

The Perfect Match:
Assortative Matching in Mergers and Acquisitions.
PRELIMINARY.
PLEASE DO NOT CIRCULATE OR CITE.

Maria Guadalupe	Veronica Rappoport	Bernard Salanié	Catherine Thomas*
INSEAD	LSE	Columbia	LSE

November 4, 2021

Abstract

We interpret M&A deals in Western Europe during the 2000s as the equilibrium of a matching process and apply recently developed empirical methods to infer features of the merger surplus function that rationalizes the observed partner matching. We find revenue productivity is complementary to merger partner input scale, consistent with mergers allowing efficient reallocation of assets within firm boundaries. However, in our sample, complementarity in revenue productivities contributes significantly more to merger surplus, especially for mergers between firms in similar industries, suggesting mergers between productive firms generate gains beyond those derived from reallocating existing resources. For the subset of merged firms where post-merger data is also available, revenues, scale, productivity, and profits all exceeded counterfactual estimates in the first few years after the merger, and the observed productivity gains are positively correlated with the surplus estimated in the matching model.

*E-mail: maria.guadalupe@insead.edu, v.e.rappoport@lse.ac.uk, bsalanie@columbia.edu, c.m.thomas@lse.ac.uk. We are grateful to Alfred Galichon, Jon Vogel, and seminar participants at Sciences Po, ESCP, HBS, HEC, IFN, Nottingham, Yale, and LSE. We thank Vaidehi Paramasweran and Siqin Ding for excellent research assistance and Paul Horsler for helpful advice about data availability. All errors are our own.

1 Introduction

Merger and acquisition activity redraws firm boundaries, bringing different firms' assets and operations under joint management. The possible motives for these transactions vary, but have in common that the owners of each party anticipate being at least as well off post merger as they were before. In other words, the merger is expected to generate a surplus relative to the sum of the standalone values of each merging party. This paper estimates the sources of anticipated merger surplus using data on who merges with whom within a matching framework.

Our objective is to uncover the extent to which merger surplus arises from anticipated complementarities between the pre-merger characteristics of the merging firms. Much of the existing empirical literature about merger motives characterizes firms in one dimension of heterogeneity, such as market to book ratios, productivity, or profitability, whereas theoretical mechanisms are often based on complementarities between different dimensions, such as productivity and capital (Jovanovic and Rousseau (2002)) or mobile and immobile capabilities, (Nocke and Yeaple (2007)). Our estimation approach is based on the idea that each firm's merger partner choice reflects their trading off the values of different complementarities between each potential partner's characteristics and their own. As we will show, there is strong evidence that the mergers in our data do reflect matching on more than one dimension. Furthermore, the contribution to merger surplus of various complementarities differs for cross-border versus domestic mergers, and for cross-industry versus within-industry mergers

For each merging firm in our sample, we decompose revenues in the year before the merger into input scale and revenue productivity. Input scale is a weighted sum of the firm's assets and employment, where the weights are estimated production function coefficients. We interpret revenue productivity as a measure of the internal capabilities with which the firm uses inputs, and, as the revenue residual, it reflects efficiency, managerial skill, as well as all intangible inputs, such as brands. These two continuous dimensions of firm heterogeneity allow us to identify the relative importance of three different components of anticipated merger surplus: the surplus arising from complementarities between partners' revenue productivities, between their input scales, and the cross-complementarity between revenue productivity and the merger partner's input scale.

Our data about merging firms and mergers come from the Bureau van Dijk database Zephyr. We focus on bilateral deals that were completed during the nine-year period 2000 to 2008 among

firms in five large Western European countries, a time and place where the shared EU institutional context meant that capital could flow relatively freely across country borders.¹

The results from the empirical matching model show that firms aim to exploit different sources of complementarity, but these vary in their importance. First, revenue productivity is complementary to the merger partner’s input scale in generating merger surplus. However, revenue productivity contributes much more to anticipated merger surplus through complementarity with the revenue productivity of the merger partner. At the mean, the complementarity between revenue productivities is three time larger than the cross complementarity between revenue productivity and the input scale of the partner firms. Finally, we also show that complementarity between the input scales of the merging firms does not contribute significantly to anticipated merger gains.

