrowing (Cloyne et al. 2019). To detect whether measurement error and nonlinearities play a role, researchers can conduct heterogeneity tests tailored to the spillover of interest. Specifically, spillovers through lenders affect households with low liquidity and asset holdings more strongly; demand spillovers have larger effects on households working in consumer-facing sectors; and labor market spillovers matter more for working-age individuals. As usual, researchers can explore bias due to multiple spillover groups by measuring leave-out means for all spillover types and can overcome bias due to nonlinearities using flexible functional forms.

7.3 Spillovers among regions

7.3.1 Spillovers and potential biases after banking deregulation.

A large literature shows that U.S. state-level banking deregulation improved real economic outcomes (starting with Hubbard and Palia 1995 and Jayaratne and Strahan 1996, reviewed by Berger, Molyneux, and Wilson 2020). While researchers have established direct effects on deregulating states themselves, there exists less work on how deregulation in one state may have affected other states. A likely spillover channel is through trade. Researchers can construct groups of states that trade intensely with each other (e.g., using the Commodity Flow Survey) and construct leave-out means at the level of state groups. A further relevant spillover channel is through labor markets. Researchers can group states that experience significant cross-state labor flows to measure leave-out means.

Researchers should test whether states with larger dependence on cross-state trade and cross-state migration are more affected by these spillovers. If there is little heterogeneity, spillovers may be driven by omitted spillover types. Since the deregulation episodes are well documented and mostly take the form of binary treatment values, measurement error is less likely to be a problem in these settings.

7.3.2 Spillovers and potential biases after fiscal stimulus.

Another commonly studied regional shock is fiscal stimulus (for a review, see Chodorow-Reich 2019). Cross-regional spillovers through trade and labor markets likely play a role, as in the case of banking deregulation. These can be estimated by forming groups of states connected through trade and labor flows.

A unique feature of fiscal stimulus is that there may be exogenous variation at both the state and the country levels (as discussed in Ramey 2019). For instance, wars lead the U.S. government to increase defense spending, which has heterogeneous effects across states depending on the predetermined locations of military production. Whether there truly is exogenous variation in

aggregate fiscal stimulus is an active debate in the literature. Here, my aim is not to comment on this debate, but to point out that, in principle, researchers can use country-level variation in a given shock to directly estimate a spillover at the country level.[18]

In the case of fiscal stimulus, researchers would need to regress a state-level outcome on exogenous spending in the state itself as well as the leave-out mean of exogenous spending in all other states. The regression could control for state fixed effects (to absorb time-invariant variation across states), but not for time fixed effects (as country-level stimulus only varies across time). The coefficient for the leave-out mean would capture the country-level spillover effect of raising stimulus in other states. In fact, Conley et al. (2021) carries out a spillover estimation in this spirit, using state- and country-level variation in defense spending induced by geopolitical shocks.

More generally, whenever researchers can identify a source of variation that is exogenous at the country level as well as at a lower level of aggregation, they can estimate country-level spillovers. Country-level variation may stem from monetary policy abroad (Jiménez et al. 2012), large political upheavals (Fuchs-Schündeln 2008), or idiosyncratic policy decisions (Romer and Romer 2004, 2010).

8 Conclusion

Large-scale macroeconomic and financial shocks affect firms and households through many complex spillover channels. By directly estimating spillovers, researchers can test which general equilibrium effects need to feature in models and how empirical estimates from one level of aggregation can inform other levels of aggregation.

Direct estimation of spillovers requires careful implementation. Spillover estimates suffer from distinct sources of mechanical bias that are not sufficiently discussed in applied research. For example, spillover estimates can be of the wrong sign, large, and statistically significant if additional spillover types operate through channels outside of the empirical model. Measurement error and nonlinear direct effects can lead to large and significant spillover estimates even if the true model

[18] Macroeconomists sometimes refer to a country-level spillover as "missing intercept" (Wolf 2021).

contains zero spillovers.

Mechanical biases are particularly concerning for researchers studying large-scale financial and macroeconomic shocks because these settings feature many types of spillover channels, nonlinear effects are common, and measurement of shocks can be difficult. Using an application to a real-world credit cut, I highlight that mechanical bias can be large in a real-world setting. Researchers may form completely erroneous judgments about which spillover channels are important, for example, by concluding that there are large sectoral spillovers when, in fact, true sectoral spillovers are zero and regional spillovers are large.

Fortunately, several practical tools allow researchers to detect and overcome mechanical bias. Testing for heterogeneous effects, flexible functional forms, and instrumental variables can overcome the problems. The examples discussed in the final section of the paper provide practical guidance to researchers interested in directly estimating spillovers.

References

Acemoglu, D. 2010. Theory, general equilibrium, and political economy in development economics. *Journal of Economic Perspectives* 24:17–32.

Adão, R., C. Arkolakis, and F. Esposito. 2020. General equilibrium effects in space: Theory and measurement. Working Paper, University of Chicago.

Ammermueller, A., and J.-S. Pischke. 2009. Peer effects in european primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics* 27:315–48.

Andersen, A., E. T. Hansen, K. Huber, N. Johannesen, and L. Straub. 2022. Disaggregated economic accounts. Working Paper, University of Chicago.

Angelucci, M., and G. De Giorgi. 2009. Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *American Economic Review* 99:486–508.

Angrist, J. D. 2014. The perils of peer effects. *Labour Economics* 30:98–108.

Auclert, A., W. S. Dobbie, and P. Goldsmith-Pinkham. 2019. Macroeconomic effects of debt relief: Consumer bankruptcy protections in the great recession. Working Paper, Stanford University.

Auclert, A., M. Rognlie, and L. Straub. 2018. The intertemporal Keynesian cross. Working Paper, Stanford University.

Auerbach, A., Y. Gorodnichenko, and D. Murphy. 2020. Local fiscal multipliers and fiscal spillovers in the usa. *IMF Economic Review* 68:195–229.

Avitabile, C. 2012. Spillover effects in healthcare programs: Evidence on social norms and information sharing. Working Paper, IDB.

Baird, S., J. A. Bohren, C. McIntosh, and B. Özler. 2018. Optimal design of experiments in the presence of interference. *Review of Economics and Statistics* 100:844–60.

