Estimating General Equilibrium Spillovers of Large-Scale Shocks

Kilian Huber*

March 2022

Abstract

Large-scale financial and macroeconomic shocks directly affect some firms and households and indirectly impact others through general equilibrium spillovers. In this paper, I describe how researchers can estimate spillovers directly 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 to a credit shock and additional examples highlight the broad relevance of the suggested methods.

*University of Chicago, kilianhuber@uchicago.edu. 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, Christian Wolf, and numerous seminar audiences for helpful conversations and comments.

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 micro data are informative about other levels of aggregation.[1] Spillovers are particularly important when researchers study large-scale financial and business cycle shocks because many firms and households are simultaneously affected and spillovers are often large.

Consider why it is helpful to understand spillovers using 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, we need 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 policymaker may wonder how subsidizing the unhealthy bank will affect the entire regional economy, and so will need to understand regional spillovers after a banking shock. In both cases, estimates of regional spillovers would allow converting 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 spillover estimation using quasi-experiments or experiments. This approach is

[1] In my definition, the terms "general equilibrium spillovers," "general equilibrium effects," and "spillovers" all refer to the same concept, namely, the effects of shocks on prices, technology, and other features of the economic environment. These effects operate not only at the country level, but also at lower levels of aggregation, such as regions, sectors, networks, etc. 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 spillover estimation. 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 investigate and overcome these biases. I illustrate the relevance of the proposed solutions with an application to a real-world credit shock and by describing further examples. Finally, I argue that direct spillover estimation can inform policy.

The Method of Direct Spillover Estimation

The paper begins with an empirical framework for direct spillover estimation. 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 spillover estimation 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 the region. Direct spillover estimation using leave-out means has several attractive features. It is

[2]Recent work attempts 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 VI).

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.

Notwithstanding these advantages, I argue that direct spillover estimation 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 the 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 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. To name a few relevant groups, a shock to firms can spill over to other firms through factor markets, product markets, input-output networks, and common lenders. 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, while 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 simultaneously 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 datasets (Bound and Krueger 1991). Such 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 non-spillover settings, 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 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 only exist when individuals face binding liquidity, borrowing, or capital constraints (Brunnermeier and Sannikov 2014; Giroud and Mueller 2017, 2019; Berg 2018; 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.

Investigating and Addressing Mechanical Bias

I turn to detecting and overcoming the sources of mechanical biases discussed so far (due to multiple spillovers, mismeasurement, and nonlinearities). To investigate 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 non-tradable 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 datasets, 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 hard 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 Direct 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: non-tradable 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 heterogeneity 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 direct spillover estimation. The results suggest that one job lost at a directly treated firm led to another 1.6 jobs lost at untreated firms in the same region. This finding implies that realistic general equilibrium models need to include strong regional amplification forces. It also offers lessons for policy. Consider a public credit program targeted at a median treated firm. A naive calculation based only on the direct effect suggests that this program would raise regional employment by 1.2 employees for 100,000 USD of lent funds. In contrast, direct spillover estimation implies much larger gains of 2.8 employees. The latter calculation requires knowledge of both direct and spillover effects and would not be possible based only on regional estimates.

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 address mechanical biases in each case. The examples highlight the broad relevance of direct 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 spillover estimation**

    (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. **Investigate 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. **Address 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, measurement error, and nonlinear direct effects can cause mechanical bias. Address each source of bias in turn.

   (b) To address 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 VI for which types of spillovers are likely relevant in different types of analyses.)

(c) To address 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). However, instruments may be hard to find in practice.

I Related Literature

This paper relates to the methodological discussion in finance and macroeconomics on how to convert estimates from micro data to higher or lower levels of aggregation. Most of this literature relies on structural, model-based approaches (Browning et al. 1999; Acemoglu 2010; Nakamura and Steinsson 2018).[3] Direct spillover estimation 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; Auerbach et al. 2020; Gathmann et al. 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

[3]Recent structural examples include: Li et al. (2016); Auclert et al. (2018, 2019); Beraja et al. (2019); Guren et al. (2020); Chodorow-Reich et al. (2021); Herreño (2021). Sarto (2018), Adão et al. (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, 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 in the other parts of finance and macroeconomics. See Berg et al. (2021) for more discussion on publications using differences-in-differences and spillover estimation.

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., due to liquidity constraints or regulatory capital thresholds) and treatment is often hard 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 forthcoming).