These findings provide an assessment of the empirical relevance of the most prominent class of theories of mergers, in which surplus arises from reallocating control of existing firm resources within expanded firm boundaries. For example, [Jovanovic and Rousseau \(2002\)](#) explain mergers as the opportunity to leverage one firm’s superior technology over the combined capital of the merged firm, and [Nocke and Yeaple \(2008\)](#) model merger surplus as arising from the complementarity between a headquarters’ assets and its merger partner’s production assets. In these theories, the merged firm’s production function has some common features: there is a cross-complementarity between different pre-merger firm assets, such as technology and capital, and also a substitutability between at least one same pre-merger asset, such as firm technology, because the superior contribution from one firm acts as a public good within expanded firm boundaries and renders the inferior contribution from the merger partner redundant.

In our data, while we find evidence of a cross-complementarity in revenue productivity and scale, consistent with reallocation merger motives, there is no evidence consistent with the predicted substitutability of firm capabilities within firm boundaries in our data. In fact the opposite is true. The positive assortative matching in revenue productivities we find is, instead, consistent with mergers creating value by generating new shared capabilities within enlarged firm boundaries.

The complementarity in revenue productivities is analogous to the finding of ‘like buying like’ along one dimension of firm heterogeneity in [Rhodes-Kropf and Robinson \(2008\)](#), who show merg-

¹This feature is important because the matching model we develop and estimate assumes away any frictions, and interprets the observed mergers as the static market equilibrium.

ing firms in their US sample have similar market-to-book ratios.² One other empirical precedent reminiscent of this complementarity is found in the context of FDI via cross-border mergers, where [Bilir and Morales \(2020\)](#) show that new technologies in parent-firm headquarters increase the productivity of foreign affiliates by a greater extent if the affiliate is also innovative.³ The anticipated complementarity in revenue productivities that we find establishes more generally that the value of the expansion via M&A depends on the existing capabilities of both firms involved.⁴

Several recent papers apply the estimation methods developed in [Fox \(2018\)](#) to data about who merges with whom to estimate sources of merger value in different data samples, and this method also allows for multiple dimensions of firm heterogeneity. [Mindruta, Moeen, and Agarwal \(2016\)](#) study alliances in the bio-pharma industry and find evidence of complementarities between firm size and also between upstream research capabilities. These authors discuss how knowledge recombination creates more value when firms with better research capabilities match with each other. Their finding therefore offers one potential reason for the complementarity in productivities that we observe in our sample.

After showing the relative strength of complementarities on average in the sample, we allow the terms in the surplus function to differ with whether or not the merging firms are in the same broadly defined industry, and with whether they are located in the same country. The complementarity between firm revenue productivities is shown to contribute around twice as much to the anticipated merger gains when merging firms are in a similar industry rather than in different industries, both within and across country borders. The complementarity between revenue productivity and input scale is present both within and across industries, but contributes significantly only to domestic mergers.

²[Braguinsky, Ohyama, Okazaki, and Syverson \(2015\)](#) find merging firms have similar measures of quantity-based productivity in their study of Japanese spinning. In an unpublished paper, [Ozcan \(2015\)](#) finds that the merging firms tend to have similar levels of patenting quality.

³For the most part, the literature on multinational firm expansion does not typically distinguish greenfield investment from M&A (see, for example, [Ramondo and Rodriguez-Clare \(2013\)](#), [McGrattan and Prescott \(2009\)](#), [Helpman, Melitz, and Yeaple \(2004\)](#)). In these models, technology is transmitted either freely or at an exogenous discount rate, and so the capabilities of the least productive firm in the merger are redundant.

⁴This finding relates to [Cohen and Levinthal \(1990\)](#), who describe how a firm's prior knowledge determines its ability to recognize, assimilate, and apply new information to generate value, and define this ability as a firm's absorptive capacity. It is perhaps not too much of a leap of logic to imagine that this phenomenon exists during post-merger integration, and that similar starting points facilitate productivity-increasing information exchange within a newly-merged firm.

David (2020) combines data on who merges with whom and who opts to remain unmerged among US firms in a dynamic model. While more profitable firms select into merging in general, he finds evidence that mergers allow resources to be efficiently reallocated within firm boundaries in that acquirers tend to be more profitable than their targets. Although it is not the largest motive for merging in our sample, there is anticipated merger surplus from reallocating joint firm assets to more productive control, consistent with the mechanism described in Jovanovic and Rousseau (2002) for the domestic mergers we observe. Input scale does not contribute significantly to anticipated merger gains in the cross-border mergers in our sample, either via a complementarity with partner scale or with partner revenue productivity.