Barrot, J.-N., and J. Sauvagnat. 2016. Input specificity and the propagation of idiosyncratic shocks in production networks. *Quarterly Journal of Economics* 131:1543–92.

Bentolila, S., M. Jansen, G. Jiménez, and S. Ruano. 2018. When credit dries up: Job losses in the great recession. *Journal of the European Economic Association* 16:650–95.

Beraja, M., E. Hurst, and J. Ospina. 2019. The aggregate implications of regional business cycles. *Econometrica* 87:1789–833.

Berg, T. 2018. Got rejected? Real effects of not getting a loan. *Review of Financial Studies* 31:4912–57.

Berg, T., M. Reisinger, and D. Streitz. 2021. Spillover effects in empirical corporate finance. *Journal of Financial Economics* 142:1109–27.

Berger, A. N., P. Molyneux, and J. O. Wilson. 2020. Banks and the real economy: An assessment of the research. *Journal of Corporate Finance* 62:101513–.

Bernstein, S., E. Colonnelli, X. Giroud, and B. Iverson. 2019. Bankruptcy spillovers. *Journal of Financial Economics* 133:608–33.

Biermann, M., and K. Huber. 2021. Tracing the international transmission of a crisis through multinational firms. Working Paper, University of Chicago.

Bloom, N., M. Schankerman, and J. Van Reenen. 2013. Identifying technology spillovers and product market rivalry. *Econometrica* 81:1347–93.

Blundell, R., M. C. Dias, C. Meghir, and J. Van Reenen. 2004. Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association* 2:569–606.

Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar. 2019. Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake. *Review of Economics and Statistics* 101:60–75.

Boning, W. C., J. Guyton, R. Hodge, and J. Slemrod. 2020. Heard it through the grapevine: The direct and network effects of a tax enforcement field experiment on firms. *Journal of Public Economics* 190:104261–.

Boot, A. W. A. 2000. Relationship banking: What do we know? *Journal of Financial Intermediation* 9:7–25.

Bound, J., and A. B. Krueger. 1991. The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics* 9:1–24.

Bramoullé, Y., H. Djebbari, and B. Fortin. 2009. Identification of peer effects through social networks. *Journal of Econometrics* 150:41–55.

Breza, E., A. G. Chandrasekhar, T. H. McCormick, and M. Pan. 2020. Using aggregated relational data to feasibly identify network structure without network data. *American Economic Review* 110:2454–84.

Browning, M., L. P. Hansen, and J. J. Heckman. 1999. Micro data and general equilibrium models. *Handbook of Macroeconomics* 1:543–633.

Brunnermeier, M. K., and Y. Sannikov. 2014. A macroeconomic model with a financial sector. *American Economic Review* 104:379–421.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi. 2021. Supply chain disruptions: Evidence from the Great East Japan Earthquake. *Quarterly Journal of Economics* 136:1255–321.

Carvalho, V. M., and A. Tahbaz-Salehi. 2019. Production networks: A primer. *Annual Review of Economics* 11:635–63.

Chaney, T., D. Sraer, and D. Thesmar. 2012. The collateral channel: How real estate shocks affect corporate investment. *American Economic Review* 102:2381–409.

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.

———. 2019. Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy* 11:1–34.

———. 2020. Regional data in macroeconomics: Some advice for practitioners. *Journal of Economic Dynamics and Control* 115:103875–.

Chodorow-Reich, G., P. T. Nenov, and A. Simsek. 2021. Stock market wealth and the real economy: A local labor market approach. *American Economic Review* 111:1613–57.

Cloyne, J., K. Huber, E. Ilzetzki, and H. Kleven. 2019. The effect of house prices on household borrowing: A new approach. *American Economic Review* 109:2104–36.

Conley, T. G., B. Dupor, M. Ebsim, J. Li, and P. McCrory. 2021. The local-spillover decomposition of an aggregate causal effect. Working Paper, FRB St. Louis.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora. 2013. Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *Quarterly Journal of Economics* 128:531–80.

Cunha, J. M., G. De Giorgi, and S. Jayachandran. 2019. The price effects of cash versus in-kind transfers. *Review of Economic Studies* 86:240–81.

DeFusco, A. A. 2018. Homeowner borrowing and housing collateral: New evidence from expiring price controls. *Journal of Finance* 73:523–73.

Duflo, E., and E. Saez. 2003. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics* 118:815–42.

Dupor, B., and P. B. McCrory. 2018. A cup runneth over: Fiscal policy spillovers from the 2009 recovery act. *Economic Journal* 128:1476–508.

Dwenger, N., F. M. Fossen, and M. Simmler. 2015. From Financial to Real Economic Crisis: Evidence From Individual Firm-Bank Relationships in Germany Discussion Paper, DIW Berlin.

Edmans, A., X. Gabaix, and D. Jenter. 2017. Executive compensation: A survey of theory and evidence. In B. Hermalin and M. Weisbach, eds., *Handbook of the Economics of Corporate Governance*, vol. 1.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. W. Walker. forthcoming. General equilibrium effects of cash transfers: Experimental evidence from Kenya. *Econometrica* .

Ellison, G., E. L. Glaeser, and W. R. Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review* 100:1195–213.

Epple, D., and R. E. Romano. 2011. Peer effects in education: A survey of the theory and evidence. In *Handbook of social economics*, vol. 1, 1053–163. Amsterdam, The Netherlands: Elsevier.

Ferracci, M., G. Jolivet, and G. J. van den Berg. 2014. Evidence of treatment spillovers within markets. *Review of Economics and Statistics* 96:812–23.

Filmer, D., J. Friedman, E. Kandpal, and J. Onishi. 2021. Cash transfers, food prices, and nutrition impacts on ineligible children. *Review of Economics and Statistics* 1–45.

Fuchs-Schündeln, N. 2008. The response of household saving to the large shock of german reunification. *American Economic Review* 98:1798–828.

Gabaix, X., and R. S. Koijen. 2022. Granular instrumental variables .

Gan, J. 2007. Collateral, debt capacity, and corporate investment: Evidence from a natural experiment. *Journal of Financial Economics* 85:709–34.