Two recent papers also provide methodological advice on estimating treatment effects in the presence of spillovers. Berg et al. (2021) show that spillovers arise naturally in standard corporate finance theory and analytically derive the bias of direct effects in models with a single spillover. I complement their insights by focusing on the estimation of spillover effects (rather than direct effects) and by studying mechanical bias arising from multiple spillovers, mismeasurement, and nonlinearities. Berg et al. (2021) estimate regional spillovers following the Commerzbank credit shock, while I consider regional and sectoral spillovers following the same shock. In another related paper, Mian et al. (2022) show theoretically how regional spillovers arise in a general equilibrium model and discuss identification assumptions in the context of regional credit shocks.[5]

Outside of finance and macroeconomics, several applied papers estimate spillovers directly, mainly in education (reviews in Epple and Romano 2011; Sacerdote 2011; List et al. 2019), development (RCTs in Miguel and Kremer 2004; Angelucci and De Giorgi 2009; Janssens 2011;

[5]The estimates in Huber (2018) and Mian et al. (2022) lead to remarkably similar conclusions on the magnitude of regional spillover effects, despite the different settings. Huber (2018, Table 11) reports that local general equilibrium effects account for roughly 60 percent of the total regional effect, while the corresponding number in Mian et al. (2022) is 80 percent. The methods differs slightly, as Mian et al. (2022) compare individual-level and region-level regressions, while I directly estimate spillovers (and associated standard errors) using a leave-out mean.

Muralidharan et al. 2017; Cunha et al. 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 et al. 2014; Lalive et al. 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.[6] 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 2001; Glaeser et al. 2003; Bramoullé et al. 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).

II Empirical Framework

II.A 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.

[6] 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 operate chiefly through the pre-defined 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.

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 object of study, but the analysis applies equally when households or other entities are directly treated.

While in reality there are many channels that 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 where researchers estimate aggregate spillovers among firms in the same country (see Section VI).

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.

Throughout the paper, I assume that treatment of all firms is exogenous to the error, such that $E(x_i \varepsilon_i) = 0 \; \forall i$. This implies that bias does not arise due to standard "endogeneity" issues of treatment being correlated with other determinants of the outcome (as in Manski's (1993) reflection problem). Instead, bias will arise mechanically due to the issues involved in constructing and estimating spillover coefficients.

The superscripts on the coefficients γ^j and λ^k indicate that spillover effects are firm-specific, since spillovers arising from two different firms are not necessarily identical. It is, however, difficult to estimate one spillover coefficient per firm in the data. Instead, researchers commonly assume that spillovers are identical for firms in the same sector or region, which facilitates estimation:

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

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

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 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. They measure the change in the outcome of firm i if the average exposure of other firms in its region or sector increases. The direct effect (i.e., the change if firm i alone is treated) is given by β. A useful way to report magnitudes are the ratios of spillover to direct effects, $\frac{\gamma}{\beta}$ and $\frac{\lambda}{\beta}$.[7] Spillover coefficients on their own are often hard to interpret because they do not always capture absolute effects, but rather, they capture effects relative to a control firm for whom direct treatment and leave-out means are all zero.[8]

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 easily be 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:

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

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

[7] Alternatively, researchers may prefer to report the share of the total effect due to general equilibrium spillovers. For instance, if spillover leave out-means are uncorrelated $\left(Corr\left(\overline{x_{r(i)}}, \overline{x_{s(i)}}\right) = 0\right)$, the share of the regional total effect driven by spillovers is $\frac{\gamma}{\beta+\gamma}$. This fact can be seen by taking expectations of equation 4 over regions.

[8] 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, 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.

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

where the total number of workers in the region is $\widetilde{N_{r(i)}}$ and $\widetilde{x_{s(i)}}$ can be defined analogously.

II.B 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 (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.

II.C Treatment May Vary Systematically Across Regions and Sectors

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

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

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 ν_i is an unobserved random error. The variables $u_{r(i)}$, $u_{s(i)}$, z_i, and ν_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 where researchers intentionally treat some groups more than others. In most naturally occurring settings, variation is also systematic. For instance, exposure to the 2008/09 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 that there is 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 hard 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.

II.D 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 log-normally distributed with mean 0 and standard deviation 1. 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 mean 0 and standard deviation 1. Throughout the paper, I report coefficients and standard errors averaged over 100 simulations.

III 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 investigate bias.

III.A Testing the Wrong Type of Spillover

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

$$y_i = x_i + \overline{x_{r(i)}} + \varepsilon_i. \tag{8}$$

Treatment varies systematically across regions and sectors (i.e., $u_{r(i)}$ 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 only be interested in spillovers among borrowers of the same bank. 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 why researchers may overlook relevant spillover types. Some economic connections between firms are not recorded in standard datasets. 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.

III.B 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 8, the regional and sectoral leave-out means are uncorrelated by construction, so it is not obvious that the coefficient on the regional leave-out mean should be biased.

To illustrate the effects of ignoring relevant spillovers, I use the simulated data based on equation 8 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 percent. In contrast, the estimated sectoral spillover coefficient on $\overline{x_{s(i)}}$ is negative and significant (Table I, column 1). The ratio of estimated spillover to direct effect is large at -33 percent. 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 on x_i. The spillover estimate is then also biased because x_i and $\overline{x_{r(i)}}$ are positively correlated (due to the common factor $u_{r(i)}$).