We model the market for corporate control as a matching problem à la Becker (1973), where each possible match between two firms generates a merger surplus that can be freely transferred between matched partners. The distributions of firms' characteristics, along with the distributions of preferences for those characteristics, determine how the surplus is shared in any realized match.⁵ In a pioneering contribution, Choo and Siow (2006) introduced a tractable model that allows for matching on unobservables as well as observables. We maintain elements of their approach and also rely on the work of Dupuy and Galichon (2014), which incorporated observed continuous characteristics in this model. Our estimation method builds on Galichon and Salanié (2021)'s moment matching, and on their iterative proportional fitting procedure (IPFP) algorithm, to compute the equilibrium.

The main advantage of our structural modeling empirical approach is that it controls for the marginal distributions of each observable firm characteristic when estimating the parameters of the surplus function. This property captures the idea that firms are competing for merger partners with characteristics that complement their own. It is the marginal distributions that serve to attach a value to each merging firm outside each possible match and ensure the emergence of an equilibrium. It would not have been possible to infer sources of merger surplus simply by examining the variance-covariance matrix inside observed matches.⁶

⁵The crucial assumption of freely transferable utility is a common assumption in the marriage market matching literature, and is, perhaps, more natural, in the M&A setting where monetary transfers between merging partners, in the form of acquisition prices, are intensely negotiated pre-merger.

⁶Sections 2 and 3 expand on this key advantage of our approach. As a simple illustration, because revenue productivity and input scale are negatively correlated within firm in our sample, more frequent merging between productive firms leads to a negative correlation between productivity and input scale within merger. However, it is

Our application of these methods to the study of mergers differs from the classical marriage problem of [Becker \(1973\)](#) in two important ways. First, we do not consider data on non-merging firms. This implies that we cannot quantify the absolute value of mergers relative to remaining standalone. This does not affect our ability to recover estimates of the relative value of different mergers, however. As our specification builds on [Choo and Siow \(2006\)](#)’s logit structure, it inherits a version of the usual independence of irrelevant alternatives property.⁷ As a consequence, the relative probabilities of different mergers are unaffected by the presence of the outside option of not merging.

The second important difference is the absence of a variable analogous to gender, which plays a crucial role in marriage markets. Because a large majority of marriages consist of two partners of different genders, most empirical work on marriage models matching as *bipartite*, in the jargon of the field.⁸ Our work instead conceives of merging firms matching as *roommates*.⁹ In particular, we do not use the “acquirer” and “target” labels in our data as a pre-existing distinction. We take the stance that the identity of the acquirer, and the target, in a merger is an endogenous outcome, related to the equilibrium transfers required between the owners of the firms taking part in the merger.¹⁰

An important welfare-relevant question is the extent to which our results are consistent with firms merging to increase their market power. The complementarity between revenue productivities in our main results could arise from the anticipated ability of mergers between high-price firms to further increase price. Nonetheless, if this were a primary motive for the mergers in our data, we would also expect to see large complementarities between merging firms’ sizes.¹¹ In particular,

not the case that more merger surplus is created when productive firms merge with small-scale partners. In fact, our full results show the opposite to be true.

⁷See [Galichon and Salanié \(2019\)](#)

⁸For a recent exception, see [Ciscato, Galichon, and Goussé \(2020\)](#).

⁹The name comes from [Gale and Shapley \(1962\)](#)’s seminal paper: “An even number of boys wish to divide up into pairs of roommates” (Example 3, page 12).

¹⁰As noted in other studies, the firms that Bureau van Dijk labels acquirers do differ, on average, from those that are labeled targets, being larger on average. The model generates additional predictions related to equilibrium transfers between merging firms as a function of firm characteristics. Relating predicted transfers to acquisitions prices is a topic for future research.

¹¹In industry-level models of this phenomenon, the equilibrium price rise post-merger is increasing in the pre-merger market shares (Williamson, 1968, Perry and Porter, 1985, Nocke and Whinston, also [Deneckere and Davidson, 1985](#) and [Farrell and Shapiro, 1990](#)). See [Asker and Nocke \(2021\)](#) for a discussion of why mergers above a certain size

complementarity in size would likely be greater when firms were closer competitors, in the same country and same industry. Our findings show the opposite. The very limited complementarity in input scales is actually smallest for mergers between firms in the same country, both across and within industries.¹² In addition, there is no evidence that the complementarity in firm revenue productivities increases in the scale of the merged firm. These findings hence suggest that the firms in our sample did not merge primarily in order to increase their joint market power and raise prices. It is worth noting, though, that the deals we observe took place within the regulatory remit of the relevant competition authorities. Any proposed mergers that were clearly anticompetitive may well have been blocked preemptively.

