



FDI spillover effects in incomplete datasets

Alex Eapen

Research School of Management, Australian
National University, Canberra, Australia

Correspondence:

A Eapen, Research School of Management,
Australian National University, Canberra,
ACT 0200, Australia.
Tel: +61 2 6125 7352;
email: alex.eapen@anu.edu.au

Abstract

Scholars studying foreign direct investment (FDI) spillovers usually examine whether productivity gains in domestic firms can be attributed to the presence of foreign firms in their industry. However, empirical estimation is often based on datasets that omit certain kinds of firms in the economy. We argue that identifying FDI spillover effects in such incomplete datasets is problematic, owing to measurement error and selection problems. Using Monte Carlo simulations, we show that spillover effect estimates from incomplete datasets are potentially biased. We discuss the theoretical implications of this, and demonstrate a weighted instrumental variable approach that could yield better spillover effect estimates in incomplete datasets.

Journal of International Business Studies (2013) 0, 1–26. doi:10.1057/jibs.2013.32

Keywords: knowledge and productivity spillovers; incomplete datasets; identification problems; Monte Carlo simulation; weighted instrumental variable estimator

INTRODUCTION

Policymakers often justify incentives such as subsidies and tax breaks to foreign investors with the argument that positive externalities from the presence of foreign firms in the economy far outweigh the economic cost of those incentives. The scholarly literature on foreign direct investment (FDI) spillovers probes this argument by examining whether there are indeed any technology externalities from foreign firms. Specifically, scholars in this field look for productivity gains in domestic firms that can be attributed to the presence of foreign firms in their industry (Altomonte & Pennings, 2009; Driffield & Love, 2007; Eapen, 2012; Gorg & Strobl, 2001; Haskel, Pereira, & Slaughter, 2007; Keller & Yeaple, 2009; Meyer & Sinani, 2009; Zhang, Li, Li, & Zhou, 2010). The basic spillover specification thus has productivity of domestic firms as dependent variable, and foreign presence, usually proxied by the share of foreign firms in total capital or employment in the industry, as key regressor. A positive coefficient estimate for foreign presence suggests the presence of technology external effects or, in other words, technology spillovers from foreign to domestic firms.

While scholars have since augmented this basic model in terms of its theoretical specification and empirical estimation, nearly all recent studies estimate spillover effects using datasets of firms that are incomplete. Manufacturing census datasets that several scholars have used (e.g., Haskel et al., 2007; Zhang et al., 2010) invariably miss data on small firms in the economy. On the other hand, secondary datasets such as Compustat and Prowess that others have

used (e.g., Keller & Yeaple, 2009; Marin & Sasidharan, 2010) are usually biased toward publicly listed firms and, as such, against foreign wholly owned subsidiaries, unlisted joint ventures between foreign and domestic firms, and at least some privately held domestic firms.

We argue that correctly identifying spillover effects¹ in incomplete datasets can be problematic. First, measures of foreign presence calculated from incomplete datasets are very likely to be inaccurate. For example, measuring foreign presence from datasets that systematically miss foreign wholly owned or joint ventures will very likely underestimate true foreign presence. Using datasets that miss domestic firms, on the other hand, can overestimate it. With such measurement error, scholars could potentially end up mapping changes in domestic firm productivity to incorrect variations in foreign presence. This, as we will show later in the paper, can result in substantially biased estimates of the spillover effect. Second, using incomplete datasets may raise selection problems. Specifically, datasets such as manufacturing census surveys that miss small firms essentially leave out a category of domestic firms that constitute the true spillover effect. And if, as prior literature has argued, small firms lack the absorptive capacity to benefit from spillovers (Altomonte & Pennings, 2009; Blalock & Simon, 2009; Zhang et al., 2010), this selective censoring can lead to overestimating the true spillover effect. Most importantly, in cases of both measurement and selection problems, we will usually lack sufficient information to distinguish biased empirical estimates from the true spillover effect. And in this sense, spillover effects are difficult to identify.

Yet this point has received very little attention in the literature.² To our knowledge, there has not yet been a formal analysis of how incompleteness of datasets affects spillover effect estimates. Nor has there been a careful consideration of how to estimate spillover effects correctly in incomplete datasets. Considering the rapidly increasing use of incomplete firm-level datasets in the literature on the one hand, and the grave theoretical and practical consequences of incorrectly identifying FDI spillovers on the other, this is a crucial gap.

We seek to bridge this gap, and in doing so make three contributions to the literature. First, we offer a formal examination of whether current spillover specifications can correctly identify the spillover effect in incomplete datasets. We present a clear exposition of the identification problem by breaking it down into its measurement error and selection

components, and show that spillover effects are correctly identified in incomplete datasets only under fairly restrictive assumptions. We also offer some predictions on the potential bias in estimation when these assumptions do not hold. Second, we use a series of Monte Carlo simulations to demonstrate what happens when researchers estimate spillover effects in incomplete datasets. These simulations capture what is most likely happening in most recent spillover estimations, and confirm that the bias in estimated spillover effects is not trivial. Our key result is that measurement error overestimates the spillover effect in secondary datasets that miss foreign firms. This is likely because, as we explain later in the paper, improvements in domestic firm productivity in such datasets end up being attributed to smaller increases in foreign presence than they should be. The effect of foreign presence, in turn, gets magnified. Our second key result is that missing small firms triggers selection and measurement error biases that together lead to overestimates of the true spillover effect. This finding carries significant implications for what we know about FDI spillovers. At the very least, it raises the possibility that prior results derived from manufacturing census datasets that miss small firms have probably overstated the actual extent of spillovers from foreign firms. Our third contribution to the literature lies in our discussion of whether we can do better to estimate spillover effects more accurately in incomplete datasets. We present the case for a weighted instrumental variable estimator that combines the use of instrumental variables with weights that are inversely proportional to each firm's sample inclusion probability, and empirically demonstrate that this approach provides better spillover effect estimates in incomplete datasets. Taken as a whole, thus, our paper stresses the need to redirect empirical research in FDI spillovers toward designs that better identify spillover effects. This is of grave importance, given that identification – the ability to generate valid inferences from empirical data – is a key benchmark for high-quality empirical research (Oxley, Rivkin, & Ryall, 2010).

The rest of the paper is organized as follows. We first review the basic FDI spillover specification and some recent extensions, and against this backdrop highlight the research problem that motivates this paper. Then, we show how identifying spillover effects in incomplete datasets is difficult, owing to measurement error and selection issues. We demonstrate this empirically through a series of Monte Carlo simulations. Subsequently, we propose a

weighted instrumental variable estimator as a potential solution to the identification problems we raise in the paper, and empirically show that it yields significantly better estimates of spillover effects in incomplete datasets. In the final section, we discuss the theoretical implications of this study.

BACKGROUND

Given that scholars in this field are interested primarily in technology spillovers from foreign to domestic firms, the basic spillover specification examines whether there are productivity differences between domestic firms that can be attributed to variations in the presence of foreign firms in their respective industries. This takes the form

$$Prod_{ij} = \beta_0 + \beta_1 FP_j + \varepsilon_{ij} \quad (1)$$

where $Prod_{ij}$ is total factor productivity of domestic firm i of industry j , FP_j is foreign presence in industry j , and β_1 , the coefficient of foreign presence, is the spillover effect. A positive estimate of β_1 indicates a positive correlation between foreign presence and domestic firm productivity, suggesting the possibility of spillovers from foreign to domestic firms. This is certainly an indirect test of spillovers (Driffield, Love, & Menghinello, 2010), but underpins the specifications commonly used in the literature.

A key issue in the above specification is that the presence of foreign firms confounds technology spillovers with a potentially negative competition effect. In other words, since not all foreign firms are potential sources of spillovers, and some exert more competitive pressures on domestic firms than they contribute to spillovers (Aitken & Harrison, 1999), the net effect of foreign presence may not identify spillovers, even when they do exist. Also, there is a potential endogeneity issue, because a positive coefficient estimate could also result from foreign firms simply choosing to enter productive domestic industries. To get around these, scholars have extended the basic spillover specification in different ways. While many recent studies seek to separate spillover and competition effects by including a control term for competition (e.g., Javorcik, 2004), Castellani and Zanfei (2006) go a step further by introducing a specification that gives greater weights in the foreign presence term to knowledge-intensive foreign subsidiaries. In this way, the foreign presence term better reflects the stock of foreign technology in the industry that could potentially spill over. Similarly, since not all domestic firms are equally capable of absorbing spillovers, disregarding this heterogeneity could conceal evidence for spillovers. Several

scholars have hence modeled spillovers to vary with the absorptive capacity of domestic firms (Blalock & Simon, 2009; Eapen, 2012; Zhang et al., 2010). In the same spirit, other scholars have conditioned the spillover effect on characteristics of the industry (Keller & Yeaple, 2009; Zamborsky, 2012) and institutional setting (Smeets & de Vaal, 2011). Another important strand of the literature argues that spillovers are better observed between foreign and domestic firms in vertically related industries (Blalock & Simon, 2009; Javorcik, 2004). On the empirical side, scholars routinely estimate these specifications using firm fixed effects or difference score estimators in order to rule out simultaneity and other endogeneity issues that could hinder identification of the spillover effect (see Keller, 2009).

Clearly, the common thread through all these extensions of the basic model is a focus on identifying the spillover effect correctly. As we explain later on, identification relies crucially on having theoretical models and empirical tests that reflect technology externalities as they exist in reality (see discussions in Altomonte & Pennings, 2009; Castellani & Zanfei, 2006, 2007; Driffield & Jindra, 2012; and also Oxley et al., 2010, for a more general treatment). And all else constant, improvements such as those described above should allow for a more precise identification of the spillover effect.

RESEARCH PROBLEM

One issue that has not received much attention this far, however, is the incomplete nature of datasets routinely used to estimate spillover effects. Practically all recent studies use datasets of firms sourced either from secondary data sources such as Compu-stat (Keller & Yeaple, 2009), Amadeus (Castellani & Zanfei, 2006), and Prowess (Marin & Sasidharan, 2010), or from firm-level census surveys of the manufacturing sector by government statistical agencies. As examples of the latter, Haskel et al. (2007) use data on UK firms from the Annual Census of Production Respondents Database collected by the UK Office of National Statistics, and Zhang et al. (2010) use data on Chinese firms from the Chinese National Bureau of Statistics (NBS). The critical feature of both secondary and manufacturing census datasets is that they do not contain all the firms in the economy. They are incomplete. The Compu-stat dataset, for example, “includes only publicly traded companies” (Keller & Yeaple, 2009: 823), and the Prowess database that some scholars have used to estimate spillover effects in India “compiles and provides detailed quantitative information from the

income statements and balance sheets of listed companies”³ (Marin & Sasidharan, 2010: 1231). This naturally implies that wholly owned foreign subsidiaries, unlisted joint ventures between foreign and domestic firms, and at least some privately held domestic firms could be missing in these datasets. Banga (2006) suggests precisely such a bias in her Indian dataset, and says it “does not include fully foreign-owned firms, nor any of the joint ventures that are not listed on any Indian stock exchange” (561).

Manufacturing census datasets are certainly more comprehensive in coverage, but they usually tend to miss small firms. As Chang and Xu (2008) explain, the Chinese manufacturing census dataset, which other scholars have also used, leaves out firms with annual sales less than 5 million Rmb. Haskel et al. (2007) and Altomonte and Pennings (2009) describe a similar size bias in the UK and Romanian datasets. And for the Amadeus-derived Community Innovation Survey (CIS)-Elios data that Castellani and Zanfei (2006) use, they say: “The CIS-Elios sample tends to over-represent large firms” (201).

It is thus reasonably fair to say that nearly all recent estimates of spillover effects have been from datasets that are incomplete or, more precisely, from datasets that either miss small or privately held foreign and domestic firms. Can spillover effects still be identified in incomplete datasets? Or, in other words, can we still correctly estimate spillover effects from datasets that do not contain all foreign or domestic firms? If not, what does this imply for theoretical conclusions we can draw from past studies? And equally importantly, given that even the most comprehensive datasets available to scholars today (e.g., manufacturing census datasets) still miss certain types of firms, is there a way to improve spillover effect estimation from incomplete datasets? These are the questions we seek to answer in the rest of this paper.

IDENTIFICATION OF SPILLOVER EFFECTS IN INCOMPLETE DATASETS

Identification, ultimately, is a causal inference problem. A common interpretation, for example, in the structural equation modeling domain, is that identification pertains to whether there is sufficient information in a model to allow it to converge to one unique set of values for the parameters. If multiple solutions exist, the model is not identified.

But fundamentally, since the parameters of interest in the model usually represent causal relationships, identifying them means being able to trace

out those causal relationships from the available data (Angrist & Krueger, 2001; Angrist & Pischke, 2009; Manski, 1995). In this sense, identification is an issue of causal inference. Endogeneity, for example, is an identification problem, because one cannot infer the direction of the true causal relationship from the available data alone. Selection problems, too, pose identification challenges, because we do not observe the distribution of the outcome for censored observations. And hence inferring anything about causal effects, that is, how the conditional distribution of the outcome variable (or more precisely, its conditional mean) varies with different sets of values of the regressors, is problematic (Manski, 1995).