Gathmann, C., I. Helm, and U. Schönberg. 2020. Spillover effects of mass layoffs. *Journal of the European Economic Association* 18:427–68.

Gautier, P., P. Muller, B. van der Klaauw, M. Rosholm, and M. Svarer. 2018. Estimating equilibrium effects of job search assistance. *Journal of Labor Economics* 36:1073–125.

Giroud, X., S. Lenzu, Q. Maingi, and H. Mueller. 2021. Propagation and amplification of local productivity spillovers. Working Paper, Columbia University.

Giroud, X., and H. M. Mueller. 2017. Firm leverage, consumer demand, and employment losses during the great recession. *Quarterly Journal of Economics* 132:271–316.

———. 2019. Firms' internal networks and local economic shocks. *American Economic Review* 109:3617–49.

Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman. 2003. The social multiplier. *Journal of the European Economic Association* 1:345–53.

Gormley, T. A., and D. A. Matsa. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies* 27:617–61.

Greenstone, M., R. Hornbeck, and E. Moretti. 2010. Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy* 118:536–98.

Greenstone, M., A. Mas, and H.-L. Nguyen. 2020. Do credit market shocks affect the real economy? Quasi-experimental evidence from the great recession and "normal" economic times. *American Economic Journal: Economic Policy* 12:200–25.

Guren, A., A. McKay, E. Nakamura, and J. Steinsson. 2020. What do we learn from cross-regional empirical estimates in macroeconomics? In *NBER Macroeconomics Annual 2020, volume 35*.

Helm, I. 2020. National industry trade shocks, local labour markets, and agglomeration spillovers. *Review of Economic Studies* 87:1399–431.

Henderson, J. V. 2003. Marshall's scale economies. *Journal of Urban Economics* 53:1–28.

Herreño, J. 2021. The aggregate effects of bank lending cuts. Working Paper, UC San Diego.

Hirano, K., and J. Hahn. 2010. Design of randomized experiments to measure social interaction effects. *Economics Letters* 106:51–3.

Howell, S. T. 2017. Financing innovation: Evidence from R&D grants. *American Economic Review* 107:1136–64.

Hubbard, R. G., and D. Palia. 1995. Executive pay and performance: Evidence from the US banking industry. *Journal of Financial Economics* 39:105–30.

Huber, K. 2018. Disentangling the effects of a banking crisis: Evidence from german firms and counties. *American Economic Review* 103:868–98.

Huber, K., V. Lindenthal, and F. Waldinger. 2021. Discrimination, managers and firm performance: Evidence from "Aryanizations" in Nazi Germany. *Journal of Political Economy* 129:2455–503.

Jaffe, A. B., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108:577–98.

Janssens, W. 2011. Externalities in program evaluation: The impact of a women's empowerment program on immunization. *Journal of the European Economic Association* 9:1082–113.

Jayaratne, J., and P. E. Strahan. 1996. The finance-growth nexus: Evidence from bank branch deregulation. *Quarterly Journal of Economics* 111:639–70.

Jenter, D., E. Matveyev, and L. Roth. 2018. Good and bad CEOs. Working Paper, LSE.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina. 2012. Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *American Economic Review* 102:2301–26.

Kerr, W. R., and R. Nanda. 2015. Financing innovation. *Annual Review of Financial Economics* 7:445–62.

Lalive, R., C. Landais, and J. Zweimüller. 2015. Market externalities of large unemployment insurance extension programs. *American Economic Review* 105:3564–96.

Lerner, J., and A. Seru. 2022. The use and misuse of patent data: Issues for finance and beyond. *Review of Financial Studies* 35:2667–704.

Leth-Petersen, S. 2010. Intertemporal consumption and credit constraints: Does total expenditure respond to an exogenous shock to credit? *American Economic Review* 100:1080–103.

Li, S., T. M. Whited, and Y. Wu. 2016. Collateral, taxes, and leverage. *Review of Financial Studies* 29:1453–500.

List, J. A., F. Momeni, and Y. Zenou. 2019. Are estimates of early education programs too pessimistic? Evidence from a large-scale field experiment that causally measures neighbor effects. Working Paper, University of Chicago.

Manresa, E. 2016. Estimating the structure of social interactions using panel data. Working Paper, NYU.

Manski, C. F. 1993. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60:531–42.

Mian, A., A. Sarto, and A. Sufi. 2022. Estimating credit multipliers. Working Paper, University of Chicago.

Mian, A., and A. Sufi. 2014. What explains the 2007-2009 drop in employment? *Econometrica* 82:2197–223.

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

Moffitt, R. A. 2001. Policy interventions, low-level equilibria, and social interactions. *Social Dynamics* 4:6–17.

Moretti, E. 2010. Local multipliers. *American Economic Review* 100:1–7.

Mueller, H., and C. Yannelis. 2022. Increasing enrollment in income-driven student loan repayment plans: Evidence from the navient field experiment. *Journal of Finance* 77:367–402.

Muralidharan, K., P. Niehaus, and S. Sukhtankar. 2017. General equilibrium effects of (improving) public employment programs: Experimental evidence from India. Working Paper, UC San Diego.

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

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Ramey, V. A. 2019. Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives* 33:89–114.

Rincke, J., and C. Traxler. 2011. Enforcement spillovers. *Review of Economics and Statistics* 93:1224–34.

Romer, C. D., and D. H. Romer. 2004. A new measure of monetary shocks: Derivation and implications. *American Economic Review* 94:1055–84.

———. 2010. The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review* 100:763–801.

Rubin, D. B. 1980. Comment: Randomization analysis of experimental data: The fisher randomization test. *Journal of the American Statistical Association* 75:591–3.

———. 1990. Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science* 5:472–80.

Sacerdote, B. 2011. Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the economics of education*, vol. 3, 249–77. Amsterdam, The Netherlands: Elsevier.

Sarto, A. 2018. Recovering macro elasticities from regional data. Working Paper, NYU.

Sharpe, S. A. 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance* 45:1069–87.

Tintelnot, F., A. K. Kikkawa, M. Mogstad, and E. Dhyne. 2021. Trade and domestic production networks. *Review of Economic Studies* 88:643–68.

Vazquez-Bare, G. forthcoming. Identification and estimation of spillover effects in randomized experiments. *Journal of Econometrics* .