III.C Addressing Bias Due to Testing the Wrong Type of Spillover

How can researchers investigate whether there is 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 non-tradable 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 non-zero 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. Including the regional leave-out mean addresses the bias (Table I, column 2).[9] Instrumental variables can also solve the bias, although instruments are hard 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 I, 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,

[9] 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.

as outlined in Section II.C.

III.D Bias Due to Multiple Non-Zero Spillovers

The bias is not limited to the case where a spillover with zero coefficient is included in the model. In an additional simulation, I assume that spillovers operate within regions and sectors with a coefficient of one. The true data-generating process is thus:

$$y_i = x_i + \overline{x_{r(i)}} + \overline{x_{s(i)}} + \varepsilon_i. \tag{9}$$

However, as above, researchers only include 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 II, column 1). While the true ratio is 100 percent, the estimated ratio is 29 percent. Instrumenting (column 2) and controlling for all relevant spillover types (column 3) overcome the bias.

IV 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.

IV.A 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:

$$y_i = \beta x_i + \gamma \overline{x_{r(i)}} + \varepsilon_i. \tag{10}$$

Imagine that direct treatment status x_i is measured with error. Observed treatment status is:

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

Measurement error η_i is normally distributed with mean 0 and standard deviation 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:

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

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

IV.B Bias Due to Measurement Error

Using simulated data, I illustrate how estimates of spillovers depend on classical error. I generate data based on equation 10, assuming that there is no spillover effect ($\gamma = 0$). In the absence of measurement error ($STV = 1$), the regression results are consistent. The estimated direct effect (coefficient on x_i^*) is close to one and significant, while the estimated spillover (coefficient on $\overline{x_{r(i)}}^*$) is small and insignificant (Table III, 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 5 percent (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 223 percent 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 0.

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 partially gets 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 that is caused by high $u_{r(i)}$) shows up in the spillover estimate.

IV.C The Direction of Bias Due to Measurement Error

The examples so far showed 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 10 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.[10]

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 non-zero 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 non-zero ($\beta = 0$ 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 III, panel B, column 1). Under random group-level variation ($u_{r(i)} = 0$), the spillover is attenuated (column 2).

IV.D 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

[10]The full derivation and definitions of C_1 and C_2 are in Appendix B.

on lending and real outcomes, relative to small losses (Brunnermeier and Sannikov 2014).

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 on w_i is close to one and significant, while the regional spillover coefficient on $\overline{w_{r(i)}}$ is small and insignificant (Table IV, 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 20 percent. 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 percent (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 on 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 there is no common component in direct exposure that could load onto the leave-out mean (columns 5 and 6).

IV.E Addressing 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 III, panel A, column 5) or if direct effects are nonlinear (Table IV, column 7).

V 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.

V.A 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 2008/09, having held positions in US 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 2008/09, 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 et al. 2021; Biermann and Huber 2021). To summarize, 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 dataset 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.[11] 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

[11] Bank relationships are available for 112,344 firms. German firms and banks usually form long-lasting relationships, as only 1.7 percent of firms add a new bank per year (Dwenger et al. 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.

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 non-tradable firms (since they sell locally).[12]

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.

V.B 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 et al. 2010; Bloom et al. 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 V, column 1). This suggests that the spillover is as large as the direct effect in a market where 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 et al. 2010; Moretti 2010; Mian and Sufi 2014; Giroud and

[12] Following Mian and Sufi (2014), I classify an industry as tradable if it exports at least 10,000 USD per worker, 500 USD 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 non-tradable producers and innovative firms with high R&D are strongly affected by local shocks, while other firms are not (Jaffe et al. 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 non-tradable and high R&D sectors (Table V, 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.

V.C 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 VI, 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 percent of observations are misclassified with low measurement error; 10 percent with medium; and 30 percent 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 percent 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.

V.D Magnitude of Regional Spillovers

The results suggest that an untreated firm in a median region (with 24 percent of other firms treated) grew by 2.7 percentage points less solely because of regional spillovers (Table V, 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.

The spillover estimate can be represented as a job-for-job effect: the employment change of untreated firms relative to the employment change of directly treated firms. Untreated firms in the median region experienced just the spillover effect, an employment decline of 2.7 percentage points. Treated firms experienced both spillover and direct effects, an employment decline of 5.4 percentage points. The elasticity of untreated employment with respect to treated employment is thus 0.5. Multiplying the elasticity by the ratio of untreated to treated employment in the median region (which is 3.2) yields a job-for-job spillover of 1.6. For each job lost at directly treated firms, 1.6 jobs are lost at untreated firms.

V.E Policy Calculation Based on Spillovers

Spillover estimates can be useful in the analysis of government policy. Consider a stabilization policy that allows Commerzbank to continue lending to a median borrower.[13] The direct effects of Commerzbank's lending cut approximately lowered the median borrower's number of employees by one and bank debt by 90,243 USD.[14] By providing 90,243 USD for Commerzbank to lend to a median firm, the government could undo the direct effect and increase employment of the treated

[13]The German government conducted a policy in this spirit. It aimed to stabilize credit supply to German firms by injecting capital and purchasing a 25 percent stake in Commerzbank through its government fund Soffin. Despite these measures, Commerzbank still reduced lending. Commerzbank was the only large nationwide lender in Germany to be subsidized by Soffin. Only three other, specialized banks received capital from Soffin (two real estate banks, Aareal Bank and Hypo Real Estate Group, and the Landesbank West LB/Portigon), which shows that Commerzbank was relatively strongly affected by the crisis.