We use identification in this sense of whether we are able to tease out our causal effect of interest correctly from the available data. And hence there are both theoretical and empirical precursors to identifying spillover effects. On the theoretical side, identification relies on having valid theoretical models of spillovers that reflect reasonably accurately how spillovers actually occur.⁴ And, on the empirical side, we need to carefully ensure that it is indeed these theoretical models and not any other mechanism that is generating the empirical data we observe. In other words, we need to ensure that the data allow us to uniquely identify the causal parameter we are interested in (Angrist & Pischke, 2009; Reeb, Sakakibara, & Mahmood, 2012).

In the following sections, we argue that missing firms from the dataset poses two separate challenges to identifying the spillover effect. The first is a measurement error problem that arises from mis-measuring the key regressor foreign presence, and the second is a selection effect due to the censored selection process in datasets such as manufacturing census surveys that leave out small firms.

Incomplete Datasets, Measurement Error, and Identification

Measures of foreign presence computed from incomplete datasets are likely to be incorrect. Secondary datasets, as we have seen, are often slanted toward publicly traded firms, and are therefore very likely to miss foreign wholly owned subsidiaries, unlisted joint ventures between foreign and domestic firms, and at least some privately held domestic firms. When foreign wholly owned and joint ventures are missing, measures of foreign presence computed from these datasets quite likely underestimate true foreign presence. Missing privately held domestic firms, on the other hand, can have the opposite

effect, and yield overestimates of foreign presence. When both unlisted foreign and domestic firms are missing simultaneously, some of the measurement error may cancel out, but it is unlikely that the underestimate in foreign presence due to missing foreign firms will exactly offset the overestimate due to missing domestic firms. Chances are that measures of foreign presence in secondary datasets will still not accurately reflect the true level of foreign presence in the industry.

Missing small firms, as most manufacturing census datasets do, also corrupts measures of foreign presence. But anticipating the exact direction of measurement error is not straightforward, because it depends on the relative sizes of foreign and domestic firms in the industry. If the small firms that are missing are mostly foreign ones, the effects are basically the same as when missing foreign firms, that is, foreign presence is underestimated. If, on the other hand, as Altomonte and Pennings (2009) suggest in the case of their Romanian dataset, small firms are predominantly domestic, then missing small firms potentially overestimates the true foreign presence. So in summary, even though we cannot always forecast the direction of measurement error, it is very likely that foreign presence is incorrectly measured in both secondary and manufacturing census datasets as a result of their incompleteness.

Could incorrect measures of foreign presence affect the estimation of the spillover effect? As we have seen, Eq. (1) represents the basic relationship we want to estimate. Denoting foreign presence measures computed from incomplete datasets in lowercase letters, fp_j , we can write

$$fp_j = FP_j + U_j \quad (2)$$

where uppercase FP_j is the true foreign presence, and U_j is the measurement error. For the moment, we will assume that U_j has the standard properties of 0 mean and constant variance. Given that what we are able to compute from (incomplete) datasets is fp_j and not FP_j , we cannot estimate the relationship between $Prod_{ij}$ and FP_j as in Eq. (1). Instead, we are able only to estimate how $Prod_{ij}$ varies with fp_j :

$$Prod_{ij} = \theta_0 + \theta_1 fp_j + \varepsilon_{ij} \quad (3)$$

To distinguish the true spillover effect, β_1 , from what we are able to estimate from incomplete datasets, we denote the latter by θ_1 . When correctly identified, θ_1 will equal the true spillover effect, at least in its expected value. But it is fairly easy to see that this is unlikely to be true. If we plug Eq. (2) into

Eq. (3), we get

$$Prod_{ij} = \theta_0 + \theta_1 (FP_j + U_j) + \varepsilon_{ij} \quad (4)$$

which, when rearranged, becomes

$$Prod_{ij} = \theta_0 + \theta_1 FP_j + (\theta_1 U_j + \varepsilon_{ij}) \quad (5)$$

The new error term $(\theta_1 U_j + \varepsilon_{ij})$ in Eq. (5) is now correlated with the regressor FP_j , because they both contain the common element U_j . Intuitively, this correlation between FP_j and the error term means that FP_j no longer captures only the effect of foreign presence. Since the error term contains all unobserved determinants of domestic firm productivity, a correlation with the error term basically means that FP_j also picks up the effects of some of these unobserved influences on domestic productivity. The estimated effect of FP_j when it is incorrectly measured is thus very likely to be biased (Carroll, Ruppert, Stefanski, & Crainiceanu, 2006; Chen, Hong, & Nekipelov, 2011). More importantly, the observed data could be a realization of the true spillover specification (Eq. (1)), but plausibly also of the measurement-error-tainted specification (Eq. (4)). We usually will not be able to tell which estimates we are observing in the data. And, in this sense, we have an identification problem.

The econometric literature on measurement error gives us some sense of how the estimate of θ_1 is related to the true parameter, β_1 . In the simplest case, if we assume classical measurement error, that is, that measurement error does not vary with the true values of foreign presence, then θ_1 is an unbiased and consistent estimator not of the spillover effect β_1 , but instead of the term $\lambda^* \beta_1$, where λ is the proportion of variance in fp_j that is due to the variance in FP_j (Carroll et al., 2006; Fuller, 1987; Kipnis, Carroll, Freedman, & Li, 1999). That is,

$$E(\theta_1) = \lambda^* \beta_1 \quad (6)$$

and

$$\lambda = \frac{Var(FP_j)}{Var(FP_j) + Var(U_j)} \quad (7)$$

Given that λ is a ratio and, as such, is bounded between 0 and 1, we can be sure that measurement error in this case has an attenuating effect. The estimated spillover effect will underestimate the true effect.

This result, however, depends on the assumption that measurement error is independent of the true values of foreign presence. For this to hold in our case, there should be no systematic differences between industries in how foreign presence is under- or overestimated. This could be true for datasets that

miss small firms, because there is no strong reason to believe that the mismeasurement of foreign presence increases or decreases with its true values in different industries. But in datasets that miss foreign firms, this assumption is unlikely to be true. Mis-measurement is likely to be more severe in industries where foreign firms are prevalent, and less so in industries where contractual forms of foreign investment, such as licensing agreements with domestic firms, are more common. When measurement error is not independent of the true values of the regressor, Carroll et al. (2006), among several others, show that the resulting spillover effect estimate can be an overestimate.⁵

Figure 1 offers a simple and intuitive visual representation of this. As we have mentioned, when the dataset misses foreign firms, the underestimate of foreign presence will be more severe in industries where foreign firms are more prevalent. When estimating spillover effects, this means that productivity differences between domestic firms get mapped onto incorrect variations in foreign presence. And when productivity gains are attributed to smaller changes in foreign presence than they should be, this can result in overestimating the spillover effect. In Figure 1, FP1 and FP2 are the true measures of foreign presence in two industries, and PROD1 and PROD2 are the corresponding conditional means of domestic firm productivity in each of those industries. The difference between PROD1 and PROD2, a , is the true productivity gain, on average, for domestic firms that is due to the different levels of foreign presence FP1 and FP2. But, all else constant, when FP2 is underestimated and measured as $fp2$ owing to missing foreign firms from the dataset, the observed productivity gain a is erroneously attributed to a significantly smaller change in foreign presence, that

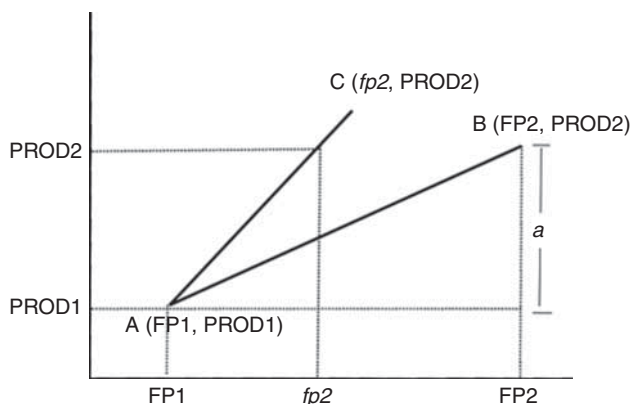


Figure 1 Illustration of the bias in estimated spillover effect when foreign presence is incorrectly measured.

is, from FP1 to $fp2$. The spillover effect from foreign firms is thus overestimated.

There are several critical points that emerge from the discussion this far. First, it is clear that the true spillover effect is not identifiable in incomplete datasets unless we know something about the reliability of our measure of foreign presence. Given Eq. (7), this means at least knowing the variances of the true foreign presence and its measurement error.

Second, if we go ahead and estimate the spillover effect anyway, we are making an implicit identification assumption that the variance of measurement error, $Var(U_i)$ in Eq. (7), is 0. Only then will our estimate represent the true effect. Given the expression in Eq. (2), this assumption implies that all the variance in fp_j is due to variance in the true foreign presence FP_j . For all practical purposes, this means we need to assume there is no measurement error. When this assumption does not hold, we are likely to obtain biased estimates of the spillover effect. And, contrary to a widely held belief, this bias need not always be downward.

Third, while for analytical ease we used the basic spillover specification in Eq. (1) to demonstrate the effect of measurement error, our observations extend to more complex specifications as well. For example, Castellani and Zanfei (2006) replace the term FP_j with a better measure of the stock of foreign technology, $FP_{j,RD}$, which gives greater weights to foreign firms that are R&D-intensive. It is very likely that this measure too is incorrect in incomplete datasets. If the most R&D-intensive affiliates are likely to be wholly owned (Buckley & Casson, 1976; Caves, 1996; Hennart, 1982; Hymer, 1976), then the term $FP_{j,RD}$ in secondary datasets (which are biased toward publicly listed firms, and hence against wholly owned subsidiaries) could grossly underestimate foreign presence and overestimate the spillover effect. Specifications that condition the spillover effect on domestic firm absorptive capacity, or on institutional characteristics, are not free from measurement error either. Using interaction terms to estimate conditioning effects is conceptually similar to estimating a main effect specification such as in Eq. (1) separately for different values of the moderating variable, and then comparing the coefficients (Aiken & West, 1991). So conditioning the spillover effect on absorptive capacity of domestic firms is akin to estimating specification (1) in subsamples of firms with high, medium, and low absorptive capacity, and then examining whether the coefficients from each subsample are significantly different. This comparison can be meaningless when the

coefficients in each subsample are estimated incorrectly. So the measurement error problems discussed above are equally relevant to these specifications as well. Moreover, researchers have shown that measurement error bias is likely to worsen with the inclusion of other covariates that are correlated with the mismeasured key regressor (Bound, Brown, & Mathiowetz, 2001; Carroll et al., 2006). Inclusion of interaction terms, which are very likely correlated with the key regressor, could potentially worsen measurement error bias.

A fourth important point is that while scholars typically employ fixed effects or difference score estimation to control for various sources of endogeneity in spillover studies, these estimation procedures worsen identification difficulties due to measurement error (Bound et al., 2001; Hausman, 2001). As Hausman (2001: 58) puts it: "One situation where the bias (due to measurement error) increases ... is when the regression specification is panel data with fixed effects". And as Bound et al. (2001: 3714) put it: "Moving from levels to changes intensifies the bias due to errors in measuring the independent variable". Also, in empirical estimation, specifications that also include a control term for competition may not help identification, because the competition term could be meaningless if the dataset leaves out foreign firms.

So, in effect, many of the theoretical and empirical extensions of the basic model that could potentially improve identification in the general case may not do so in incomplete datasets.

Incomplete Datasets, Selection Effects, and Identification

The second challenge to identifying spillover effects in incomplete datasets is the censored selection process by which firms are included in the dataset. This is largely a problem with estimations based on manufacturing census datasets where small firms are missing, but could also apply to secondary datasets if the privately held domestic firms that are missing also happen to be the small firms in the economy.⁶ In this section, as in the previous one, we first lay out the issues formally, and then show that identification of the spillover effect is difficult.

Let Z be a binary term that denotes whether or not a given firm is observed in the dataset. Then, in manufacturing census datasets, and other datasets such as Amadeus where small firms are left out from the data, $Z=0$ for small firms, and $Z=1$ for all other firms. As is clear from Eq. (1), our primary interest is in the conditional mean function $E(Prod_{ij}|FP_j)$, the

change in average domestic firm productivity for different levels of foreign presence. Following Manski (1995) and Horowitz and Manski (1998), we can write out this conditional expectation $E(Prod_{ij}|FP_j)$ as

$$E(Prod_{ij} | FP_j) = E(Prod_{ij} | FP_j, Z=1) * P(Z=1 | FP_j) + E(Prod_{ij} | FP_j, Z=0) * P(Z=0 | FP_j) \quad (8)$$

That is, the true conditional mean of domestic productivity equals the weighted sum of conditional means for firms that are observed and missing from the dataset, with the censoring probabilities as weights. One can immediately see that the only component of the expression in Eq. (8) that can be computed from incomplete datasets is $E(Prod_{ij}|FP_j, Z=1)$, the conditional mean of domestic firm productivity for firms that are observed. The censored sampling process provides no information on any of the other components in Eq. (8), which are $E(Prod_{ij}|FP_j, Z=0)$, $P(Z=0|FP_j)$, and $P(Z=1|FP_j)$. Identifying the true spillover effect is hence impossible. Manski (1995: 24) makes this point in general terms where y is the dependent variable and x the regressor, when he says: "Censoring makes it impossible to learn anything about the expected value $E(y|x)$ conditional on x ".