Verner, E., and G. Gyöngyösi. 2020. Household debt revaluation and the real economy: Evidence from a foreign currency debt crisis. *American Economic Review* .

Wolf, C. 2021. The missing intercept: A demand equivalence approach. Working Paper, MIT.

Tables

Table 1: Testing for the wrong spillover biases estimates

	(1)	(2)	(3)	(4)
Coefficient for x_i	1.626***	0.999***	0.995***	0.998***
(true coefficient = 1)	(0.059)	(0.008)	(0.037)	(0.012)
Coefficient for $\overline{x_{s(i)}}$	-0.530***	0.001	-0.012	0.004
(true coefficient = 0)	(0.051)	(0.009)	(0.127)	(0.033)
Coefficient for $\overline{x_{r(i)}}$		1.000***		
(true coefficient = 1)		(0.009)		
Group-level variation		Systematic		Random
Estimator	OLS	OLS	IV	OLS

The variable x_i is the direct treatment status of firm i, which is in sector $s(i)$ and region $r(i)$; and $\overline{x_{s(i)}}$ and $\overline{x_{r(i)}}$ are the average treatment status of all other firms in $s(i)$ and $r(i)$, respectively, apart from firm i (leave-out means). The IV specification in column 3 instruments for x_i and $\overline{x_{s(i)}}$ using z_i and $\overline{z_{s(i)}}$. Systematic variation means that $u_{s(i)}$ and $u_{r(i)}$ (from Equation (7)) are lognormally distributed with a mean of zero and a standard deviation of one. Random variation indicates that $u_{s(i)}$ and $u_{r(i)}$ are zero for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table 2: Testing for just one type of spillover biases estimates

	(1)	(2)	(3)	(4)
Coefficient for x_i	1.626***	0.995***	0.999***	0.998***
(true coefficient = 1)	(0.059)	(0.037)	(0.008)	(0.012)
Coefficient for $\overline{x_{s(i)}}$	0.470***	0.988***	1.001***	1.004***
(true coefficient = 1)	(0.051)	(0.127)	(0.009)	(0.033)
Coefficient for $\overline{x_{r(i)}}$			1.000***	0.999***
(true coefficient = 1)			(0.009)	(0.009)
Group-level variation		Systematic		Random
Estimator	OLS	IV	OLS	OLS

The variable x_i is the direct treatment status of firm i, which is in sector $s(i)$ and region $r(i)$; and $\overline{x_{s(i)}}$ and $\overline{x_{r(i)}}$ are the average treatment status of all other firms in $s(i)$ and $r(i)$, respectively, apart from firm i (leave-out means). The IV specification in column 2 instrument for x_i and $\overline{x_{s(i)}}$ using z_i and $\overline{z_{s(i)}}$. Systematic variation means that $u_{s(i)}$ and $u_{r(i)}$ (from Equation (7)) are lognormally distributed with a mean of zero and a standard deviation of one. Random variation indicates that $u_{s(i)}$ and $u_{r(i)}$ are zero for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table 3: Mismeasurement due to classical error biases spillover estimates

A: Specifications with zero true spillover effect

	(1)	(2)	(3)	(4)	(5)
Coefficient for x_i^*	0.999***	0.863***	0.754***	0.469***	1.000***
(true coefficient = 1)	(0.009)	(0.010)	(0.010)	(0.009)	(0.029)
Coefficient for $\overline{x_{r(i)}}^*$	-0.000	0.129***	0.229***	0.474***	0.001
(true coefficient = 0)	(0.011)	(0.012)	(0.013)	(0.019)	(0.103)
Measurement error	None	Low	Medium	High	High
Estimator	OLS	OLS	OLS	OLS	IV

B: Specifications with true spillover effect

	(1)	(2)
Coefficient for x_i^*	0.521***	0.700***
(true coefficient = 1)	(0.009)	(0.011)
Coefficient for $\overline{x_{r(i)}}^*$	1.365***	0.693***
(true coefficient = 1)	(0.032)	(0.045)
Measurement error	High	High
Estimator	OLS	OLS
Group-level variation	Systematic	Random

The variable x_i is the direct treatment status of firm i, which is in sector $s(i)$ and region $r(i)$; and $\overline{x_{r(i)}}$ is the average treatment status of all other firms in $r(i)$, apart from firm i (leave-out means). An asterisk indicates that the variable is observed and may contain measurement error. The signal-to-total-variance ratio of x_i is 95% for low measurement error, 90% for medium measurement error, and 70% for high measurement error. The IV specification in panel A, column 5 instruments for x_i^* and $\overline{x_{r(i)}}^*$ using z_i and $\overline{z_{r(i)}}$. Systematic variation means that $u_{s(i)}$ and $u_{r(i)}$ (from Equation (7)) are lognormally distributed with a mean of zero and a standard deviation of one. Random variation indicates that $u_{s(i)}$ and $u_{r(i)}$ are zero for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table 4: Mismeasurement due to nonlinearity biases spillover estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coefficient for w_i	1.002***		1.000***		1.002***		
(true coefficient = 1)	(0.012)		(0.002)		(0.017)		
Coefficient for $\overline{w_{r(i)}}$	-0.002		-0.000		-0.004		
(true coefficient = 0)	(0.013)		(0.002)		(0.051)		
Coefficient for x_i		0.793***		4.594***		0.503***	0.773***
(true coefficient > 0)		(0.013)		(0.363)		(0.012)	(0.021)
Coefficient for $\overline{x_{r(i)}}$		0.148***		7.648***		-0.003	-0.019
(true coefficient = 0)		(0.016)		(2.025)		(0.032)	(0.102)
Regressor correctly specified	Yes	No	Yes	No	Yes	No	No
Definition of w_i	$w_i = x_i \; if \; x_i > 0$		$w_i = x_i^2 \; if \; x_i > 0$		$w_i = x_i \; if \; x_i > 0$		$w_i = x_i \; if \; x_i > 0$
Group-level variation	OLS	OLS	OLS	OLS	OLS	OLS	IV
Group-level variation	Systematic		Systematic		Random		Systematic