Finally, we also examine the post-merger performance of the firms in our data. For around one third of the mergers, the firm-level post-merger revenues, profits, employment, and total assets for the first five years after the merger are available in the Bureau van Dijk database ORBIS. We construct a hypothetical counterfactual estimate for each of these variables had the firms not merged using data on the median growth rate over the relevant time period of all firms in the relevant country and industry. Actual revenues and input scale are around 10% larger on average than in the non-merging counterfactual for each of the first five years after the merger. The observed profit growth tends to be positive over the same time period while counterfactual profit growth is negative, related to the fact that the post-merger period coincides with the aftermath of the financial crisis in Europe after 2007. These findings are consistent with firm owners having made sensible decisions by merging. However, there is large variation in all of these post-merger performance measures.

The post-merger performance data also offer some support to the estimates of the matching model. In the five years after merging, revenue productivity growth relative to the constructed counterfactual is shown to be significantly positively correlated with the estimates of expected merger surplus delivered by the matching model, even after controlling for the growth that can be attributed to all firm-level factors. A one standard deviation increase in the estimated merger surplus is associated with revenue productivity growth in the merged firm that is 3% higher than

threshold trigger antitrust concerns.

¹²The absence of a complementarity in merging firms' input scales also suggests that firms in this sample were not merging purely for the sake of empire building, serving the objectives of the managers without increasing the value of the firm (see [Roll, 1986](#), [Jensen, 1986](#), and [Shleifer and Vishny, 1989](#)).

in the estimated counterfactual.¹³

The rest of the paper proceeds as follows. Section 2 describes the empirical setting and the data used. Section 3 presents the matching model, the estimation framework, the identification strategy, and the computation approach. The more technical details are given in the Appendix. Section 4 presents the main results describing the estimated merger surplus function. Section 5 compares the estimated anticipated merger gains with the merger gains observed in a subsample of firms. Section 6 concludes.

2 Empirical Setting and Data

We study 4,447 two-firm mergers that took place over the years 2000 to 2008 between firms headquartered in one of five large Western European countries. The data come from the Bureau van Dijk database Zephyr and include merger characteristics as well as firm characteristics around the time of the merger. The five countries are France, Germany, Italy, Spain, and the UK, and the time period is the sustained growth period preceding the financial crisis of 2008. A practical advantage of these data (relative to commonly-used datasets such as Compustat in the USA), is that non-listed firms are included due to European financial reporting requirements.

We limit the sample to deals that led to a change in the identity of the controlling owner of one of the merging parties, and to those where data to measure firm productivity and input scale are available for each of the firms. This means that in all five countries, our data tend to exclude smaller firms, where less data is reported. Zephyr contains each firm’s primary four-digit industry code according to the NACE Revision 2 classification. We include both manufacturing and services firms in the sample, and group all firms into one of five broadly-defined industry groups.¹⁴

The matching model that we describe in more detail in the following section is estimated from variation in the relative frequency of observed mergers between firms with particular observable

¹³Stiebale (2016) shows that innovation, as measured by patent activity, increases after merging, relative to the sum of stand-alone levels, which offers a reason for matching on productivities that resembles the finding in Mindruta, Moeen, and Agarwal (2016). In contrast, in their study of mergers of US firms, Blonigen and Pierce (2016) find no effect of mergers on productivity and instead show large increases in market power and markups for the combined entities.

¹⁴We exclude firms in financial service or the public sector, i.e. banks, insurance companies, public administration and defence. We also exclude firms in the extraction of crude petroleum and natural gas, code 6, because the industry-specific productivity coefficients were inconsistent with some of the structure imposed in later parts of the paper.

characteristics. The firms are characterized by the two discrete dimensions country and industry group, and the two continuous characteristics input scale and revenue productivity.