Studies that overlook this selection issue are in effect only modeling the spillover effect for firms that are observed, that is, $E(Prod_{ij}|FP_j, Z=1)$. From Eq. (8), we can see that this will equal the true effect only when there is no censoring in the selection process, that is, when each firm is equally likely to be observed or missing, $P(Z=0|FP_j) = P(Z=1|FP_j) = 0.5$, and when the distributions of productivity for observed and missing firms are the same, that is, $E(Prod_{ij}|FP_j, Z=0) = E(Prod_{ij}|FP_j, Z=1)$. So, in effect, studies estimating spillover effects from datasets that miss small firms are implicitly making identification assumptions of no censoring and no heterogeneity in spillover effects.

We know that neither of these assumptions is tenable. Small firms are more likely than other firms to be missing in census datasets, and also, owing to their lower levels of absorptive capacity, less likely to experience productivity gains from the presence of foreign firms (Altomonte & Pennings, 2009; Blalock & Simon, 2009; Zhang et al., 2010). This link between size and spillover heterogeneity has been well articulated in the literature. For productivity gains from spillovers to materialize, domestic firms need to possess complementary capabilities that

allow them to implement new (foreign) technologies. And several scholars have argued that large firms are more likely to have better complementary capabilities. For example, Blalock and Simon (2009: 1099) say:

Among those firms that gain access to the multinationals' technology, those with greater complementary capabilities should be able to exploit the technology better, and thus enjoy greater increases in productivity Generally, larger firms are likely to have more of these complementary capabilities.

Similarly, Zhang et al. (2010: 974) suggest that:

Relative to small firms, large domestic firms' greater stock of internal resources and knowledge can be used as complementary assets to utilize the variety of technologies and management practices brought by foreign firms.

Besides complementary capabilities, adopting new technology often entails additional investment in plant and equipment. In this respect, too, large firms are potentially more likely, relative to small firms, to be able to raise external finance for making these investments (Zhang et al., 2010). For these and other reasons, scholars have linked the size of firms with their ability to absorb spillovers, and have argued that "relative to small firms, large domestic firms have greater absorptive capacity to recognize and understand the variety of technologies and management practices brought by foreign firms" (Zhang et al., 2010: 973).

So, in sum, there is very little reason to believe that the productivity distributions for small firms that are missing are the same as those for all other non-missing firms. In fact, it is quite likely, not only that $E(Prod_{ij}|FP_j, Z=1) > E(Prod_{ij}|FP_j, Z=0)$, especially at higher levels of foreign presence, but also that, on average, the productivity of small firms at higher levels of foreign presence is negative, owing to market-stealing effects (Aitken & Harrison, 1999). And if $E(Prod_{ij}|FP_j, Z=0)$ in Eq. (8) is for this reason negative, it could very well be that studies ignoring the selection effect and thus modeling only $E(Prod_{ij}|FP_j, Z=1)$ are overestimating the true spillover effect.

There are three critical points that summarize this discussion. First, in order to correctly identify the spillover effect and model $E(Prod_{ij}|FP_j)$, we need information on the conditional distribution of productivity for unobserved small firms, as well as the censoring probabilities as depicted in Eq. (8). These are usually unknown in most datasets, which makes identifying the true spillover effect problematic. Second, studies that do not take the selection effect

into consideration are basically leaving out the second half of the expression in Eq. (8). This identifies the spillover effect only under the assumption that there is no selective censoring of small firms, and that the distribution of outcomes for small firms does not differ from that of non-missing firms. And third, as we have just discussed, it is plausible that spillover effects are overestimated when these identifying assumptions do not hold.

MONTE CARLO SIMULATION

To empirically demonstrate the consequences of estimating spillover effects in incomplete datasets, we use a series of Monte Carlo simulations. We chose simulations as our key methodology for three main reasons. First, simulations are most useful when formal logic alone does not provide all answers to the research question.⁷ While we presented measurement error and selection problems separately in our discussion above, in most empirical studies they occur simultaneously. For example, in datasets that miss small firms, there is potential measurement error in computing the key regressor, and there is a selection problem. Through formal logic, we can deduce the potential effects of these problems individually, but there is no real way to predict their combined effect. Monte Carlo simulations are the only way to know. Second, formal logic cannot tell us much about the magnitude of the potential biases in estimated spillover effects. We know from the analysis above that spillover effects in incomplete datasets may not be properly identified, but the question still remains as to whether the estimated effects are substantially different from the true effect. Since deductive logic cannot tell us much about magnitudes, we need simulations. Finally, the simulations we describe below allow us a controlled environment similar to that of a laboratory experiment where we have complete control over the manipulations. We do have to make some basic (but defensible) assumptions in simulations, but this is a worthwhile trade-off, considering the control we have over the manipulations. For example, we are able to isolate the problems of measurement errors and selection effects to test their individual effects on the spillover effect estimate, and then introduce them together to observe their combined effect. And all the while, given the controlled environment, we are fairly certain that the effects we observe are purely the result of our manipulations. There are no other confounding effects.

The basic principle underlying Monte Carlo simulation is that the behavior of a statistic such as the

estimated spillover effect in different types of incomplete samples “can be assessed by the empirical process of actually drawing lots of (incomplete) samples and observing this behavior” (Mooney, 1997:4). The basic mechanics are fairly simple. We start by defining the population dataset, that is, the dataset of all firms in the economy. We will assume that the Chinese manufacturing census dataset (see Chang & Xu, 2008, and Zhang et al., 2010, for a description of this dataset) is the true population dataset.⁸ Then, as we describe in detail below, we define the “true” spillover effect β_1 . So now we have a population of firms, and know the true spillover effect. Then, to assess the properties of estimates of this effect in incomplete samples, we draw samples that miss either foreign or small firms, and estimate the spillover effect within each sample.⁹ We describe the specific sampling rules we used to draw these incomplete samples below in more detail. For each type of incompleteness, we draw 10,000 different samples and regress domestic firm productivity on measures of foreign presence calculated from the data in the sample. From each regression, we thus have an estimate of the spillover effect θ_1 (see Eq. (3)). Using these estimates, we then construct the sampling distribution of θ_1 , and examine whether expected value equals the true, known, spillover effect. If the expected value equals the true effect (which we have fixed upfront), we conclude that the spillover effect estimate is unbiased. That is, accurate inferences about the true spillover effect can be made from estimates of the effect based on incomplete samples. On the other hand, if the expected value falls far away from the true effect, we conclude that incomplete data do bias estimates of the spillover effect. In other words, we are not correctly identifying the spillover effect in such datasets.

We run a series of three simulations. In the first, which serves as the baseline case, we test the effect of measurement error alone, with a very basic spillover specification. The second simulation is similar, except that we use Castellani and Zanfei’s (2006) measure of foreign presence, which gives greater weights to R&D-intensive foreign subsidiaries. In the third simulation, we explore more complex effects of incomplete datasets on the estimation of spillover effects. We examine the combined effect of measurement error and selection effects, as well as how several recent theoretical specifications perform in incomplete datasets. For example, leveraging the panel data nature of the Chinese dataset, we explore what happens when fixed-effects estimation is employed, as well as when a control term is included

for industry competition. Taken together, these simulations provide us with a clear picture of the effects of estimating spillover effects in incomplete datasets, starting from the simplest specification in Simulation 1 to more complex specifications in Simulation 3. A summary of the simulations is provided in Table 1, and we describe them in more detail below.

Simulation 1

As mentioned before, we treat the set of firms included in the manufacturing census dataset from the Chinese NBS as our true population. In Simulation 1, we use only data for the year 2006, which include close to 270,000 firms from 29 different two-digit SIC industries. (In later simulations we extend this, and use a panel dataset for years 1999–2006.) The next step is fixing the true spillover effect, and for this we first need to compute foreign presence variables.

Prior literature has mainly used two kinds of measures of foreign presence. The first is the weighted average of foreign equity ownership of all firms in the industry, using each firm’s share in total sales (Javorcik, 2004) or total industry employment (Aitken & Harrison, 1999) as weights. Javorcik (2004), for example, describes her measure as foreign equity participation averaged over all firms in the (industry), weighted by each firm’s share in the (industry) output (612), which is represented as

$$FP_{j, equity} = \frac{\sum \text{Foreign ownership share}_{ij} \times \text{Sales}_{ij}}{\sum \text{Sales}_{ij}} \quad (9)$$

The second commonly used measure of foreign presence is a ratio, usually of foreign to total employment in the industry (Chang & Xu, 2008; Meyer & Sinani, 2009):

$$FP_{j, emp} = \frac{\text{Number of employees in foreign firms}_j}{\text{Total number of employees}_j} \quad (10)$$

Given that we have information on each firm’s size and foreign ownership, we are able to calculate the foreign presence for each industry in the dataset. Since we consider this dataset as our “true” population, we also treat measures of foreign presence computed from it as “true” values of foreign presence. We represent true foreign presence with the uppercase letters FP_j .

Now we are ready to specify the true spillover effect in the population. We do this by assigning a spillover coefficient to each domestic firm as a draw from a normal distribution with +2 as mean.¹⁰ So while each domestic firm’s spillover effect is

Table 1 Summary of simulations

	Simulation 1	Simulation 2	Simulation 3
Objectives	Baseline simulation: Test the effects of measurement error on basic spillover specifications. Selection effects held constant	Testing the behavior of spillover estimates when using Castellani and Zanfei's (2006) measure of foreign presence	Showing the effects of selection and measurement problems while using more complex spillover specifications (e.g., fixed effects, controls for competition)
Population data	Cross-section data for the year 2006 from the China NBS manufacturing census dataset. We assume this is complete	Same as in Simulation 1	Panel data on Chinese firms from the China NBS manufacturing census dataset for years 1999–2006
Fixing the true spillover effect	For each domestic firm, we assigned a spillover effect (β_1) as a draw from a normal distribution with mean +2. We also fixed $\beta_0 = 10$ and then calculated the productivity of each domestic firm from the linear combination $Prod_{ij} = \beta_0 + \beta_1 FP_j$ Since we have two measures of the true foreign presence, $FP_{j,equity}$ and $FP_{j,emp}$, in effect we computed two productivity terms for each firm, $Prod_{ij,equity}$ and $Prod_{ij,emp}$, by plugging the appropriate foreign presence term into the linear combination	Same as in Simulation 1, except that the linear combination we used to calculate each domestic firms productivity was $Prod_{ij} = \beta_0 + \beta_1 fp_{j,RD}$	We estimated the spillover effect from the population dataset using the following specifications, and for each specification assumed that the estimated spillover effect was the true β_1 $Prod_{ijt} = \theta_0 + \theta_1 FP_{jt,equity} + \eta_{it} + \varepsilon_{ijt}$ $Prod_{ijt} = \theta_0 + \theta_1 FP_{jt,emp} + \eta_{it} + \varepsilon_{ijt}$ $Prod_{ijt} = \theta_0 + \theta_1 FP_{jt,equity} + \theta_2 herf_{ijt} + \eta_{it} + \varepsilon_{ijt}$ $Prod_{ijt} = \theta_0 + \theta_1 FP_{jt,emp} + \theta_2 herf_{ijt} + \eta_{it} + \varepsilon_{ijt}$ In each case, we assumed that θ_1 is the true β_1
Sampling rules applied	(1) Draw samples that miss foreign firms with ownership greater than 0.9 (2) Draw samples that miss small firms, that is, with sales revenues (or employee numbers) less than the median value in the population	Same as in Simulation 1	Draw samples that miss small firms, that is, with sales revenues (or employee numbers) less than the median value in the population
Models estimated in each (incomplete) sample	The specifications: $Prod_{ij} = \theta_0 + \theta_1 fp_{j,equity}$ $Prod_{ij} = \theta_0 + \theta_1 fp_{j,emp}$	The specification: $Prod_{ij} = \theta_0 + \theta_1 fp_{j,RD}$	All the above models, but replacing $FP_{j,equity}$ and $FP_{j,emp}$ with their sample counterparts $fp_{j,equity}$ and $fp_{j,emp}$

different, it comes from the same underlying distribution. Also, by drawing from the same distribution, we are keeping selection effects constant.¹¹

Now that we have β_1 for each firm, we compute the productivity of each domestic firms as the linear combination

$$Prod_{ij} = \beta_0 + \beta_1 FP_j \quad (11)$$

where, as usual, $Prod_{ij}$ is the total factor productivity of domestic firm i , FP_j is the “true” foreign presence in the industry j to which firm i belongs, β_0 is a constant (we used the value 10), and β_1 is the “true” spillover effect for each firm in the population. Given that we have two measures of foreign presence (see Eqs. (9) and (10)), we actually specified two separate foreign presence–productivity

relationships for each firm:

$$Prod_{ij,equity} = \beta_0 + \beta_1 FP_{j,equity} \quad (12)$$

$$Prod_{ij,emp} = \beta_0 + \beta_1 FP_{j,emp} \quad (13)$$

In doing this, we have just specified the exact process that creates variance in the dependent variable. $Prod_{ij,equity}$ and $Prod_{ij,emp}$ now depend solely on variations in $FP_{j,equity}$ and $FP_{j,emp}$, respectively, and since β_1 is drawn from a distribution with mean 2, we expect the estimated spillover effect in its expected value to equal 2.