In columns 1, 2, 5, and 6, $w_i = x_i \; if \; x_i > 0$ and $w_i = 0 \; if \; x_i \leq 0$. In columns 3 and 4, $w_i = x_i^2 \; if \; x_i > 0$ and $w_i = 0 \; if \; x_i \leq 0$. The variable x_i is the direct treatment status of firm i in region $r(i)$ and sector $s(i)$; and $\overline{x_{(i)}}$ is the average treatment status over all other firms in region $r(i)$, apart from firm i (leave-out mean). The variable $\overline{w_{r(i)}}$ is the leave-out mean of w_i in region $r(i)$, where w_i is defined as described in the column. The IV specification in column 7 instruments for x_i and $\overline{x_{r(i)}}$ using z_i and $\overline{z_{r(i)}}$. Systematic variation means that $u_{r(i)}$ (from Equation (7)) is lognormally distributed with a mean of zero and a standard deviation of one. Random variation indicates that $u_{r(i)}$ is 0 for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table 5: Application: Testing for the wrong spillover

	(1)	(2)	(3)	(4)
Coefficient for x_i	-0.030***	-0.027***	-0.031**	-0.026***
	(0.007)	(0.007)	(0.013)	(0.009)
Coefficient for $\overline{x_{s(i)}}$	-0.030*	-0.015	-0.045	-0.007
	(0.018)	(0.018)	(0.031)	(0.024)
Coefficient for $\overline{x_{r(i)}}$		-0.114**	-0.213***	-0.067
		(0.051)	(0.077)	(0.055)
Sectors in sample	All sectors		Nontradable and high R&D	Tradable and low R&D
Observations	45,252	45,252	14,810	30,442

The variable x_i is the direct treatment status of firm i, which is in sector $s(i)$ and region $r(i)$; and $\overline{x_{s(i)}}$ and $\overline{x_{r(i)}}$ are the average treatment status of all other firms in $s(i)$ and $r(i)$, respectively, apart from firm i (leave-out means). All specifications control for firm log age, export share (fraction of exports out of total revenue), import share (fraction of imports out of total costs), and fixed effects for four firm size bins (1–49, 50–249, 250–999, and over 1,000 employees), industry fixed effects at the level of the one-digit WZ classification, and a fixed effect for firms in the former GDR. Standard errors are clustered by region.

Table 6: Application: Measurement error

	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient for x_i^*	-0.027***	-0.023***	-0.024***	-0.009	-0.021**	-0.004
	(0.007)	(0.006)	(0.006)	(0.006)	(0.010)	(0.007)
Coefficient for $\overline{x_{r(i)}}^*$	-0.123**	-0.155***	-0.160***	-0.256***	-0.346***	-0.214**
	(0.050)	(0.054)	(0.058)	(0.086)	(0.128)	(0.094)
Measurement error	None	Low	Medium	High	High	High
Sectors in sample	All sectors				Nontradable and high R&D	Tradable and low R&D
Observations	45,252	45,252	45,252	45,252	14,810	30,442

The variable x_i^* is the observed direct treatment status of firm i in region $r(i)$; and $\overline{x_{r(i)}}^*$ is the average observed treatment status of all other firms in $r(i)$, apart from firm i (leave-out means). The binary variable is misclassified for a random 5% of observations in the case of low, 10% in the case of medium, and 30% in the case of high measurement error. All specifications control for firm log age, export share (fraction of exports out of total revenue), import share (fraction of imports out of total costs) and fixed effects for four firm size bins (1–49, 50–249, 250–999, and over 1,000 employees), industry fixed effects at the level of the one-digit WZ classification, and a fixed effect for firms in the former GDR. Standard errors are clustered at the regional level.

Internet Appendix

Appendix A Estimating Spillover Effects Through Networks

The estimation issues studied in the paper are relevant for researchers using variation at the individual and group level to estimate spillovers. In this section, I show that similar issues apply to the estimation of spillover effects through networks. Network analysis requires slightly different notation, but the intuition is similar. I then explicitly show how measurement error and nonlinear direct effects can bias the estimates of network spillovers.

Appendix A.1 Setup of a Network

Researchers often study how networks amplify shocks. For instance, an active literature focuses on the transmission of firm-level shocks to other firms, through production or financial linkages (Barrot and Sauvagnat 2016; Boehm, Flaaen, and Pandalai-Nayar 2019; Carvalho and Tahbaz-Salehi 2019; Carvalho et al. 2021; Tintelnot et al. 2021). For the sake of concreteness, I describe the following analysis using the language of supply linkages in production networks, but the insights are more general.

A typical specification to analyze production networks is:

$$y_i = \theta\, x_i + \delta\, \overline{x_{(i)}} + \varepsilon_i, \tag{A1}$$

where y_i is a firm-level outcome and x_i is the direct treatment status of firm i. The average treatment status of firms that are direct suppliers to firm i is:

$$\overline{x_{(i)}} = \frac{\sum_{j \neq i} \left(x_j \cdot \mathbb{1}\left\{ j \textit{ supplies to } i \right\} \right)}{N_i}, \tag{A2}$$

where $\mathbb{1}\left\{ j \textit{ supplies to } i \right\}$ indicates whether firm j is a supplier to firm i. The number of suppliers to firm i is N_i. In the general network case and in all simulations below, links are directed, so that

a link from j to i (j supplies i) does not imply a link from i to j.[A1]

The key assumption is how direct treatment status is determined. I specify that:

$$x_i = r_i + \sum_{j \neq i} \left(r_j \cdot \mathbb{1}\left\{ j \; supplies \; to \; i \right\} \right) + u_i. \tag{A3}$$

The first term r_i is a random factor associated with firm i. The second term is the sum of all factors associated with the suppliers to firm i. The third term u_i is a random error. The variables r_i, ε_i,and u_i are uncorrelated, and each component is independently distributed across firms.

The second term implies that the treatment status of each firm is correlated with the treatment status of its suppliers. Such correlated treatment status occurs naturally if the creation of supply links is correlated with the process determining treatment status. For instance, if firms linked to the same supplier happen to be located in the same region (as in the case of sectoral clustering) and if treatment status is regionally concentrated (as in the case of natural disasters), then treatment status can be approximated by Equation (A3). Note that treatment status is still exogenous (i.e., uncorrelated with the error term ε_i in Equation (A1)). Correlated treatment status simply means that the process determining treatment status is not exogenous to supply links. In experimental settings, treatment status is less likely to be correlated with suppliers' treatment status because researchers can randomize treatment status independently of regional concentration or other types of clustering.