[14]The median firm in the sample had 36 employees and held 880,424 USD in bank debt (598,928 Euro using the 2008 average exchange rate) in 2008. The direct effect on employment was roughly 2.7 percent (Table V, column 2) and on bank debt was 10.25 percent (Table 4, column 3 of Huber 2018). Multiplying the percentage impact by the initial value yields the absolute changes of 1 employee and 90,243 USD in bank debt.

firm by one. Regional employment would then increase by an additional 1.6 employees, using the earlier estimate of the job-for-job spillover. This implies that, by providing 100,000 USD, the government would increase regional employment by 2.8 employees.

The estimate of 2.8 employees is close to recent estimates of the effect of fiscal stimulus. For instance, the impact of the 2009 American Recovery and Reinvestment Act, averaged across studies, was 2.1 jobs per 100,000 USD 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 remain below 10 percent even in recessions. This makes the lending policy relatively effective from a net present value perspective.

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

VI Further Examples of Direct Spillover Estimation

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

VI.A Spillovers Among Firms

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 et al. 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 policymakers and firm managers. For instance, regional policymakers 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.[15]

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 et al. 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 Andersen et al. 2022). If such data are not available, researchers can identify a subgroup of demand-dependent firms by focusing on producers of non-tradable 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

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

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).

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 pre-existing 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, while 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, heterogeneity tests inspired by theoretical models are useful. To be concrete, consider the following heterogeneity tests for the spillover types discussed above. Theory predicts that regional spillovers are larger among non-tradable 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 address 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.

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 et al. 2018; Edmans et al. 2017). Apart from the spillover channels discussed so far, spillovers onto firms in the same labor market are more relevant after managerial shocks, while 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 where manager identities and compensation are public information), but multiple spillovers remain a concern.

VI.B Spillovers Among Households

Spillovers After Borrowing and Consumption Shocks

Several studies find that shocks to housing wealth raise 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 pre-determined 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 VI.A. 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 non-tradable 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.

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 investigate 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 address nonlinearities using flexible functional forms.

VI.C Spillovers Among Regions

Spillovers and Potential Biases After Banking Deregulation

A large literature shows that US state-level banking deregulation improved real economic outcomes (starting with Hubbard and Palia 1995 and Jayaratne and Strahan 1996, recently reviewed by Berger et al. 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.

Spillovers and Potential Biases After Fiscal Stimulus

Another commonly studied regional shock is fiscal stimulus (review in 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 US government to increase defense spending, which has heterogeneous effects across states depending on the pre-determined locations of military production. Whether there truly is exogenous variation in aggregate fiscal stimulus is an active debate in the literature. My aim here 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.[16]

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 on the leave-out mean would capture the country-level spillover effect of raising stimulus in other states. In fact, a recent paper by Conley et al. (2021) carries out a spillover estimation in this spirit, using state- and country-level variation in defense spending induced by geopolitical

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

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).

VII Conclusion

Large-scale macroeconomic and financial shocks affect firms and households through many complex spillover channels. By using direct spillover estimation, 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 spillover estimation 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 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 address the problems. The examples discussed in the final section of the paper provide practical guidance to researchers interested in implementing direct spillover estimation.

References

Acemoglu, Daron, "Theory, General Equilibrium, and Political Economy in Development Economics," *Journal of Economic Perspectives*, 2010, *24* (3), 17–32.

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito, "General Equilibrium Effects in Space: Theory and Measurement," 2020.

Ammermueller, Andreas and Jörn-Steffen Pischke, "Peer Effects in European Primary Schools: Evidence From the Progress in International Reading Literacy Study," *Journal of Labor Economics*, 2009, *27* (3), 315–348.

Andersen, Asger, Emil Toft Hansen, Kilian Huber, Niels Johannesen, and Ludwig Straub, "Disaggregated Economic Accounts: Measurement and Implications for Fiscal Policy," 2022.

Angelucci, Manuela and Giacomo De Giorgi, "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 2009, *99* (1), 486–508.

Angrist, Joshua D., "The Perils of Peer Effects," *Labour Economics*, 2014, *30*, 98–108.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub, "The Intertemporal Keynesian Cross," 2018.

__ , Will S. Dobbie, and Paul Goldsmith-Pinkham, "Macroeconomic Effects of Debt Relief: Consumer Bankruptcy Protections in the Great Recession," 2019.

Auerbach, Alan, Yuriy Gorodnichenko, and Daniel Murphy, "Local Fiscal Multipliers and Fiscal Spillovers in the USA," *IMF Economic Review*, 2020, *68* (1), 195–229.