The next step was to create a dataset of the estimated spillover effect in various kinds of incomplete samples. To do this, we followed certain sampling rules, and under each sampling rule drew 10,000 independent samples. For each sample, we

calculated two measures of foreign presence in the sample, $\hat{p}_{i,equity}$ and $\hat{p}_{i,emp}$. These are calculated using the same formulas as in Eqs. (9) and (10), but this time using the incomplete samples we have drawn from the population (hence the lowercase letters). We then estimated a model similar to the specification in Eq. (3) by regressing our measures of productivity on foreign presence. Specifically, we ran the following regressions:

$$Prod_{ij,equity} = \theta_0 + \theta_1 \hat{p}_{i,equity} \quad (14)$$

$$Prod_{ij,emp} = \theta_0 + \theta_1 \hat{p}_{i,emp} \quad (15)$$

We tried to replicate the patterns of incompleteness in past studies in the way we drew samples. Essentially this meant excluding firms from different parts of the foreign ownership and size distributions (please see endnote 9 for the justification for why we did this). We defined high levels of foreign ownership as when more than 90% of a firm's equity was owned by foreign parents. To understand what happens when foreign firms with high levels of foreign ownership – that is, foreign wholly owned firms – are missing from the data, we drew samples that contained only 10% of firms that had foreign ownership greater than 90%, but contained all other firms in the population. Similarly, to understand the effect of missing small firms, we drew samples with only 10% of firms in the small category, that is, with sales or employment less than the median value in the population. As part of robustness tests, we replaced the sampling criteria with different threshold values as well. For example, we re-ran our simulations defining foreign firms as those with greater than 50% and 10% foreign ownership, respectively, and defining small firms as those smaller than the 25th percentile value in our data. As we discuss later, our results are robust across these different definitions.

Simulation 2

Whereas the previous simulation gives us some baseline results, the following two simulations seek to explore the effect of estimating slightly more complex spillover specifications in incomplete datasets. In particular, Simulation 2 uses an arguably better measure of foreign presence, $FP_{i,RD}$, which gives greater weights to knowledge-intensive foreign firms. While the core features of the simulation remain the same as in Simulation 1, there are differences in the way we specify the true spillover effect, and in the model we estimate in each sample draw.

To fix the true spillover effect, we computed the productivity of each domestic firm from the following linear combination:

$$Prod_{ij,RD} = \beta_0 + \beta_1 FP_{i,RD} \quad (16)$$

where, as in Simulation 1, β_0 was fixed at 10, and β_1 for each firm was drawn from a normal distribution with mean +2. But, unlike in Simulation 1, foreign presence was calculated as in Castellani and Zanfei (2006: 182) as “foreign capital weighted by each foreign firms R&D intensity”:

$$\begin{aligned} \text{Foreign Presence}(FP_{i,RD}) \\ = \sum (\text{Foreign ownership share}_{ij} \times K_{ij} \times RD_{ij}) \end{aligned} \quad (17)$$

Accordingly, the model estimated in each incomplete sample draw was also amended to

$$Prod_{ij,RD} = \theta_0 + \theta_1 \hat{p}_{i,RD} \quad (18)$$

where $\hat{p}_{i,RD}$ was Castellani and Zanfei's foreign presence measure calculated from the incomplete sample.

Simulation 3

In the previous two simulations, since we specified the spillover effect for each firm as a draw from the same distribution, we were keeping selection effects constant. In other words, no category of firm had its effects drawn from a distribution with higher or lower mean values. In Simulation 3, we relax this assumption and allow for both selection and measurement error effects. This more closely resembles the problems that real studies face. We do this by changing the way we specify the true effect. Instead of assigning effects for each firms as draws from the same distribution, we simply estimate the spillover effect in the full Chinese dataset, that is, in our population, and assume that estimate to be the true spillover effect.

Then, as in the case of the previous simulations, we draw incomplete samples and estimate the spillover effect in each. Unlike in previous simulations, if the estimated spillover effect in samples that miss small firms differs from that in the population, it could now be due to both selection effects and measurement error effects. In other words, it could be because we leave out a category of firms that do not benefit from spillovers, and because we are incorrectly measuring foreign presence.

Besides allowing for selection and measurement error effects, in this simulation we also use two further specifications that supposedly allow for better identification: a model with firm fixed effects, and one that includes a control for competition.

As noted before, fixed effects are usually employed to correct for a host of endogeneity issues, and a control for competition is usually included to partial out spillover from competition effects. Simulation 3 examines how these identification strategies perform in incomplete datasets.

So, in effect, the main differences from Simulations 1 and 2 lie in the specifications we run in the incomplete samples, and in how we specify the true spillover effect. To be clear, we estimated the following models in the incomplete datasets:

$$Prod_{ijt} = \theta_0 + \theta_1 fp_{jt, equity} + \eta_i + \varepsilon_{ijt} \quad (19)$$

$$Prod_{ijt} = \theta_0 + \theta_1 fp_{jt, emp} + \eta_i + \varepsilon_{ijt} \quad (20)$$

$$Prod_{ijt} = \theta_0 + \theta_1 fp_{jt, equity} + \theta_2 herfi_{jt} + \eta_i + \varepsilon_{ijt} \quad (21)$$

$$Prod_{ijt} = \theta_0 + \theta_1 fp_{jt, emo} + \theta_2 herfi_{jt} + \eta_i + \varepsilon_{ijt} \quad (22)$$

where η_i denotes firm-specific fixed effects, and $herfi_{jt}$, our proxy for competition, was a market concentration measure computed as “the sum of the squared markets shares of the four largest producers in a given sector” (Javorcik, 2004: 614). For the first two specifications (Eqs. (19) and (20)), we used the estimates we obtained from a fixed-effects regression

of productivity¹² on foreign presence (measured first as $FP_{j, equity}$ and subsequently as $FP_{j, emp}$) in the population dataset as our “true” spillover effects. Similarly, for Specifications (21) and (22), we assumed true spillover effects to be the ones we obtained from estimating the same models in the full dataset.

RESULTS

The results of our simulations are presented in Tables 2 to 4. In each table, we report the true spillover effect as well as the expected value of estimated spillover effects in models estimated from different types of incomplete datasets. The comparison between the true effects and the estimated ones, in terms both of statistical significance and of magnitude, gives us a sense of what happens when the implicit identification assumptions in estimating spillover effects from incomplete datasets do not hold.

The results of Simulation 1, the baseline case, are in Table 2. Panel A shows the true and estimated spillover effects for a spillover specification using $fp_{j, equity}$ as the key regressor, and Panel B shows the same for one using $fp_{j, emp}$ instead. Columns 2 and 5 show the true spillover effect, and Columns 3 and 6

Table 2 Results of Simulation 1: The effect of measurement error on spillover effect estimates in basic models

Panel A		
(1)	(2)	(3)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ij} = \theta_0 + \theta_1 fp_{j, equity} + \varepsilon_{ij}$
1 Missing foreign firms	2.0	3.52
2 Missing small firms (in terms of sales)	2.0	1.94
3 Missing small firms (in terms of employees)	2.0	1.91
Panel B		
(4)	(5)	(6)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ij} = \theta_0 + \theta_1 fp_{j, emp} + \varepsilon_{ij}$
1 Missing foreign firms	2.0	3.92
2 Missing small firms (in terms of sales)	2.0	1.86
3 Missing small firms (in terms of employees)	2.0	1.89

Notes: The null hypothesis that there is no difference between the mean estimated spillover effect and the true effect was rejected in a *t*-test ($p < 0.001$) for all simulations.

$fp_{j, equity}$ is measured as: $\Sigma \text{Foreign ownership share}_{ij} * \text{Sales}_{ij} / \Sigma \text{Sales}_{ij}$.

$fp_{j, emp}$ is measured as: $\text{Number of employees in foreign firms}_j / \text{Total number of employees}_j$.

show the average estimated effect over 10,000 regressions. The rows correspond to different types of incompleteness in the data.

The results suggest that spillover effects from datasets that miss foreign firms overestimate the true spillover effect. While the true effect is 2, the average estimated effects are 3.52 and 3.92 when using $\hat{fp}_{j,equity}$ and $\hat{fp}_{j,emp}$, respectively, as regressors. These differences are statistically significant in a *t*-test ($p < 0.001$). To provide a sense of the magnitude of overestimation, we calculated that while the true increase in productivity between firms in industries at the 10th and 90th percentile values of foreign presence is 5%, it is (over)estimated as 10% and 13%, respectively, when using $\hat{fp}_{j,equity}$ and $\hat{fp}_{j,emp}$ as regressors. Note that in this simulation we kept selection effects constant, so the difference between true and estimated spillover effects is due purely to measurement error. Missing small firms, on the other hand, underestimates the true spillover effect. While differences between true and estimated spillovers are statistically significant, the magnitudes are relatively small. The expected values of estimated coefficients vary between 1.86 and 1.94, depending on whether $\hat{fp}_{j,equity}$ or $\hat{fp}_{j,emp}$ is the regressor, and on whether we censor small firms based on sales revenues or number of employees. All we can say from these results is that measurement error due to missing small firms puts a downward pressure on the estimated spillover effect, but perhaps not in the order of a substantial magnitude.

Simulation 2 examines whether an arguably better measure of foreign presence, $\hat{fp}_{j,RD}$, does a better job of identifying the spillover effect. From the results reported in Table 3, it appears that using Castellani and Zanfei's measure is almost equally bad when missing foreign firms, and worse when missing small

firms. When missing foreign firms, the estimated spillover effect is 3.33, and when missing small firms it is 1.53. These are significantly different from the true spillover effect, both statistically as well as in terms of effect sizes.

Table 4 reports the results of Simulation 3. As mentioned, here we test more complex specifications to see whether they are able to estimate the true spillover effect correctly. Given that we are now allowing for selection effects as well, we look only at cases where small firms are missing. Missing foreign firms should not lead to selection problems, since the estimation is done only on the group of domestic firms. There are four specifications we test in total: two models with fixed effects and $\hat{fp}_{j,equity}$ and $\hat{fp}_{j,emp}$, respectively, as regressors (results are in Panels A and B), and the same two models but with a control term for competition (results are in Panels C and D). The true spillover effects are in Columns 2, 5, 8, and 11, and the estimated effects are in Columns 3, 6, 9, and 12. Across all specifications, missing small firms results in an overestimation of the true spillover effect. To get a sense of magnitudes, for a basic fixed-effects model the true difference in productivity between firms at the 10th and 90th percentiles of foreign presence is around 3.5% when $\hat{fp}_{j,equity}$ is used to measure foreign presence. When we use $\hat{fp}_{j,emp}$ instead, the true productivity difference is in the order of 2.5%. The estimated magnitudes in datasets where small firms are missing, however, are 6% and 4%, respectively, almost twice the actual effects. Unlike in Simulations 1 and 2, these overestimations are due not only to measurement error, but to a combination of measurement error and selection effects.

If we re-run the models in Simulation 3 by replacing the $\hat{fp}_{j,equity}$ and $\hat{fp}_{j,emp}$ terms with their true

Table 3 Results of Simulation 2: The effect of measurement error on spillover effect estimates when using Castellani and Zanfei's (2006) measure of foreign presence

(1)	(2)	(3)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ij} = \theta_0 + \theta_1 \hat{fp}_{j,RD} + \varepsilon_{ij}$
1 Missing foreign firms	2.0	3.33
2 Missing small firms (in terms of sales)	2.0	1.41
3 Missing small firms (in terms of employees)	2.0	1.40

Notes: The null hypothesis that there is no difference between the mean estimated spillover effect and the true effect was rejected in a *t*-test ($p < 0.001$) for all simulations.

$\hat{fp}_{j,RD}$ is measured as: $\Sigma (Foreign\ ownership\ share_{ij} * K_{ij} * RD_{ij})$.

Table 4 Results of Simulation 3: The effect of the selection and measurement error on spillover effect estimation in a fixed-effects model

Panel A

(1)	(2)	(3)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ijt} = \theta_0 + \theta_1 \hat{fp}_{jt, equity} + \eta_i + \varepsilon_{ijt}$
1 Missing small firms (in terms of sales)	1.67	2.61
2 Missing small firms (in terms of employees)	1.67	2.08

Panel B

(4)	(5)	(6)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ijt} = \theta_0 + \theta_1 \hat{fp}_{jt, emp} + \eta_i + \varepsilon_{ijt}$
1 Missing small firms (in terms of sales)	1.13	1.92
2 Missing small firms (in terms of employees)	1.13	1.61

Panel C

(7)	(8)	(9)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ijt} = \theta_0 + \theta_1 \hat{fp}_{jt, equity} + \theta_2 herfi_{jt} + \eta_i + \varepsilon_{ijt}$
1 Missing small firms (in terms of sales)	1.20	2.10
2 Missing small firms (in terms of employees)	1.20	1.64

Panel D

(10)	(11)	(12)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ijt} = \theta_0 + \theta_1 \hat{fp}_{jt, emp} + \theta_2 herfi_{jt} + \eta_i + \varepsilon_{ijt}$
1 Missing small firms (in terms of sales)	0.89	1.57
2 Missing small firms (in terms of employees)	0.89	1.28

Notes: The null hypothesis that there is no difference between the mean estimated spillover effect and the true effect was rejected in a *t*-test ($p < 0.001$) for all simulations.