Appendix A.2 Effects of Measurement Error on Network Spillover Estimates

To highlight the consequences of measurement error in network analysis, I run 100 simulations. In each simulation, I generate a random network among 500 firms with density 0.002. This implies that firms have on average one supplier, with a standard deviation of one. I assume that r_i is lognormally distributed with a mean of zero and a standard deviation of one. The error terms ε_i and u_i are drawn from a normal distribution with a mean of zero and a standard deviation of 0.1.

[A1] The model in Equation (A1) can be generalized to include not just the treatment status of direct links, but also the treatment status of second order links (i.e., the treatment status of a supplier's supplier) and further higher order links (as in Carvalho et al. 2021). The intuition below also applies to such higher order analyses.

I generate data where the true direct effect is one ($\theta = 1$) but the network spillover effect is zero ($\delta = 0$). If treatment status is measured without error, a regression of the firm outcome on x_i and $\overline{x_{(i)}}$ produces consistent estimates (Table A.I, column 1). However, with measurement error, the network spillover effect is positive and significant (column 2).[A2] The ratio of network spillover to direct effect is 24%.

The intuitive reason for the bias in the network analysis is similar to above. There is a common factor in direct treatment status and supplier's treatment status. The common factor is relatively stronger, and measurement error is relatively weaker, in the measure of suppliers' treatment status. As a result, some of the true direct effect loads onto the spillover estimate.

Table A.I: Estimates of network spillovers are biased under measurement error and nonlinear direct effects

	(1)	(2)	(3)	(4)
Coefficient for x_i^*	1.000	0.656		0.411
(true coefficient = 1)	(0.001)	(0.042)		(0.050)
Coefficient for $\overline{x_{(i)}}^*$	0.000	0.158		0.077
(true coefficient = 0)	(0.002)	(0.036)		(0.021)
Coefficient for w_i			1.000	
(true coefficient = 1)			(0.003)	
Coefficient for $\overline{w_{(i)}}$			0.000	
(true coefficient = 0)			(0.004)	
Measurement error	No	Yes	No	No
True direct effects are nonlinear	No	No	Yes: $w_i = x_i^2 \ if \ x_i > 0$	

In columns 1 and 2, the true data-generating equation is $y_i = x_i + \varepsilon_i$. The variable x_i^* is the observed direct treatment status of firm i and $\overline{x_{(i)}}^*$ is the observed average treatment status over all suppliers of firm i. The variables are measured correctly in columns 1, 3, and 4. The variables are measured with error in column 2, so that the signal-to-total-variance ratio of x_i is 0.7. In columns 3 and 4, the true data-generating equation is $y_i = w_i + \varepsilon_i$, where $w_i = x_i^2 \ if \ x_i > 0$ and $w_i = 0 \ if \ x_i \leq 0$.The reported coefficients and standard errors are averaged over 100 simulations.

[A2]The specification of measurement error is the same as in Section 5.1 above. Direct treatment status x_i can only be measured with error, such that $x_i^* = x_i + \eta_i$. Measurement error η_i is drawn from a normal distribution with a mean of zero and a standard deviation of σ. It is uncorrelated with ε_i, r_i, and u_i. I set σ so that the signal-to-total variance ratio equals 0.7.

Appendix A.3 Effects of Nonlinear Direct Effects on Network Spillover Estimates

The network spillover estimate can also be biased if the true direct effect is nonlinear. To analyze the impact of nonlinearity, I define:

$$w_i = \begin{cases} x_i^2 & if\ x_i > 0, \\ 0 & otherwise. \end{cases}$$

Direct treatment status x_i is determined as in Equation (A3) above.[A3] I assume that the true direct effect of w_i is one ($\theta = 1$) and the network spillover effect is zero ($\delta = 0$), so that the true data-generating process is given by:

$$y_i = w_i + \varepsilon_i. \tag{A4}$$

If researchers specify the nonlinear relationship between x_i and y_i correctly, the regression produces consistent estimates (Table A.I, column 3). But if researchers use linear regressors, as is standard practice, the estimates are biased and the ratio of network spillover to direct effect is 19% (column 4).

The reason for the bias is, once again, the factor r_i that is common to the direct treatment status of firm i and suppliers' treatment status. The coefficient for x_i estimates a linear direct effect. Conditional on this linear effect, there remains a nonlinear correlation between suppliers' treatment status and the outcome y_i, induced by the factor r_i in suppliers' treatment status. This leads to a significant, large, and inconsistent estimate of the network spillover.

[A3] The random network and other random terms also follow the calibration above. The only difference is that the mean of the random error u_i is negative for the purpose of this section (equal to the negative of the 90th percentile of the distribution of $r_i + \sum_{j \neq i} (r_j \cdot \mathbb{1}\{j\ supplies\ to\ i\})$). If this mean was not negative, almost all observations would have positive x_i and there would not be a nonlinear direct effect of x_i and y_i.

Appendix B Converting Percentage Effects into Absolute Changes

Assume that a researcher uses the percentage change in employment as outcome:

$$\frac{\triangle n_i}{n_i} = \beta\, x_i + \gamma \overline{x_{r(i)}} + \lambda\, \overline{x_{s(i)}} + \alpha + \varepsilon_i, \tag{A5}$$

and wants to calculate the change in the absolute number of jobs driven by spillovers. The firm-level change in jobs is:

$$\triangle n_i = \frac{\triangle n_i}{n_i} \times n_i = \left(\beta\, x_i + \gamma \overline{x_{r(i)}} + \lambda\, \overline{x_{s(i)}} + \alpha + \varepsilon_i\right) \times n_i, \tag{A6}$$

where the second equality comes from inserting Equation (A5).