Avitabile, Ciro, "Spillover Effects in Healthcare Programs: Evidence on Social Norms and Information Sharing," 2012. IDB Working Paper 380.

Baird, Sarah, J. Aislinn Bohren, Craig McIntosh, and Berk Özler, "Optimal Design of Experiments in the Presence of Interference," *Review of Economics and Statistics*, 2018, *100* (5), 844–860.

Barrot, Jean-Noël and Julien Sauvagnat, "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks," *Quarterly Journal of Economics*, 2016, *131* (3), 1543–1592.

Bentolila, Samuel, Marcel Jansen, Gabriel Jiménez, and Sonia Ruano, "When Credit Dries Up: Job Losses in the Great Recession," *Journal of the European Economic Association*, 2018, *16* (3), 650–695.

Beraja, Martin, Erik Hurst, and Juan Ospina, "The Aggregate Implications of Regional Business Cycles," *Econometrica*, 2019, *87* (6), 1789–1833.

Berg, Tobias, "Got Rejected? Real Effects of Not Getting a Loan," *Review of Financial Studies*, 2018, *31* (12), 4912–4957.

__ , Markus Reisinger, and Daniel Streitz, "Spillover Effects in Empirical Corporate Finance," *Journal of Financial Economics*, 2021, *142* (3), 1109–1127.

Berger, Allen N., Phil Molyneux, and John O.S. Wilson, "Banks and the Real Economy: An Assessment of the Research," *Journal of Corporate Finance*, 2020, p. 101513.

Bernstein, Shai, Emanuele Colonnelli, Xavier Giroud, and Benjamin Iverson, "Bankruptcy Spillovers," *Journal of Financial Economics*, 2019, *133* (3), 608–633.

Biermann, Marcus and Kilian Huber, "Tracing the International Transmission of a Crisis Through Multinational Firms," 2021.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and John Van Reenen, "Evaluating the Employment Impact of a Mandatory Job Search Program," *Journal of the European Economic Association*, 2004, *2* (4), 569–606.

Boehm, Christoph E., Aaron Flaaen, and Nitya Pandalai-Nayar, "Input linkages and the Transmission of Shocks: Firm-Level Evidence From the 2011 Tōhoku Earthquake," *Review of Economics and Statistics*, 2019, *101* (1), 60–75.

Boning, William C., John Guyton, Ronald Hodge, and Joel Slemrod, "Heard It Through the Grapevine: The Direct and Network Effects of a Tax Enforcement Field Experiment on Firms," *Journal of Public Economics*, 2020, *190*, 104261.

Boot, Arnoud W. A., "Relationship Banking: What Do We Know?," *Journal of Financial Intermediation*, 2000, *9* (1), 7–25.

Bound, John and Alan B. Krueger, "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," *Journal of Labor Economics*, 1991, *9* (1), 1–24.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of Peer Effects Through Social Networks," *Journal of Econometrics*, 2009, *150* (1), 41–55.

Breza, Emily, Arun G. Chandrasekhar, Tyler H McCormick, and Mengjie Pan, "Using Aggregated Relational Data to Feasibly Identify Network Structure Without Network Data," *American Economic Review*, 2020, *110* (8), 2454–84.

Browning, Martin, Lars Peter Hansen, and James J. Heckman, "Micro Data and General Equilibrium Models," *Handbook of Macroeconomics*, 1999, *1*, 543–633.

Brunnermeier, Markus K. and Yuliy Sannikov, "A Macroeconomic Model With a Financial Sector," *American Economic Review*, 2014, *104* (2), 379–421.

Carvalho, Vasco M. and Alireza Tahbaz-Salehi, "Production Networks: A Primer," *Annual Review of Economics*, 2019, *11*, 635–663.

__ , Makoto Nirei, Yukiko Saito, and Alireza Tahbaz-Salehi, "Supply Chain Disruptions: Evidence From the Great East Japan Earthquake," 2020.

Chaney, Thomas, David Sraer, and David Thesmar, "The Collateral Channel: How Real Estate Shocks Affect Corporate Investment," *American Economic Review*, 2012, *102* (6), 2381–2409.

Chodorow-Reich, Gabriel, "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence From the 2008-9 Financial Crisis," *Quarterly Journal of Economics*, 2014, *129* (1), 1–59.

__ , "Geographic Cross-sectional Fiscal Spending Multipliers: What Have We Learned?," *American Economic Journal: Economic Policy*, 2019, *11* (2), 1–34.

__ , "Regional Data in Macroeconomics: Some Advice for Practitioners," *Journal of Economic Dynamics and Control*, 2020, p. 103875.

__ , Plamen T. Nenov, and Alp Simsek, "Stock Market Wealth and the Real Economy: A Local Labor Market Approach," *American Economic Review*, 2021, *111* (5), 1613–57.