$\hat{fp}_{j, equity}$ is measured as: $\Sigma \text{Foreign ownership share}_{ij} * \text{Sales}_{ij} / \Sigma \text{Sales}_{ij}$.

$\hat{fp}_{j, emp}$ is measured as: $\text{Number of employees in foreign firms}_j / \text{Total number of employees}_j$.

values $FP_{j, equity}$ and $FP_{j, emp}$, we are in effect ruling out the effect of measurement error. Any remaining estimation bias is solely due to selection effects. The results of this exercise are in Table 5. Panels A and B shows results for specifications where $FP_{j, equity}$ and $FP_{j, emp}$, respectively, were the regressors. Note that when measurement error is kept constant in

the manner we have just described, a fixed-effect spillover model with $FP_{j, equity}$ as regressor yields on average a spillover effect of 2.80 (see Row 1, Column 3, in Table 5). The same model with $\hat{fp}_{j, equity}$ as regressor, meaning that both measurement error and selection problems are in play, yields 2.61 (see Row 1, Column 3, in Table 4). So the effect of

Table 5 The effect of the selection problem alone on spillover effect estimation in a fixed-effects model

Panel A		
(1)	(2)	(3)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ij} = \theta_0 + \theta_1 FP_{j,equity} + \varepsilon_{ij}$
1 Missing small firms (in terms of sales)	1.67	2.80
2 Missing small firms (in terms of employees)	1.67	2.40
Panel B		
(4)	(5)	(6)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ij} = \theta_0 + \theta_1 FP_{j,emp} + \varepsilon_{ij}$
1 Missing small firms (in terms of sales)	1.13	2.14
2 Missing small firms (in terms of employees)	1.13	1.78

Notes: The simulation uses $FP_{j,equity}$ and $FP_{j,emp}$ as regressors. So the effects we see are not due to measurement error, but to selection problems alone.

$FP_{j,equity}$ is measured as: $\Sigma \text{Foreign ownership share}_{ij} * \text{Sales}_{ij} / \Sigma \text{Sales}_{ij}$.

$FP_{j,emp}$ is measured as: $\text{Number of employees in foreign firms}_j / \text{Total number of employees}_j$.

measurement error, from these two results, is that it attenuates the spillover estimate, and this is consistent with the results of Simulations 1 and 2.

The overarching conclusion one could draw from the results across all the simulations is that measurement error results in overestimates of the true spillover effect when the dataset misses foreign firms, whereas it slightly underestimates it when small firms are missing. But when we allow both selection and measurement error problems, as the results of Simulation 3 show, the spillover effect is very likely overestimated.

Sensitivity Tests

To ensure confidence in these results, we examined their robustness to several parameters of the simulation. In Simulations 1 and 2, we assigned the true spillover effect for each firm as a random draw from a normal distribution with mean equal to 2. We first tested whether the assumption of normally distributed spillover effects was crucial to our results. We assigned spillover effects from a log-normal distribution, and our results remained the same. We also checked the robustness of our results to other mean values – positive and negative, as well as larger and smaller – and obtained substantially similar results.

We also varied the cut-off points of foreign ownership at which foreign firms were classified as wholly owned. In the simulations we reported, we defined wholly owned firms as those with foreign ownership greater than 90%. We examined the effects of also including in that category firms with more than 50% foreign ownership, that is, majority foreign-owned joint ventures. The effect of missing wholly owned firms then was much more pronounced than what we report in our main analyses. The mean estimated effects for specifications with $fp_{j,equity}$, $fp_{j,emp}$, and $fp_{j,RD}$ as independent variables were 4.25, 4.91, and 5.53, respectively. When we repeated it with firms with more than 10% foreign ownership considered foreign, the biases were even more severe. The mean estimated effects this time for specifications with $fp_{j,equity}$, $fp_{j,emp}$, and $fp_{j,RD}$ as independent variables were 4.78, 6.75, and 11.27, respectively. This means that the results we report here are the most conservative ones.

We used the median sales revenue and employment in our dataset as the threshold below which we considered firms as “small”. Given that firm size is log-normally distributed (Growiec, Pammolli, Riccaboni, & Stanley, 2008; Reichstein & Jensen, 2005; Simon & Bonini, 1958), this is a defensible assumption. The median-sized firm is still on the

lower end of the size distribution. Nevertheless, we also examined the consequences of redefining small firms as those smaller than the 25th percentile of firms in our dataset. Estimation biases are marginally smaller than what we report here. For example, instead of the estimated coefficient of 2.61 we report in Table 4 for a fixed-effect model in Simulation 3 that uses $fp_{j,equity}$ as the regressor (Table 4, Panel A, Row 1, Column 3), with size defined as less than the 25th percentile value in the data, we get 2.08. Similarly, for the model with $fp_{j,emp}$ as regressor (Table 4, Panel B, Row 1, Column 6), we get 1.70 instead of 1.92. These are still significant overestimates of the true spillover effects, both statistically and in terms of magnitude, and in this sense the results are substantially similar to the ones we report here.

Finally, while we examined the sensitivity of Castellani and Zanfei's (2006) foreign presence measure to measurement error in Simulation 2, we did so using a specification that did not employ fixed effects. Moreover, in Simulation 2 we kept selection effects constant. But in most practical uses of the measure, scholars use panel datasets that allow them to control for stable firm attributes through fixed effects (e.g., Castellani & Zanfei, 2007), and in all likelihood selection problems are not constant, as we hold them in the simulation. Hence, to test the robustness of our results in Simulation 2 to the inclusion of fixed effects, and allow for selection problems as well, we ran another series of simulations.¹³ These were similar in their mechanics to the specifications we ran in Simulation 3, except that we used $fp_{j,RD}$ as our key regressor. The results are reported in Table 6. The main result that emerges from these simulations is that using fixed effects still does not yield correct estimates of the spillover effect so long as the dataset is incomplete. Consistent with the results of Simulation 2, missing foreign firms overestimates the true spillover effect, even with

fixed effects. But unlike in Simulation 2, we get an overestimate of the true effect when small firms are missing. While the true effect is a coefficient of magnitude 1.34, fixed-effects specifications estimate it as either 2.62 or 2.13, depending on whether the size bias we infused was on the basis of sales revenues or number of employees. This deviation from the results of Simulation 2 is in fact unsurprising, given that while Simulation 2 held selection effects constant, we allowed both measurement and selection effects in these simulations. Given that small firms generally lack the absorptive capacity to benefit from spillovers, the estimated spillover effect from a dataset that misses small firms is likely to be overstated, irrespective of how foreign presence is measured, and whether or not fixed effects are included. In fact, the results in Table 6 are consistent with the results of missing small firms in Simulation 3, where we allowed both measurement error and selection effects to operate.

THE WEIGHTED INSTRUMENTAL VARIABLE ESTIMATOR

Given our simulation results, the critical question for future research is whether we can get better estimates of spillover effects in incomplete datasets. In this section, we demonstrate how a weighted instrumental variable estimation approach (Krasker & Welsch, 1985; Visek, 2009) can do just that. We first present the theoretical logic behind why this approach yields better approximations of the spillover effect in incomplete datasets, and follow that with an empirical demonstration using data from the Chinese NBS manufacturing census.

The gist of our argument is as follows. Recall that there are two hurdles to identification: measurement error and selection problems. A common approach to overcoming measurement error is the use of instrumental variables (Hausman, 2001). As the

Table 6 Simulation results of using Castellani and Zanfei's (2006) measure of foreign presence in panel data with fixed effects

(1)	(2)	(3)
Type of bias in the data	True spillover effect	Mean of the estimated spillover effect from 10,000 independent samples Model estimated: $Prod_{ijt} = \theta_0 + \theta_1 fp_{jt, RD} + \eta_i + \varepsilon_{ijt}$
1 Missing foreign firms	1.34	2.67
2 Missing small firms (in terms of sales)	1.34	2.62
3 Missing small firms (in terms of employees)	1.34	2.13

Notes: $fp_{j, RD}$ is Castellani and Zanfei's measure of foreign presence, measured as: $\Sigma (Foerign\ ownership\ share_{ij} * K_{ij} * RD_{ij})$. The term η_i denotes a firm-specific time-invariant effect.

economics literature has amply demonstrated, instrumenting regressors that are measured with error can yield better estimates of their causal effects. In our context, this means that employing instruments for foreign presence variables could potentially help overcome some of the measurement error biases we have documented. To deal with the selection problem, if we consider missing small firms as equivalent to an over-representation of large firms in the data, then a potential solution is to use weighted regressions wherein large firms contribute less information to the estimation process. The weighted instrumental variable estimator combines the utility of both these approaches, as we explain in more detail below.

Weighted Instrumental Variable Estimator

The basic mechanism by which an instrumental variable approach overcomes endogeneity (see Cameron & Trivedi, 2009; Murray, 2006, among several others) and measurement error problems (Bound et al., 2001; Carroll et al., 2006; Hausman, 2001) is well documented. The key idea is that when we instrument the error-prone variable with another regressor, under certain strict conditions, this will yield consistent parameter estimates (e.g., Bound et al., 2001; Hausman, 2001).¹⁴ This approach works when the instrument is correlated with the regressor, but is uncorrelated with the error term of the regression. So, for the model depicted in Eq. (1), if we denote v as a potential instrument, then the latter requirement can be represented as $E(\varepsilon_{ij}|v) = 0$ (Cameron & Trivedi, 2009). Moreover, the instrument itself should not be a relevant explanatory variable in the main model. In other words, any influence the instrument has on the dependent variable should be through its association with the key regressor. In cases where the key regressor is measured with error, an additional condition is that the instrument is also uncorrelated with the measurement error, that is, with the U_j term in Eq. (2) (Hausman, 2001).

To facilitate our explanation, we will represent the instrumental variable estimator as

$$\hat{\beta}_{IV} = (V'X)^{-1}(V'Y) \quad (23)$$

where X with n rows and k columns is the data matrix, Y is the dependent variable vector, and V contains the instruments.

While the instrumental variable approach resolves measurement error bias, it does not do much to correct for the selection issue. Usually, selection problems can be treated by Heckman-type corrections (e.g., Martin, 2013; Shaver, 1998). But, in our

case, we do not empirically observe the selection process, and are thus unable to model it. Manski (1995) and Horowitz and Manski (1998) recommend reporting bounds for the parameter estimates when faced with selection problems, but this also requires a knowledge at least of the censoring probabilities. In our case, we do not know this either. In fact, given that what we know is nothing more than the fact that small firms are likely to be missing (and not details such as how many, or with what likelihood), there is unlikely to be a perfect solution to selection problems arising from not observing them.

However, we could potentially get better approximations of the true spillover effect by employing weights in the estimation that are inversely related to the size of the firms in the dataset. The intuition behind this is fairly straightforward. The selection problems we discuss in this paper have broad similarities with those that arise from non-response or from stratified sampling in survey research. All these problems result in some groups of observations being over- or under-represented in the data available for analysis. When this happens, it is important to reinstate the relative importance of each observation in the data to match that in the population, before estimating population parameters. One way to achieve this is by weighting observations inversely to their likelihood of being included in the sample (Kish, 1990; Porter, 1973). By assigning weights in this way, firms that are over-represented in the sample contribute relatively less to the estimation of the population parameter than firms that are under-represented. We can represent the mechanics of this procedure as follows.

Let w_i represent the weight attached to observation i , such that w_i is inversely related to the probability of inclusion of observation i in the sample. Following Dumouchel and Duncan (1983), we can require that w_i be proportional to π_j/n_j , where π_j represents the proportion of the full population corresponding to industry j that observation i belongs to, and n_j represents the size of a random sample from the j th industry. Note that the inverse of w_i is the probability of observation i from industry j being included in the sample. In other words, w_i is inversely related to the over-representation of i in the sample. Now, let W be a diagonal matrix with w_i as diagonal elements. Then, in the relatively simple case of linear regression model with multiple covariates, the weighted least square estimator is

$$\hat{\beta}_W = (X'WX)^{-1}(X'WY) \quad (24)$$

It is straightforward to see that, when all the weights equal 1, this reduces to the usual OLS estimator $\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y)$. The key difference, intuitively, is that whereas the least-squares estimator minimizes the square of deviations $\Sigma(Y_i - x_i'\beta)$ to obtain coefficient estimates, the weighted least-squares estimator minimizes $\Sigma w_i(Y_i - x_i'\beta)$. Firms with higher weights contribute more to the estimation process. The most useful characteristic of the weighted estimator for our purpose is that it is a consistent estimator of population parameters when the sample is not representative of the population (Dumouchel & Duncan, 1983). It is approximately unbiased as well (Pfefferman, 1993, 1996).