Appendix B.1 Conversion if Firm Size Is Symmetric

Under the assumption that all firms in a region are of roughly equal size before treatment, $n_i \approx \overline{n}$, it is easy to take region-level averages of Equation (A6) to arrive at an approximate equation for the change in jobs:

$$\overline{\triangle n}^{r(i)} \approx \left((\beta + \gamma)\, \overline{x}^{r(i)} + \lambda\, \overline{x_{s(i)}}^{r(i)} + \alpha + \overline{\varepsilon}^{r(i)}\right) \times \overline{n},$$

The total effect on jobs in the region is then:

$$Total\ Effect = \frac{d\overline{\triangle n_i}^{r(i)}}{d\overline{x}^{r(i)}} \approx (\beta + \gamma) \times \overline{n},$$

whereas the direct effect, in the absence of any regional spillover effects, is:

$$Direct\ Effect = \frac{d\overline{\triangle n_i}^{r(i)}}{d\overline{x}^{r(i)}} \mid (\gamma = 0) \approx \beta \times \overline{n}.$$

The equalities are approximate because they rely on the assumption that firms are roughly evenly sized. In the symmetric case, the conversion to jobs has no effect on ratios, since the scaling factor

$\overline{n}$ cancels out. The share of the number of jobs due to direct effects remains $\frac{\beta}{\beta+\gamma}$ and the share due to spillovers $\frac{\gamma}{\beta+\gamma}$.

Appendix B.2 Conversion if Firm Size Varies

If treatment varies with firm size, it is more accurate to calculate the effect on jobs separately for different parts of the firm size distribution. To do so, assume that all firms in a given size bin k are of roughly equal size before treatment, so that $n_i \approx \overline{n}^k$ for all firms i in size bin k. Then average Equation (A6) across all firms in region $r(i)$ and size bin k to get the bin-specific change in jobs:

$$\overline{\triangle n}^{r(i),k} \approx \left(\beta\, \overline{x}^{r(i),k} + \gamma \overline{x_{r(i)}}^{r(i),k} + \lambda\, \overline{x_{s(i)}}^{r(i),k} + \alpha + \overline{\varepsilon}^{r(i),k}\right) \times \overline{n}^k.$$

Unlike in the symmetric case, the average direct treatment status of firms and the average regional leave-out mean are not necessarily equal. Specifically, average treatment status of firms in size bin k is $\overline{x}^{r(i),k}$ and the average regional leave-out mean in size bin k (including in the averaging those firms in other size bins but in the same region) is $\overline{x_{r(i)}}^{r(i),k}$. The average change in the number of jobs across all firms in region $r(i)$ is then the weighted average of bin-specific changes in jobs:

$$\overline{\triangle n}^{r(i)} = \sum_k \left[\overline{\triangle n}^{r(i),k} \times \omega^{r(i),k}\right],$$

where $\omega^{r(i),k}$ is the fraction of firms in size bin k.

The total regional effect now depends on how treatment differs across different firm size bins. As a result, treatment is characterized by a vector $\overrightarrow{dx^{r(i)}}$. The vector contains average direct treatment status in each size bin (values of $d\overline{x}^{r(i),k}$ for each k) and average regional leave-out means in each size bin (values of $d\overline{x_{r(i)}}^{r(i),k}$ for each k).

The total regional effect of treatment is:

$$Total\ Effect = \frac{d\overline{\triangle n_i}^{r(i)}}{\overrightarrow{dx^{r(i)}}} \approx \sum_k \left[\left(\beta\, d\overline{x}^{r(i),k} + \gamma d\overline{x_{r(i)}}^{r(i),k}\right) \times \overline{n}^k \times \omega^{r(i),k}\right],$$

whereas the direct effect, in the absence of any regional spillover effects, is:

$$Direct\ Effect = \frac{d\overline{\triangle n_i}^{r(i)}}{\overrightarrow{dx^{r(i)}}} \mid (\gamma = 0) \approx \sum_k \left[\beta\, d\overline{x}^{r(i),k} \times \overline{n}^k \times \omega^{r(i),k}\right].$$

The ratios of direct and spillover effects now depend on the relative size of treatments affecting different size bins and cannot be easily simplified.

Since researchers are typically be interested in the effect in a representative region, the region-specific superscript $r(i)$ can be dropped and averages across the full sample can be used for the calculation:

$$Total\ Effect \approx \sum_k \left[\left(\beta\, d\overline{x}^k + \gamma d\overline{x_{r(i)}}^k\right) \times \overline{y}^k \times \omega^k\right],$$

$$Direct\ Effect \approx \sum_k \left[\beta\, d\overline{x}^k \times \overline{y}^k \times \omega^k\right].$$

Appendix B.3 Example of Conversion

For concreteness, consider a simple example. In each region, there are two size bins (L and M). A fraction w of firms are in bin L ($\omega^{r(i),L} = w$). Treatment is binary and only firms in the L bin are directly treated ($d\overline{x}^{r(i),L} = 1$ and $d\overline{x}^{r(i),M} = 0$). The leave-out mean for M firms is thus simply $\frac{N^L}{N-1}$, whereas the leave-out mean for L firms is $\frac{N^L-1}{N-1}$. The total effect is composed of the direct effect and spillovers affecting L firms plus the spillovers affecting M firms:

$$Total\ Effect \approx \left[\left(\beta + \gamma\left(\frac{N^L-1}{N-1}\right)\right) \times \overline{n}^L \times w\right] + \left[\gamma\frac{N^L}{N-1} \times \overline{n}^M \times (1-w)\right].$$

The direct effect affects only L firms:

$$Direct\ Effect \approx \beta \times \overline{n}^L \times w.$$

The setup has so far assumed homogeneous treatment effects across size bins, but the setup here generalizes to settings where researchers estimate heterogeneous treatment effects across size bins, i.e., where coefficients β^k and γ^k vary with k.

Appendix C Derivation of the Bias Due to Measurement Error

The true model is:

$$y_i = \beta x_i + \gamma \overline{x_{r(i)}} + \varepsilon_i. \tag{A7}$$

Direct treatment status x_i is measured with error. The observed variables are:

$$x_i^* = x_i + \eta_i = u_{r(i)} + z_i + v_i + \eta_i,$$

$$\overline{x_{r(i)}^*} = \overline{x_{r(i)}} + \overline{\eta_{r(i)}} = u_{r(i)} + \overline{z_{r(i)}} + \overline{v_{r(i)}} + \overline{\eta_{r(i)}}.$$

I assume that the variables ε_i, $u_{r(i)}$, z_i, and v_i are uncorrelated with each other.