Cloyne, James, Kilian Huber, Ethan Ilzetzki, and Henrik Kleven, "The Effect of House Prices on Household Borrowing: A New Approach," *American Economic Review*, 2019, *109* (6), 2104–36.

Conley, Timothy Guy, Bill Dupor, Mahdi Ebsim, Jingchao Li, and Peter McCrory, "The Local-Spillover Decomposition of an Aggregate Causal Effect," 2021. FRB St. Louis Working Paper 2021-006.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do Labor Market Policies Have Displacement Effects? Evidence From a Clustered Randomized Experiment," *Quarterly Journal of Economics*, 2013, *128* (2), 531–580.

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran, "The Price Effects of cash versus in-kind transfers," *Review of Economic Studies*, 2019, *86* (1), 240–281.

DeFusco, Anthony A., "Homeowner Borrowing and Housing Collateral: New Evidence From Expiring Price Controls," *Journal of Finance*, 2018, *73* (2), 523–573.

Duflo, Esther and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment," *Quarterly Journal of Economics*, 2003, *118* (3), 815–842.

Dupor, Bill and Peter B. McCrory, "A Cup Runneth Over: Fiscal Policy Spillovers From the 2009 Recovery Act," *Economic Journal*, 2018, *128* (611), 1476–1508.

Dwenger, Nadja, Frank M. Fossen, and Martin Simmler, "From Financial to Real Economic Crisis: Evidence From Individual Firm-Bank Relationships in Germany," 2015. DIW Berlin Discussion Paper 1510.

Edmans, Alex, Xavier Gabaix, and Dirk Jenter, "Executive Compensation: A Survey of Theory and Evidence," in Benjamin Hermalin and Michael Weisbach, eds., *Handbook of the Economics of Corporate Governance*, Vol. 1 2017.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W. Walker, "General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya," *Econometrica*, forthcoming.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr, "What Causes Industry Agglomeration? Evidence From Coagglomeration Patterns," *American Economic Review*, 2010, *100* (3), 1195–1213.

Epple, Dennis and Richard E. Romano, "Peer Effects in Education: A Survey of the Theory and Evidence," in "Handbook of Social Economics," Vol. 1, Elsevier, 2011, pp. 1053–1163.

Ferracci, Marc, Grégory Jolivet, and Gerard J. van den Berg, "Evidence of Treatment Spillovers Within Markets," *Review of Economics and Statistics*, 2014, *96* (5), 812–823.

Filmer, Deon, Jed Friedman, Eeshani Kandpal, and Junko Onishi, "Cash Transfers, Food Prices, and Nutrition Impacts on Ineligible Children," *Review of Economics and Statistics*, 2021, pp. 1–45.

Fuchs-Schündeln, Nicola, "The Response of Household Saving to the Large Shock of German Reunification," *American Economic Review*, 2008, *98* (5), 1798–1828.

Gan, Jie, "Collateral, Debt Capacity, and Corporate Investment: Evidence From a Natural Experiment," *Journal of Financial Economics*, 2007, *85* (3), 709–734.

Gathmann, Christina, Ines Helm, and Uta Schönberg, "Spillover Effects of Mass Layoffs," *Journal of the European Economic Association*, 2020, *18* (1), 427–468.

Gautier, Pieter, Paul Muller, Bas van der Klaauw, Michael Rosholm, and Michael Svarer, "Estimating Equilibrium Effects of Job Search Assistance," *Journal of Labor Economics*, 2018, *36* (4), 1073–1125.

Giroud, Xavier and Holger M. Mueller, "Firm Leverage, Consumer Demand, and Employment Losses During the Great Recession," *Quarterly Journal of Economics*, 2017, *132* (1), 271–316.

__ **and** __ , "Firms' Internal Networks and Local Economic Shocks," *American Economic Review*, 2019, *109* (10), 3617–49.

__ **, Simone Lenzu, Quinn Maingi, and Holger Mueller**, "Propagation and Amplification of Local Productivity Spillovers," 2021.

Glaeser, Edward L., Bruce I Sacerdote, and Jose A Scheinkman, "The Social Multiplier," *Journal of the European Economic Association*, 2003, *1* (2-3), 345–353.

Gormley, Todd A. and David A. Matsa, "Common Errors: How To (And Nt To) Control for Unobserved Heterogeneity," *Review of Financial Studies*, 2014, *27* (2), 617–661.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen, "Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence From the Great Recession and "Normal" Economic Times," *American Economic Journal: Economic Policy*, 2020, *12* (1), 200–225.

__ **, Richard Hornbeck, and Enrico Moretti**, "Identifying Agglomeration Spillovers: Evidence From Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Guren, Adam, Alisdair McKay, Emi Nakamura, and Jón Steinsson, "What Do We Learn From Cross-Regional Empirical Estimates in Macroeconomics?," in "NBER Macroeconomics Annual 2020, volume 35" 2020.

Helm, Ines, "National Industry Trade Shocks, Local Labour Markets, and Agglomeration Spillovers," *Review of Economic Studies*, 2020, *87* (3), 1399–1431.