A promising way to resolve measurement error and selection problems in estimating spillover effects in incomplete datasets is thus to combine the mechanics of instrumental variable estimation with those of weighted regression. In other words, an instrumental variable estimator that also weights observations differently on the basis of their selection probabilities could potentially attenuate the identification problems we have highlighted in the paper. The weighted instrumental variable estimator (Krasker & Welsch, 1985; Visek, 2009) does precisely that. Using the same matrix notation described above, we can represent this estimator as

$$\hat{\beta}_{WIV} = (V'WX)^{-1}(V'WY) \quad (25)$$

In the following empirical demonstration, we show how the weighted instrumental variable estimator provides a better approximation to the true spillover effect in the Chinese NBS data, which, as we know, misses small firms (Chang & Xu, 2008).

Empirical Demonstration

To implement the weighted instrumental variable solution, we need good instruments for the foreign presence variable, as well as weights that adequately reflect the censoring or selection probability for each observation.

Finding good instruments is hard. The credibility of any potential instrument usually rests on how convincingly the researcher presents it as meeting the conditions of a valid instrument. So, in this sense, the “cloud of uncertainty that hovers over instrumental variable estimates is never entirely dispelled” (Murray, 2006: 120). In our case, however, the cloud is certainly lighter, because as well as offering persuasive arguments for the validity of our instruments, we are able to show that redoing our Monte Carlo simulation using our instruments (together with size-based weights) gives us estimates

that are closer to the true spillover effect. It also helps that there is some precedent in the literature for the instruments we use (see Griffith, Redding, & Simpson, 2003).

We instrumented the problematic variable, $\hat{p}_i$, with variables depicting the state of technology progress in the United States, one of the biggest sources of FDI into China. Specifically, and as in Griffith et al. (2003), we instrumented foreign presence in each industry with lagged values of the growth in total factor productivity and R&D intensity for the same US industry sector. We obtained data for the former from the NBER-CES Manufacturing Industry database (Becker & Gray, 2009); R&D intensity data came from the OECD-STAN database (OECD, 2010). The justification for using these variables as instruments is that they are correlated with the presence of foreign firms in various industries in China, but are not direct determinants of the productivity of Chinese domestic firms. Growth in productivity and R&D intensity in the United States indicate the improving state of technology in the United States for firms in a given industry, which could in turn prompt FDI into China. But it is unlikely that these variables are directly correlated with the productivity of domestic Chinese firms. Any correlation is likely to be through the presence of US firms in China. Furthermore, there is very little reason to believe that any of the instruments are correlated with the measurement error in foreign presence. Measurement error is an artifact of the incompleteness of the dataset, and this has nothing to do with the state of technology of firms in the United States.

The final condition that needs to be satisfied is that our instruments are uncorrelated with any other unobserved influence on the productivity of Chinese domestic firms. We cannot completely rule this out, because if there are global improvements in technology for a particular industry, this could reflect in productivity gains for both US and Chinese firms. On the other hand, such a cross-country correlation in productivity is unlikely for two reasons. First, much of the empirical evidence in the economic growth literature points to vast differences in productivity levels and growth rates between countries (Baily & Solow, 2001; Hall & Jones, 1999), and between industries in different countries (Costello, 1993; Harrigan, 1999). Moreover, the evidence for any cross-country convergence in productivity over time, at least for the manufacturing sector, is extremely weak (Bernard & Jones, 1996). Second, scholars are increasingly in agreement on

the notion that cross-country productivity differences are not driven primarily by different factor endowments, but are “fundamentally related to differences in social infrastructure across countries” (Hall & Jones, 1999: 84). Social infrastructure refers to the quality of the local economic institutions, including the security of local property rights, provision of public goods, economic and political governance and so on. When there are distortions in the allocative efficiency of the local economic infrastructure, productivity performance is likely to be poor (Acemoglu & Dell, 2010; Bartelsman, Haltiwanger, & Scarpetta, 2013; Hsieh & Klenow, 2009; King & Levine, 1993; Rajan & Zingales, 1998). If productivity is shaped heavily by the local economic institutions, then it is unlikely that productivity across countries – especially across the United States and China – shares common antecedents. Together with the empirical evidence of the lack of convergence in productivity (Bernard & Jones, 1996), this argument makes it easier to propose productivity in the United States as an instrument for foreign presence in China.

Now we can move on to the second part of operationalizing the weighted instrumental variable solution, that is, deciding what weights to use. We know that selection problems in census datasets arise from large firms being over-represented in the dataset, and hence that we want to devalue the contribution of those firms to the estimation of the spillover effects. So w_i , the diagonal element of the weighting matrix W in Eq. (25), should decrease with the size of each firm. While this much is straightforward, there is very little theoretical guidance on the specific form of the weight to use. In fact, to use weights in our case, we first need to decide what measure of firm size to use, and then the specific weighting structure (i.e., to what extent we should down-weight large firms). We chose sales revenue as our measure of firm size, because we know that small-firm censoring in the Chinese dataset is on the basis of a sales revenue threshold (Chang & Xu, 2008). So, naturally, sales revenue would be an ideal variable on which to try to weight firms. In additional tests, however, we also used a relative size measure – each firm’s market share – to see whether the principle underlying our estimation approach extends to other size measures as well. And as we report below, using market share also gives us reasonably good approximations of the true spillover effect. Coming to the second issue of the specific weighting structure, we have several options. While a straightforward possibility is to

weight firms simply by the inverse of their firm size, if we want to down-weight large firms more heavily, we could use the squared or cubed inverse of firm size. In all cases, the weights dilute the contribution of larger firms in the estimation process, but to different degrees of extent.

To make an informed choice between weighting structures, we resorted to model selection tests. Specifically, we used the method proposed by Vuong (1989) to evaluate different model specifications. The intuition behind the Vuong test is that there is a true model (with a given weighting structure) that underlies the data generation process in our size-censored dataset. We of course do not know what this model or weighting structure is. Instead, we have competing models, one with the inverse of firm size as weighting structure, and the others with the squared or cubed inverse of firm size. The Vuong test uses the “Kullback–Leibler Information Criterion (KLIC) to measure the closeness of a model to the truth” (Vuong, 1989: 307), and provides a directional test of which of a pair of competing models is closest to the unknown true data generation model. We followed Dechow’s (1994) implementation of the Vuong test, described on pages 38–40 of her paper, and compared a model using the inverse of firm size as weighting structure with two other models using the squared and cubed inverses of firm size instead. The results of these are in Panel A of Table 7. The quantities reported in the table are Vuong Z statistics obtained from a paired comparison of models using weights defined in the rows of the table (w_1) with those using weights shown in the columns of the table (w_2). As Dechow (1994) describes, a positive and significant Z statistic implies that a model using the w_1 weight is superior to one using w_2 , and vice versa. Our first comparison was between using the squared inverse of firm size and using just the inverse of firm size. We obtained a significant and negative Vuong statistic ($Z = -16.81$, $p < 0.001$), which implies that the latter model is closer to the unknown true data generation process than the former. In other words, using inverse of firm size, rather than squared inverse of firm size, is closer to the true data generation process in our size-censored dataset. We then compared the winning model with one that uses the cubed inverse of firm size as weights. Again, we obtained a negative and significant Vuong statistic ($Z = -28.45$, $p < 0.001$), which implies that, still, weighting with the inverse of firm size is closer to the true model. Furthermore, it seems that weighting with the cubed inverse is worse than weighting with the squared inverse

($Z = -3.47$, $p < 0.001$). While we do not report this in the table, we also checked whether weighting firms by the square root of the inverse of their firm size would bring us closer to the true model. The answer, again, was no; weighting by the inverse of firm size was still the better model. Taken together, these results seem to suggest that punishing large firms any more with weights greater than the inverse of their size, or punishing them any less with weights less than the inverse of their size, only takes us farther away from the true model. Of all the different weighting structures, the inverse of firm size seems to be the one that is closest to the truth.

Now that we have reasonably good instruments for the foreign presence variable, and, from the model selection tests based on Vuong (1989), reasonably strong justification to use the inverse of firm size as weights, we are ready to run the weighted instrumental variable estimation. As in all the previous simulations, we wanted to see whether estimates of the spillover effect from this approach are, on average, any better than those we get from specifications that do not use this approach. We ran a Monte Carlo simulation similar to Simulation 3 above. The only difference was in the models we estimated. The results are reported in Panel B of Table 7. To facilitate comparison, we report three sets of estimates. The first is from a standard fixed-

effects specification, that is, one that does not use the weighted instrumental variable approach. This is in Column 3. The next result, in Column 4, is from using instrumental variable estimation but without weights. And the final result, in Column 5, is from a weighted instrumental variable model. For brevity, we report only the case where we cull small firms on the basis of sales.

There are a couple of noteworthy points in the results. Using instrumental variables alone without correcting for the size bias, which some scholars have done in the literature (e.g., Haskel et al., 2007; Keller & Yeaple, 2009) does not estimate the true spillover effect correctly. In fact, comparing Columns 3 and 4 in Panel B of Table 7, one can see that the overestimation is even worse than what one gets in a standard fixed-effects estimation. When using size-based weights in addition to instrumental variables, on the other hand, the attenuation of bias is significant. The approximation of the true spillover effect is much better when we down-weight each firm by its size. While the true spillover coefficient is 1.67, using a weighted instrumental variable estimator with weights specified as the inverse of firm size gives us, on average, a coefficient of 1.64. Compared with the coefficient estimate of 2.61 from a standard fixed-effects model estimated from datasets that miss small firms, this is a significant improvement in estimation

Table 7 Simulation results of using the weighted instrumental variable estimator to estimate FDI spillover effects

Panel A: Results of Vuong model selection tests to compare different weighting structures				
	(1)	(2)		
(1) $w_1 = (\text{Firm size})^{-1}$				
(2) $w_1 = (\text{Firm size})^{-2}$	-16.81***			
(3) $w_1 = (\text{Firm size})^{-3}$	-28.45***	-33.47***		
Panel B: Instrumental variable regression with fixed effects and size-based weights				
(1)	(2)	(3)	(4)	(5)
Type of bias in the data	True effect	Standard specification with fixed effects	Instrumental variable estimation (with fixed effects, but without weights)	Weighted IV estimation (with fixed effects and inverse of firm size as weights)
1 Missing small firms (in terms of sales)	1.67	2.61	4.88	1.64

Notes: *** shows significance at $p < 0.001$.

The comparison is between pairs of weighted instrumental variable models. In the first model in the pair, the weights are defined as w_1 . That is, the diagonal elements of the W matrix in the estimator $\hat{\beta}_{wiv} = (V'WX)^{-1}(V'WY)$ are defined as w_1 . In the second model of the pair, the weights are w_2 .

The quantities reported in the table are Vuong Z statistics. A positive and significant Vuong Z statistic implies that the model using weights w_1 is superior to one that uses w_2 . Conversely, a negative Vuong Z statistic implies that the model using w_2 as weights is better than one using w_1 .

Firm size measured as average sales revenue of the firm across years that it was in the data.

We instrumented $fp_{i, equity}$, measured as $\Sigma \text{Foreign ownership share}_{ij} * \text{Sales}_{ij} / \Sigma \text{Sales}_{ij}$, using the instrumental variables described in the text.

The coefficient reported in Column 3 is the mean of the estimated spillover effect in 10,000 independent samples when using a fixed-effects specification. The coefficients in Columns 4 and 5 are the means from 10,000 independent samples obtained from instrumental variable [$\beta_{iv} = (V'X)^{-1}(V'Y)$] and weighted instrumental variable [$\beta_{wiv} = (V'WX)^{-1}(V'WY)$] estimators, respectively.

accuracy. It is also worthwhile to note that, of all the specifications tested in this paper, this combination of instrumental variables and weights gives us the best approximation to the true spillover effect.

As an additional test, we also examined what happens if we use market share, instead of sales revenue, as weights in the weighted instrumental variable approach. Again, we used the Vuong test to compare models where each firm was weighted by the inverse of its market share with others that used the squared or cubed inverse of market share. When using market share, Vuong tests suggested that a model with the squared inverse of market share was superior to those with the inverse of market share or cubed inverse of market share. Using the squared inverse of market share as weights for each firm, we re-ran the Monte Carlo simulation. Again, we were able to obtain a very good approximation to the true spillover effect. Whereas the true effect was 1.67, we obtained an average coefficient of 1.60 from the weighted instrumental variable approach.

It is important to note that the specific weights we use in this empirical demonstration are not the critical part of our empirical recommendation. We cannot be sure that weights defined as the inverse of firm size or the squared inverse of market share, which worked in our dataset, will work equally well in other datasets.¹⁵ What is generalizable from our demonstration is, first, the principle underlying the weighted instrumental variable solution, and, second, the process of its implementation. So scholars trying to adopt this empirical correction should think about weighting structures that are appropriate for their data. What we have demonstrated is that a reasonably good choice of instruments, combined with a weighting structure selected on the basis of a rigorous model selection test (Vuong, 1989), can provide significantly better approximations to the true spillover effect in incomplete datasets.

DISCUSSION

Our aim in this paper was to examine the identification of spillover effects in incomplete datasets. We have made three key arguments. First, we argued that most recent specifications can identify the true spillover effect only if certain identifying assumptions are made, and in most cases these assumptions are unlikely to be true in incomplete datasets. Second, we showed through our simulations that when these assumptions do not hold, the estimated spillover effect in incomplete datasets is highly unlikely to represent the true effect. In all likelihood, we are

overestimating the true effect. And third, we presented a weighted instrumental variable estimator as a potentially better way of estimating spillover effects when the dataset is incomplete.