The OLS estimator of γ is:

$$\begin{aligned}
\widehat{\gamma} &= \frac{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)(y_i - \overline{y_i}) \sum_i \left(x_i^* - \overline{x_i^*}\right)^2 - \sum_i \left(x_i^* - \overline{x_i^*}\right)(y_i - \overline{y_i}) \sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)}{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)^2 \left(x_i^* - \overline{x_i^*}\right)^2 - \left(\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)\right)^2} \\
&= \frac{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(\beta\left(x_i - \overline{x_i}\right) + \gamma\left(\overline{x_{r(i)}} - \overline{\overline{x_{r(i)}}}\right) + \left(\varepsilon_i - \overline{\varepsilon_i}\right)\right) \sum_i \left(x_i^* - \overline{x_i^*}\right)^2}{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)^2 \left(x_i^* - \overline{x_i^*}\right)^2 - \left(\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)\right)^2} \\
&\quad - \frac{\sum_i \left(x_i^* - \overline{x_i^*}\right)\left(\beta\left(x_i - \overline{x_{r(i)}}\right) + \gamma\left(\overline{x_{r(i)}} - \overline{\overline{x_{r(i)}}}\right) + \left(\varepsilon_i - \overline{\varepsilon_i}\right)\right) \sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)}{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)^2 \left(x_i^* - \overline{x_i^*}\right)^2 - \left(\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)\right)^2} \\
&= \frac{\sum_i \left(\beta\left(x_i - \overline{x_i}\right)\left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right) + \gamma\left(\overline{x_{r(i)}} - \overline{\overline{x_{r(i)}}}\right)\left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right) + \left(\varepsilon_i - \overline{\varepsilon_i}\right)\left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\right) \sum_i \left(x_i^* - \overline{x_i^*}\right)^2}{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)^2 \left(x_i^* - \overline{x_i^*}\right)^2 - \left(\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)\right)^2} \\
&\quad - \frac{\sum_i \left(\beta\left(x_i - \overline{x_{r(i)}}\right)\left(x_i^* - \overline{x_i^*}\right) + \gamma\left(\overline{x_{r(i)}} - \overline{\overline{x_{r(i)}}}\right)\left(x_i^* - \overline{x_i^*}\right) + \left(\varepsilon_i - \overline{\varepsilon_i}\right)\left(x_i^* - \overline{x_i^*}\right)\right) \sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)}{\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)^2 \left(x_i^* - \overline{x_i^*}\right)^2 - \left(\sum_i \left(\overline{x_{r(i)}^*} - \overline{\overline{x_{r(i)}^*}}\right)\left(x_i^* - \overline{x_i^*}\right)\right)^2},
\end{aligned} \tag{A8}$$

where the first equality is the definition of the OLS estimator with two regressors. The second equality comes from substituting the true Equation (A7) for y_i. The third equality comes from rearranging terms.

The probability limit of the OLS estimator is:

$$
\begin{aligned}
plim\,\widehat{\gamma} = & \frac{\left(\beta\, Cov\left(x_i, \overline{x^*_{r(i)}}\right) + \gamma Cov\left(\overline{x_{r(i)}}, \overline{x^*_{(i)}}\right) + Cov\left(\varepsilon_i, \overline{x^*_{r(i)}}\right)\right) V\left(x^*_i\right)}{V\left(\overline{x^*_{r(i)}}\right) V\left(x^*_i\right) - Cov\left(\overline{x^*_{r(i)}}, x^*_i\right)^2} \\
& - \frac{\left(\beta\, Cov\left(x_i, x^*_i\right) + \gamma Cov\left(\overline{x_{r(i)}}, x^*_i\right) + Cov\left(\varepsilon_i, x^*_i\right)\right) Cov\left(\overline{x^*_{r(i)}}, x^*_i\right)}{V\left(\overline{x^*_{r(i)}}\right) V\left(x^*_i\right) - Cov\left(\overline{x^*_{r(i)}}, x^*_i\right)^2} \\
= & \frac{\left(\beta\, V\left(u_{r(i)}\right) + \gamma\left(V\left(u_{r(i)}\right) + \frac{V(z_i)+V(u_i)}{\overline{N}-1}\right)\right)\left(V\left(u_{r(i)}\right) + V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)}{\left(V\left(u_{r(i)}\right) + V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)\left(V\left(u_{r(i)}\right) + \frac{V(z_i)+V(u_i)+V(\eta_i)}{\overline{N}-1}\right) - V\left(u_{r(i)}\right)^2} \\
& - \frac{\left(\gamma\left(V\left(u_{r(i)}\right) + V\left(z_i\right) + V\left(u_i\right)\right) + \gamma\left(V\left(u_{r(i)}\right)\right)\right) V\left(u_{r(i)}\right)}{\left(V\left(u_{r(i)}\right) + V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)\left(V\left(u_{r(i)}\right) + \frac{V(z_i)+V(u_i)+V(\eta_i)}{\overline{N}-1}\right) - V\left(u_{r(i)}\right)^2}
\end{aligned}
$$

where the first equality comes from substituting covariances and variances for the probability limits of the individual terms in Equation (A8). The second equality comes from solving for the covariances and variances. $\overline{N}$ is the average number of firms per group. Finally, rearranging gives:

$$
\begin{aligned}
plim\,\widehat{\gamma} = & \beta \frac{\left(\overline{N}-1\right) V\left(u_{r(i)}\right) V\left(\eta_i\right)}{\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)^2 + \overline{N} V\left(u_{r(i)}\right)\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)} \\
& + \gamma \frac{\left(\overline{N}-1\right) V\left(u_{r(i)}\right)\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)}{\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)^2 + \overline{N} V\left(u_{r(i)}\right)\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)} \\
& + \gamma \frac{\left(V\left(z_i\right) + V\left(u_i\right)\right)\left(V\left(u_{r(i)}\right) + V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)}{\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)^2 + \overline{N} V\left(u_{r(i)}\right)\left(V\left(z_i\right) + V\left(u_i\right) + V\left(\eta_i\right)\right)}.
\end{aligned}
$$