Henderson, J. Vernon, "Marshall's Scale Economies," *Journal of Urban Economics*, 2003, *53* (1), 1–28.

Herreño, Juan, "The Aggregate Effects of Bank Lending Cuts," 2021.

Hirano, Keisuke and Jinyong Hahn, "Design of Randomized Experiments to Measure Social Interaction Effects," *Economics Letters*, 2010, *106* (1), 51–53.

Howell, Sabrina T., "Financing Innovation: Evidence From R&D Grants," *American Economic Review*, 2017, *107* (4), 1136–64.

Hubbard, R. Glenn and Darius Palia, "Executive Pay and Performance: Evidence From the US Banking Industry," *Journal of Financial Economics*, 1995, *39* (1), 105–130.

Huber, Kilian, "Disentangling the Effects of a Banking Crisis: Evidence From German Firms and Counties," *American Economic Review*, 2018, *103* (3), 868–898.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 1993, *108* (3), 577–598.

Janssens, Wendy, "Externalities in Program Evaluation: the Impact of a Women's Empowerment Program on Immunization," *Journal of the European Economic Association*, 2011, *9* (6), 1082–1113.

Jayaratne, Jith and Philip E. Strahan, "The Finance-Growth Nexus: Evidence From Bank Branch Deregulation," *Quarterly Journal of Economics*, 1996, *111* (3), 639–670.

Jenter, Dirk, Egor Matveyev, and Lukas Roth, "Good and Bad CEOs," 2018.

Jiménez, Gabriel, Steven Ongena, José-Luis Peydró, and Jesús Saurina, "Credit Supply and Monetary Policy: Identifying the Bank Balance-sheet Channel with Loan Applications," *American Economic Review*, 2012, *102* (5), 2301–26.

Kerr, William R. and Ramana Nanda, "Financing Innovation," *Annual Review of Financial Economics*, 2015, *7*, 445–462.

Lalive, Rafael, Camille Landais, and Josef Zweimüller, "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 2015, pp. 3564–3596.

Lerner, Josh and Amit Seru, "The Use and Misuse of Patent Data: Issues for Finance and Beyond," *Review of Financial Studies*, forthcoming.

Leth-Petersen, Søren, "Intertemporal Consumption and Credit Constraints: Does Total Expenditure Respond to an Exogenous Shock to Credit?," *American Economic Review*, 2010, *100* (3), 1080–1103.

Li, Shaojin, Toni M Whited, and Yufeng Wu, "Collateral, Taxes, and Leverage," *Review of Financial Studies*, 2016, *29* (6), 1453–1500.

List, John A., Fatemeh Momeni, and Yves Zenou, "Are Estimates of Early Education Programs Too Pessimistic? Evidence From a Large-Scale Field Experiment That Causally Measures Neighbor Effects," 2019.

Manresa, Elena, "Estimating the Structure of Social Interactions Using Panel Data," 2016.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Mian, Atif and Amir Sufi, "What Explains the 2007-2009 Drop in Employment?," *Econometrica*, 2014, *82* (6), 2197–2223.

__ , Andrés Sarto, and Amir Sufi, "Estimating Credit Multipliers," 2022.

Miguel, Edward and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 2004, *72* (1), 159–217.

Moffitt, Robert A., "Policy Interventions, Low-Level Equilibria, and Social Interactions," *Social Dynamics*, 2001, *4* (45-82), 6–17.

Moretti, Enrico, "Local Multipliers," *American Economic Review*, 2010, *100* (2), 1–7.

Mueller, Holger and Constantine Yannelis, "Increasing Enrollment in Income-Driven Student Loan Repayment Plans: Evidence From the Navient Field Experiment," *Journal of Finance*, 2022, *77* (1), 367–402.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar, "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence From India," 2017. NBER Working Paper 23838.

Nakamura, Emi and Jón Steinsson, "Identification in Macroeconomics," *Journal of Economic Perspectives*, 2018, *32* (3), 59–86.

Petersen, Mitchell A., "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, 2009, *22* (1), 435–480.

Ramey, Valerie A., "Ten Years After the Financial Crisis: What Have We Learned From the Renaissance in Fiscal Research?," *Journal of Economic Perspectives*, 2019, *33* (2), 89–114.

Rincke, Johannes and Christian Traxler, "Enforcement Spillovers," *Review of Economics and Statistics*, 2011, *93* (4), 1224–1234.

Romer, Christina D. and David H. Romer, "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 2004, *94* (4), 1055–1084.

__ **and** __ , "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 2010, *100* (3), 763–801.

Rubin, Donald B., "Comment: Randomization Analysis of Experimental Data: The Fisher Randomization Test," *Journal of the American Statistical Association*, 1980, *75* (371), 591–593.

__ , "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies," *Statistical Science*, 1990, *5* (4), 472–480.

Sacerdote, Bruce, "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, Elsevier, 2011, pp. 249–277.