In this section, we reflect on the implications of our study for what we know this far about FDI spillovers, and suggest a few other ways we can improve the identification of spillover effects in empirical research.

Implications for Theory

The most direct implication of this paper, given that practically all empirical work this far has been based on incomplete datasets, is the possibility that past empirical results have not correctly identified the true spillover effect. Just to give a couple of examples, our results pose a challenge for findings based on secondary datasets such as Compustat. It is plausible that positive findings in such studies (e.g., Keller & Yeaple, 2009) are at least partly driven by measurement error. Keller and Yeaple do use instrumental variables as a means to overcome measurement error problems, but without some kind of weighting as well, as we have seen in Table 7, it is unlikely that this approach would have yielded unbiased estimates of the spillover effect.

In a similar vein, and as another example, our results imply that we cannot derive conclusive results from Castellani and Zanfei's (2006) study. We know that their CIS-Elios data, derived from the Amadeus database, which "contains information on the top 250,000 companies in Europe" (197), is biased toward large firms. While we do not have reliable grounds to say that Amadeus consistently misses foreign firms as well, we cannot say the same for the CIS-Elios dataset. As Castellani and Zanfei (2006) explain, their dataset merges Amadeus with the CIS survey in Europe, and in preparing the data for analysis, they dropped firms that were not present in all years of the dataset. In other words, they restricted their analysis to a balanced panel of firms. As a result of all this (arguably necessary) data management, we think it is possible that the dataset they ultimately use for the analysis also misses foreign firms. Their study does not make a claim either way. Given this, and our results in Tables 3 and 6, it is plausible that their specification does not capture the real spillover effect.

Our results also pose a challenge to what we know about spillovers from manufacturing census datasets. Again, evidence for spillovers in such datasets (e.g., Haskel et al., 2007) could be an artifact of selection problems. And, most importantly, we do

not know with any amount of certainty whether the observed correlations in the data reported in these studies are a result of the true spillover mechanism, or simply of selection problems.

While the analysis and simulations we report here have focused primarily on spillovers between foreign and domestic firms in horizontally related industries, our arguments apply equally well to estimations of vertical spillovers. The basic vertical spillover specification regresses domestic firm productivity on the presence of foreign firms in vertically related industries. *Prima facie*, there is no reason to expect that our discussion on measurement errors is invalid for vertical spillovers. Traditional measures of foreign presence, whether in horizontally or in vertically related industries, are likely to be incorrect in incomplete datasets. And this can in turn result in the same identification problems and estimation bias we have highlighted in the paper. The consequences of selection problems for vertical spillovers are slightly less straightforward to anticipate. Our key argument has been that small firms typically have lower levels of absorptive capacity, and thus estimations in datasets that leave them out might overestimate the true spillover effect. In vertical spillovers, one could argue that absorptive capacity plays a less crucial role, because in the absence of competition between them, foreign firms might be more willing to help domestic firms upgrade their productivity. In other words, even if small firms lack absorptive capacity, by virtue of greater effort from a willing foreign firm, knowledge flows can still occur. If this argument is true, then missing small firms should not have any tangible selection effect on the estimation of vertical spillover effects. However, this conclusion is based on the assumption that the only impediment to spillovers is the competitive tension between foreign and domestic firms. That is, once there is no competition, spillovers will easily occur. But, on the other hand, recent literature suggests that there are knowledge transfer difficulties, irrespective of the level of competition. For example, in a recent theoretical paper, Eapen (2012: 258) says:

Despite the willingness of foreign firms, prior literature has shown that technology upgrading of suppliers, even in the framework of formal contracts, is still difficult ... so while motivation-related (constraints) ... might be more relevant to horizontal spillovers, ... transfer-related constraints that arise from the difficulties of integrating new knowledge also very much apply to vertical spillovers.

So in essence, despite the lack of competition in vertical settings, there are still impediments to

spillovers that arise from the domestic firm's ability to absorb and integrate spillovers. And if these impediments are greater for small firms (as they very likely might be), then small firms are less likely to benefit from vertical spillovers. In fact, Blalock and Simon (2009) do find some evidence that vertical spillovers are stronger for larger than for smaller firms. If this is true, then missing small firms still creates a selection problem and a possible upward bias in estimation of spillover effects in incomplete datasets.

We need to be clear, however, that we are not dismissing prior studies entirely. As we have mentioned before, the first step to identification is to construct valid theoretical models that represent how spillovers occur. And, in this sense, prior studies including the ones we have mentioned above make significant strides forward. Identification without good theoretical models is impossible. Nor are we questioning the existence of spillovers. Our arguments and empirical simulations say nothing about whether or not there are spillovers from foreign to domestic firms. That was never our objective. What we do suggest is the need to think carefully about how our empirical approaches fare in identifying spillover effects from incomplete datasets. And our key result is that current approaches could be lacking in this regard, given that most datasets miss firms in systematic ways. Hence we need either to improve their ability to identify spillovers within current estimation frameworks, or to explore newer empirical methodologies. We have suggested weighted instrumental variable estimation as a way to improve empirical estimation in current datasets. In the following, we briefly point out a few other ways we can improve the identification of spillover effects.

Alternative Empirical Designs

Keeping weighted instrumental variable estimation aside, the perfect solution to the problems we raise in this paper is, of course, to collect data on all firms relevant to the research. This is more likely to work, if at all possible, in single-industry studies. Chung, Mitchell, and Yeung (2003), for example, study the effect of Japanese automotive assemblers in the United States on the productivity of US auto-part suppliers. Their coverage of US suppliers is fairly comprehensive; their data exclude "private and public 38 companies with only marginal automotive component involvement" (204). On the independent variable side, their coverage of Japanese FDI in automobile assembly is complete. They include data

on all Japanese auto assemblers in the United States during the period of their study. Studies of these sort are less susceptible to the biases we report in this paper, and hence, despite the practical difficulties in collecting the data, we recommend emulating them.

A second option is to explore the use of a broader range of methodologies in estimating FDI spillover effects. Clearly, qualitative approaches such as adding to the existing few case studies are a way to avoid the problems of bias that we document, but at the cost of generalizable results. Innovative use of survey data that probes directly into when and where FDI spillovers occur is another alternative (e.g., Driffield et al., 2010). Still another alternative is to use computational simulations to model spillovers. Simulation-based research has generally received less-than-deserved attention in the management literature (Harrison, Lin, Carroll, & Carley, 2007), and to our knowledge has never been used in the FDI spillover literature. Yet there have been productive uses of simulations to study various kinds of spillovers. For example, Meagher and Rogers (2004) have examined R&D spillovers through simulations, and Cantner and Pyka (1998) explore the role of absorptive capacity in technology spillovers. Neither of these studies models spillovers from foreign to domestic firms separately, but given that FDI spillovers are a special case of the phenomenon that these scholars study, modeling FDI spillovers in a simulation is feasible. A detailed specification of the simulation is beyond the scope of this paper, but certainly prior simulation-based work in related fields can serve as useful guideposts (besides Meagher & Rogers, 2004, and Cantner & Pyka, 1998, see also Bae & Coe, 2008; Centola & Macy, 2007).

CONCLUSION

Our work here hopefully brings identification issues arising from incomplete datasets to the forefront in empirical research on FDI spillovers. There are three key points that we have made. First, owing to measurement error and selection effects, it is possible to identify spillover effects in incomplete datasets only when we make certain restrictive assumptions. Second, as our Monte Carlo simulations show, spillover effect estimates are very likely to be biased when the implicit identification assumptions do not hold. And third, we have demonstrated that an instrumental variable estimation approach that uses size-based weights could offer a better approximation to spillover effects when scholars are forced to use incomplete datasets. Besides, we also see potential in using other modeling frameworks as well.

Given the far-reaching policy implications of empirical results in this line of research, it is crucial that the theoretical and empirical tools we use are capable of correctly identifying spillover effects. We are optimistic that our research reported here will help push the field closer toward that goal, and simultaneously toward what scholars agree constitute the hallmarks of high-quality empirical research (Oxley et al., 2010).

ACKNOWLEDGEMENTS

I thank the editor and three anonymous reviewers, Anthea Zhang, Anthony Obeyesekere, Dave McKendrick, Grzeg Trojanowski, Peter Zamborsky, Rajiv Krishnan, Rekha Krishnan, Roger Smeets, Shivani Sourindre, and seminar participants at the universities of Melbourne, Sydney, and Auckland for their comments. I particularly thank Gaurab Aryal for helpful discussions on identification and model selection, Sea-Jin Chang for access to Chinese National Bureau of Statistics data, and Joachim Mai from Intersect for facilitating access to their supercomputer to run the simulations.

NOTES

¹Identification, in general, pertains to whether there is enough information available to allow a model to converge to one unique set of values for the parameters. If, say, two models or sets of parameters fit the data equally well, we usually cannot tell which one of the two we are estimating. In this sense, both models are not identified. But fundamentally, since the parameters of interest in the model usually represent causal relationships, identifying them means being able to trace out the causal relationships from the available data (e.g., Angrist & Krueger, 2001; Angrist & Pischke, 2009; Manski, 1995). So identifying the spillover effect basically means ensuring that the theoretical model we specify and estimate reflects the true spillover effect (cf. Koopmans, 1949; Manski, 1995). We explain this further later in the paper.

²Identification concerns in the literature this far have been around simultaneity of domestic firm productivity and foreign presence (Keller, 2009), selection effects from not observing firms after they exit (Haskel et al., 2007), and input endogeneity in the production function that could bias the estimates of input elasticities (Castellani & Zanfei, 2006; Levinsohn & Petrin, 2003; Marschak & Andrews, 1944). Identification challenges in incomplete datasets due to measurement error and selection problems have not been formally examined. Haskel et al. (2007) and Keller and Yeaple (2009) do discuss measurement error, but not as an identification problem. Moreover, they take

the view that measurement error always biases the spillover effect estimates downward, and that instrumental variable regression overcomes measurement error issues. We will show later in the paper that neither is necessarily true. Also, while selection issues discussed in the literature pertain to survival biases arising from not being able to observe firms after they exit, our selection issue pertains to completely censored data on small firms, which, we will argue, poses very difficult identification challenges.

³The Prowess database does contain other firms as well, but as Marin and Sasidharan (2010: 1231) themselves say, these tend to be large public sector enterprises.

⁴In fact, Koopmans (1949), who introduced the term “identification” into the econometric literature, puts it this way: “Statistical inference from observations to economic behaviour parameters can be made in two steps: inference from the observations to the parameters of the assumed joint distribution of the observations, and inference from that distribution to the parameters of the structural equations describing economic behaviour. The latter problem of inference (is) described by the term identification problem” (125). So, for identification, the assumed joint distribution of observations (i.e., the theoretical model) should adequately reflect the set of structural equations that describe the true spillover process.

⁵If we express the relationship between mismeasured foreign presence (fp_i) and true foreign presence (FP_i) as a linear function such as: $\gamma_0 + \gamma_1 FP_i + \xi_i$, where ξ_i is an error term independent of FP_i , and has 0 mean and constant variance, then what the estimated spillover effect θ_1 is really identifying is a term $\lambda\beta_1$, where $\lambda = \gamma_1 \text{Var}(FP_i) / [\gamma_1^2 \text{Var}(FP_i) + \text{Var}(U_i)]$. In this case, it is possible for the estimated effect θ_1 to overestimate the true effect β_1 as long as $\gamma_1 \neq 1$ (Carroll et al., 2006: 47).

⁶Secondary datasets, as we have seen, tend to be biased towards publicly listed firms, and are thus likely to miss privately held foreign and domestic firms. Missing foreign firms can lead to measurement error, as we have seen in the previous section, but it should not raise selection problems. This is because the spillover specification is usually estimated only on the domestic firms in the dataset. Missing (privately held) domestic firms, on the other hand, can lead to selection problems if the missing domestic firms also tend to be the smaller ones in the economy.

⁷We thank an anonymous reviewer for making this point.

⁸The Chinese dataset, as we have indicated before, does have a size bias. However, this is not a problem, given that the objective of our simulation is not to

assess FDI spillovers in China. Instead, we want to examine what happens when, assuming the Chinese dataset is the full population, spillover effects are estimated from subsamples of the dataset that are systematically incomplete. If our objective was to estimate spillover effects for China, then, indeed, the Chinese dataset might be inappropriate, given its size bias.

⁹Ideally, we should examine the effect of missing privately held firms and small firms, because these mirror the specific ways in which secondary datasets and manufacturing census surveys that have been used in the past are biased. But from the Chinese NBS data that we had access to, although we had data on firm size, we did not know whether each firm was publicly listed or not. Hence, as the next best alternative, we examined the effects on estimated spillover effects of missing foreign firms. The logic is that if datasets are biased towards public firms, then it is quite likely that foreign wholly owned ventures are missing. So the case of missing foreign firms we examine is a subset of the bigger problem of datasets being biased towards publicly listed firms.

¹⁰We initially assume a true spillover effect of +2, but in later robustness tests replace this with several other positive and negative values drawn from normal as well as alternative underlying distributions.