Sarto, Andrés, "Recovering Macro Elasticities From Regional Data," 2018.

Sharpe, Steven A., "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships," *Journal of Finance*, 1990, *45* (4), 1069–87.

Tintelnot, Felix, Ayumu Ken Kikkawa, Magne Mogstad, and Emmanuel Dhyne, "Trade and Domestic Production Networks," *Review of Economic Studies*, 2020.

Vazquez-Bare, Gonzalo, "Identification and Estimation of Spillover Effects in Randomized Experiments," *Journal of Econometrics*, forthcoming.

Verner, Emil and Győző Gyöngyösi, "Household Debt Revaluation and the Real Economy: Evidence From a Foreign Currency Debt Crisis," *American Economic Review*, 2020.

Wolf, Christian, "The Missing Intercept: A Demand Equivalence Approach," 2021.

Tables

Table I: Testing for the wrong spillover biases estimates

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

Notes: 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 log-normally distributed with mean 0 and standard deviation 1. Random variation indicates that $u_{s(i)}$ and $u_{r(i)}$ are 0 for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

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

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

Notes: 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 log-normally distributed with mean 0 and standard deviation 1. Random variation indicates that $u_{s(i)}$ and $u_{r(i)}$ are 0 for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table III: Mismeasurement due to classical error biases spillover estimates

Panel A: Specifications with zero true spillover effect

	(1)	(2)	(3)	(4)	(5)
Coefficient on 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 on $\overline{x_{r(i)}}^*$	-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

Panel B: Specifications with true spillover effect

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

Notes: 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 percent for low measurement error, 90 percent for medium measurement error, and 70 percent 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 log-normally distributed with mean 0 and standard deviation 1. Random variation indicates that $u_{s(i)}$ and $u_{r(i)}$ are 0 for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table IV: Mismeasurement due to nonlinearity biases spillover estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coefficient on w_i	1.002***		1.000***		1.002***		
(true coefficient = 1)	(0.012)		(0.002)		(0.017)		
Coefficient on $\overline{w_{r(i)}}$	-0.002		-0.000		-0.004		
(true coefficient = 0)	(0.013)		(0.002)		(0.051)		
Coefficient on x_i		0.793***		4.594***		0.503***	0.773***
(true coefficient > 0)		(0.013)		(0.363)		(0.012)	(0.021)
Coefficient on $\overline{x_{r(i)}}$		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

Notes: 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 log-normally distributed with mean 0 and standard deviation 1. Random variation indicates that $u_{r(i)}$ is 0 for every firm. The reported coefficients and standard errors are averaged over 100 simulations.

Table V: Application: Testing for the wrong spillover

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

Notes: 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 VI: Application: Measurement error

	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient on 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 on $\overline{x_{r(i)}}^*$	-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			Non-tradable and high R&D	Tradable and low R&D
Observations	45,252	45,252	45,252	45,252	14,810	30,442

Notes: 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 percent of observations in the case of low, 10 percent in the case of medium, and 30 percent 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.

Online 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.A 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 et al. 2019; Carvalho and Tahbaz-Salehi 2019; Carvalho et al. 2020; Tintelnot et al. 2020). 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:

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

where $\mathbb{1}\left\{ j\, supplies\, to\, i \right\}$ 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 that there is also a link from i to j.[A1]

[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.

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.B 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 log-normally distributed with mean 0 and standard deviation 1. The error terms ε_i and u_i are drawn from a normal distribution with mean 0 and standard deviation 0.1.

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

2020). The intuition below also applies to such higher order analyses.

is positive and significant (column 2).[A2] The ratio of network spillover to direct effect is 24 percent.

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 on x_i^*	1.000	0.656		0.411
(true coefficient = 1)	(0.001)	(0.042)		(0.050)
Coefficient on $\overline{x_{(i)}}^*$	0.000	0.158		0.077
(true coefficient = 0)	(0.002)	(0.036)		(0.021)
Coefficient on w_i			1.000	
(true coefficient = 1)			(0.003)	
Coefficient on $\overline{w_{(i)}}$			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 $x_i > 0$	

Notes: 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.

Appendix A.C 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:

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

[A2] The specification of measurement error is the same as in Section IV.A 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 mean 0 and standard deviation 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.

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 percent (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 on 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.

Appendix B Derivation of the Bias Due to Measurement Error

The true model is:

$$y_i = \beta x_i + \gamma \overline{x_{r(i)}} + \varepsilon_i. \tag{A5}$$

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

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

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

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

[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.

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{A6}
$$

where the first equality is the definition of the OLS estimator with two regressors. The second equality comes from substituting the true equation A5 for y_i. The third equality comes from rearranging terms.

The probability limit of the OLS estimator is:

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

where the first equality comes from substituting covariances and variances for the probability limits of the individual terms in equation A6. The second equality comes from solving for the covariances and variances.

$\overline{N}$ is the average number of firms per group. Finally, rearranging gives:

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