¹¹If we wanted to allow for selection effects as well, we would use different distributions for different size categories of firms. We would assign spillover effects as draws from a distribution with a higher mean, say 3, for large firms that are better able to absorb spillovers, and from a normal distribution with smaller mean, say 0, for small firms that lack absorptive capacity. In this particular simulation, however, we wanted to keep effects of selection effects constant, and hence drew spillover effect values for all firms from the same distribution.

¹²We calculated the dependent variable, productivity of domestic firms, as follows. We first estimated a Cobb–Douglas production function using the Levinsohn and Petrin (2003) routine for each of the industries in our sample, and obtained industry-specific coefficients of capital and labor. Then, for each domestic firm, we calculated the residual, that is, the difference between the firm’s actual output and its expected output, given its usage of capital and labor. This residual, as in most prior FDI spillover studies, is our measure of total factor productivity (Arnold, 2005).

¹³We thank an anonymous reviewer for suggesting that we run these simulations.

¹⁴These results pertain to the asymptotic, large sample properties of this estimator. In finite samples,

there might still be some bias even, though the estimator is consistent in infinite samples. This is called the “finite-sample bias of IV” (Cameron & Trivedi, 2009: 176).

¹⁵We thank an anonymous reviewer for pointing out that it is incorrect for us to claim that a specific

weighting scheme that worked in our dataset will work equally well for scholars using other datasets. Since the extent of size bias varies from dataset to dataset, the only way to know which weight to use, we think, is to compare them in model selection tests, as we describe in our empirical demonstration.

REFERENCES

- Acemoglu, D., & Dell, M. 2010. Productivity differences between and within countries. *American Economic Journal: Macroeconomics*, 2(1): 169–188.
- Aiken, L. S., & West, S. G. 1991. *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications.
- Aitken, B. J., & Harrison, A. E. 1999. Do domestic firms benefit from direct foreign investment? Evidence from Venezuela. *The American Economic Review*, 89(3): 605–618.
- Altomonte, C., & Pennings, E. 2009. Domestic plant productivity and incremental spillovers from foreign direct investment. *Journal of International Business Studies*, 40(7): 1131–1148.
- Angrist, J., & Krueger, A. B. 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4): 69–85.
- Angrist, J., & Pischke, J.-S. 2009. *Mostly harmless econometrics*. Princeton, NJ: Princeton University Press.
- Arnold, J. M. 2005. *Productivity estimation at the plant level: A practical guide*. Unpublished manuscript. Available at <http://www.jensarnold.de>.
- Bae, J., & Coo, J. 2008. Information loss, knowledge transfer cost and the value of social relations. *Strategic Organization*, 6(3): 227–258.
- Baily, M. N., & Solow, R. M. 2001. International productivity comparisons built from the firm level. *Journal of Economic Perspectives*, 15(3): 151–172.
- Banga, R. 2006. The export-diversifying impact of Japanese and US foreign direct investments in the Indian manufacturing sector. *Journal of International Business Studies*, 37(4): 558–568.
- Bartelsman, E. J., Haltiwanger, J. C., & Scarpetta, S. 2013. Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1): 305–334.
- Becker, R. A., & Gray, W. B. 2009. NBER-CES Manufacturing Industry Database. Available from <http://www.nber.org/data/nbprod2005.html>, accessed 9 February 2011.
- Bernard, A. B., & Jones, C. I. 1996. Comparing apples to oranges: Productivity convergence and measurement across industries and countries. *American Economic Review*, 86(5): 1216–1238.
- Blalock, G., & Simon, D. H. 2009. Do all firms benefit equally from downstream FDI? The moderating effect of local suppliers’ capabilities on productivity gains. *Journal of International Business Studies*, 40(7): 1095–1112.
- Bound, J., Brown, C., & Mathiowetz, N. 2001. Measurement error in survey data. In J. J. Heckman, & E. Leamer (Eds), *Handbook of econometrics*, 3705–3843. Amsterdam, The Netherlands: Elsevier.
- Buckley, P. J., & Casson, M. 1976. *The future of the multinational enterprise*. London: Palgrave Macmillan.
- Cameron, A. C., & Trivedi, P. K. 2009. *Microeconometrics using Stata*. College Station, TX: Stata Press.
- Cantner, U., & Pyka, A. 1998. Absorbing technological spillovers: Simulations in an evolutionary framework. *Industrial and Corporate Change*, 7(2): 369–397.
- Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. 2006. *Measurement error in nonlinear models: A modern perspective*. London, UK: Chapman & Hall.
- Castellani, D., & Zanfei, A. 2006. *Multinational firms, innovation and productivity*. Cheltenham: Edward Elgar.
- Castellani, D., & Zanfei, A. 2007. Multinational companies and productivity spillovers: Is there a specification error? *Applied Economic Letters*, 14(14): 1047–1051.
- Caves, R. E. 1996. *Multinational enterprise and economic analysis*. Cambridge, UK: Cambridge University Press.
- Centola, D., & Macy, M. 2007. Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3): 702–734.
- Costello, D. M. 1993. A cross-country, cross-industry comparison of productivity growth. *Journal of Political Economy*, 101(2): 207–222.
- Chang, S.-J., & Xu, D. 2008. Spillovers and competition among foreign and local firms in China. *Strategic Management Journal*, 29(5): 495–518.
- Chen, X., Hong, H., & Nekipelov, D. 2011. Nonlinear models of measurement errors. *Journal of Economic Literature*, 49(4): 901–937.
- Chung, W., Mitchell, W., & Yeung, B. 2003. Foreign direct investment and host country productivity: The American automotive component industry in the 1980s. *Journal of International Business Studies*, 34(2): 199–218.
- Dechow, P. M. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics*, 18(1): 3–42.
- Driffield, N., & Jindra, B. 2012. Challenging the production function approach to assess the developmental effects of FDI. *European Journal of Development Research*, 24(1): 32–37.
- Driffield, N., & Love, J. H. 2007. Linking FDI motivation and host economy productivity effects: Conceptual and empirical analysis. *Journal of International Business Studies*, 38(3): 460–473.
- Driffield, N., Love, J. H., & Menghinello, S. 2010. The multinational enterprise as a source of international knowledge flows: Direct evidence from Italy. *Journal of International Business Studies*, 41(2): 350–359.
- Dumouchel, W. H., & Duncan, G. J. 1983. Using sample survey weights in multiple regression analyses of stratified samples. *Journal of the American Statistical Association*, 78(383): 535–543.
- Eapen, A. 2012. Social structure and technology spillovers from foreign to domestic firms. *Journal of International Business Studies*, 43(3): 244–263.
- Fuller, W. A. 1987. *Measurement error models*. New York: John Wiley.
- Gorg, H., & Strobl, E. 2001. Multinational companies and productivity spillovers: A meta-analysis. *The Economic Journal*, 111(475): F723–F739.
- Griffith, R., Redding, S., & Simpson, H. 2003. *Productivity convergence and foreign ownership at the establishment level*. CEPR Discussion Paper No. 3765, London, UK, Centre for Economic Policy Research.
- Growiec, J., Pammolli, F., Riccaboni, M., & Stanley, H. E. 2008. On the size distribution of business firms. *Economics Letters*, 98(2): 207–212.
- Hall, R. E., & Jones, C. I. 1999. Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1): 83–116.
- Harrigan, J. 1999. Estimation of cross-country differences in industry production functions. *Journal of International Economics*, 47(2): 267–293.
- Harrison, J. R., Lin, Z., Carroll, G. R., & Carley, K. M. 2007. Simulation modeling in organizational and management research. *Academy of Management Review*, 32(4): 1229–1245.
- Haskel, J. E., Pereira, S., & Slaughter, M. J. 2007. Does inward foreign direct investment boost the productivity of domestic firms? *Review of Economics and Statistics*, 89(3): 482–496.

- Hausman, J. 2001. Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives*, 15(4): 57–67.
- Hennart, J.-F. 1982. *A theory of multinational enterprise*. Ann Arbor, MI: University of Michigan Press.
- Horowitz, J. L., & Manski, C. 1998. Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations. *Journal of Econometrics*, 84(1): 37–58.
- Hsieh, C.-T., & Klenow, P. J. 2009. Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4): 1403–1448.
- Hymer, S. H. 1976. *The international operations of national firms: A study of direct foreign investment*. Boston, MA: MIT Press.
- Javorcik, B. S. 2004. Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *The American Economic Review*, 94(3): 605–627.
- Keller, W. 2009. *International trade, foreign direct investment, and technology spillovers*. National Bureau of Economic Research Working Paper Series no. 15442.
- Keller, W., & Yeaple, S. 2009. Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States. *Review of Economics and Statistics*, 91(4): 821–831.
- King, R. G., & Levine, R. 1993. Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics*, 108(3): 717–737.
- Kipnis, V., Carroll, R. J., Freedman, L. S., & Li, L. 1999. Implications of a new dietary measurement error model for estimation of relative risk: Application to four calibration studies. *American Journal of Epidemiology*, 150(6): 642–651.
- Kish, L. 1990. Weighting: Why, when, and how? *Proceedings of the Survey Research Methods Section, American Statistical Association*, 121–130.
- Koopmans, T. C. 1949. Identification problems in economic model construction. *Econometrica*, 17(2): 125–144.
- Krasker, W. S., & Welsch, R. E. 1985. Resistant estimation for simultaneous-equations models using weighted instrumental variables. *Econometrica*, 53(6): 1475–1488.
- Levinsohn, J., & Petrin, A. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2): 317–342.
- Manski, C. F. 1995. *Identification problems in the social sciences*. Cambridge, MA: Harvard University Press.
- Marin, A., & Sasidharan, S. 2010. Heterogeneous MNC subsidiaries and technological spillovers: Explaining positive and negative effects in India. *Research Policy*, 39(9): 1227–1241.
- Marschak, J., & Andrews, W. 1944. Random simultaneous equations and the theory of production. *Econometrica*, 12(3/4): 143–205.
- Martin, X. 2013. Solving theoretical and empirical conundrums in international strategy research: Linking foreign entry mode choices and performance. *Journal of International Business Studies*, 44(1): 28–41.
- Meagher, K., & Rogers, M. 2004. Network density and R&D spillovers. *Journal of Economic Behavior & Organization*, 53(2): 237–260.
- Meyer, K. E., & Sinani, E. 2009. When and where does foreign direct investment generate positive spillovers? A meta-analysis. *Journal of International Business Studies*, 40(7): 1075–1094.
- Mooney, C. Z. 1997. *Monte Carlo simulation*. London, UK: Sage Publications.
- Murray, M. P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20(4): 111–132.
- OECD. 2010. STAN R&D: Research and Development Expenditure in Industry. STAN: OECD Structural Analysis Statistics (database). Available at <http://dx.doi.org/10.1787/data-00032-en>, accessed 9 February 2011.
- Oxley, J., Rivkin, J. W., & Ryall, M. D. 2010. The strategy research initiative: Recognizing and encouraging high-quality research in strategy. *Strategic Organization*, 8(4): 377–386.
- Pfefferman, D. 1993. The role of sampling weights when modeling survey data. *International Statistical Review*, 61(2): 317–337.
- Pfefferman, D. 1996. The use of sampling weights for survey data analysis. *Statistical Methods in Medical Research*, 5(3): 239–261.
- Porter, R. D. 1973. On the use of survey sample weights in the linear model. *Annals of Economic and Social Measurement*, 2(2): 140–157.
- Rajan, R. G., & Zingales, L. 1998. Financial dependence and growth. *American Economic Review*, 88(3): 559–586.
- Reeb, D., Sakakibara, M., & Mahmood, I. P. 2012. From the editors: Endogeneity in international business research. *Journal of International Business Studies*, 43(3): 211–218.
- Reichstein, T., & Jensen, M. B. 2005. Firm size and firm growth rate distributions? The case of Denmark. *Industrial and Corporate Change*, 14(6): 1145–1166.
- Shaver, M. Y. 1998. Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Science*, 44(4): 571–585.
- Simon, H. A., & Bonini, C. P. 1958. The size distribution of business firms. *American Economic Review*, 48(4): 607–617.
- Smeets, R., & de Vaal, A. 2011. *Knowledge diffusion from FDI and intellectual property rights*. CPB Discussion paper no. 168, CPB Netherlands Bureau for Economic Research, The Hague.
- Vissek, J. A. 2009. Consistency of the instrumental weighted variables. *Annals of the Institute of Statistical Mathematics*, 61(3): 543–578.
- Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2): 307–333.
- Zamborsky, P. 2012. Intangibles of host country effects of foreign direct investment. *Journal of International Business and Economics*, 12(2): 123–137.
- Zhang, Y., Li, H., Li, Y., & Zhou, L.-A. 2010. FDI spillovers in an emerging market: The role of foreign firms' country origin diversity and domestic firms' absorptive capacity. *Strategic Management Journal*, 31(9): 969–989.

About the Author

Alex Eapen is a senior lecturer at the Australian National University's Research School of Management. He earned his PhD from Tilburg University, and was a finalist for the 2007 Gunnar Hedlund doctoral dissertation prize. His research interests include the theory of the multinational enterprise (MNE), and the subsequent impact of MNE entry on domestic firms, for example, through technology spillovers.

Accepted by Mariko Sakakibara, Area Editor, 29 May 2013. This paper has been with the author for two revisions.