We decompose variation in firm revenues into variation in the use of labor and capital inputs and variation in revenue productivity. We use the data on revenues, the number of firm employees and total assets the year before the merger to estimate production functions at the two-digit NACE code level. For each industry code, the log of revenues of firm j at time t , r_{jt} , is:

$$r_{jt} = z_{jt} + (\alpha_L l_{jt} + \alpha_K k_{jt}), \quad (1)$$

where l_{jt} is the log of firm i 's total number of employees and k_{jt} is the log of firm i 's total assets. Both revenues and total assets are deflated with sector-specific PPIs. Given the available data, we use material costs as an input that is correlated with per-period employment in the control function approach following [Akerberg, Caves, and Frazer \(2015\)](#). The estimated coefficients α_L and α_K are industry specific, and are calculated using data on a large number of firms in each industry from another Bureau van Dijk dataset, ORBIS.

Table 1 summarizes the mergers in our data along the two discrete characteristics and describes the probabilities of selecting into domestic and within-industry group mergers for firms by country and by industry group. Domestic, within-industry group mergers are 65% of the total and domestic cross-industry mergers make up 22%. Cross-border, within-industry group mergers are 9% of total deals, and, finally, 4% are both cross-border and cross-industry group. These statistics indicate that domestic and within-industry group mergers create more value for the merging firms, all else equal. The total numbers are in the first row and first column of each panel. The country-level rows sum to 100% across country-partner columns in panel A, and the industry group-level rows sum to 100% across industry-group columns in panel B. In panel A, the large diagonal elements, presented in bold, reflect the large share of domestic mergers, that is, the within-country endogamy. Panel B shows the extent of within-industry group endogamy.

Table 2 shows descriptive statistics about the firms participating in the four different types of mergers (international or domestic, across or within industry group). Cross-border mergers involve firms with larger input scales, and, of these mergers, those that are cross-industry have higher mean revenue productivity. For domestic mergers, those that are also within industry are slightly larger

Table 1: Endogamy on Discrete Characteristics

Panel A: By Country						
	N Firms	Merging with firm in:				
		Germany	Spain	France	UK	Italy
Total	8894	628	2029	1914	3312	1011
Germany	628	72	4	8	11	5
Spain	2029	2	90	3	3	2
France	1914	2	4	85	6	3
UK	3312	2	2	4	91	1
Italy	1011	3	4	6	4	83

Panel B: By Industry Group						
	N Firms	Merging with firm in:				
		Basic Manuf.	Sophisticated Manuf.	Other	High-Tech Serv.	Non HT Serv.
Total	8894	828	1839	2321	1177	2729
Basic Manuf.	828	74	7	14	0	5
Sophisticated Manuf.	1839	3	75	12	3	7
Other	2321	5	10	72	3	10
High-Tech Serv.	1177	0	5	5	72	18
Non-High-Tech Serv.	2729	2	4	9	8	77

Table 2: Size and Productivity by Type of Merger

	N mergers	Input Scale		TFPr	
		Mean	s.d.	Mean	s.d.
Cross-border, cross-industry group	160	8.25	2.47	2.70	1.14
Domestic, cross-industry group	981	7.55	2.30	2.57	1.02
Cross-border, within-industry group	403	8.56	2.70	2.51	1.01
Domestic, within-industry group	2,903	7.58	2.20	2.58	1.05

scale and slightly more productive, however these groups of firms look similar overall.

Next, Table 3 describes the mergers along the two observable continuous variables used in the matching model—input scale and revenue productivity. To do this, we select one firm from each merger at random to be firm x , and the other firm is firm y . Thus, we have 4,447 observations, corresponding to the number of deals in the data. Columns 1 and 2 regress the input scale of a firm on its merger partner’s scale and revenue productivity, and, while the coefficient on input scale is significant in the first column, it is insignificant once we control for country/industry fixed effects (column 2). From this, we would (mistakenly, as we will see) infer that matching on these characteristics is random. In turn, columns 3 and 4 regress revenue productivity of a firm on the partner firm’s characteristics. Here we find that both coefficients are significant, and that the largest is between revenue productivities. It is ten times larger than the coefficient relating revenue productivity and scale. Noting that we selected firm x and firm y at random within each merger pair, it is hard to learn anything about the correlations of firm characteristics within mergers from columns 2 and 4.

Chiappori, Orefice, and Quintana-Domeque (2012) show how to use these regressions as a way to test whether the observable characteristics, of which we have two that are continuous, can be summarized by a lower-dimensional index, which in our case would be a one-dimensional scale. The intuition is that if the joint surplus of a merger depended only on some linear combination of scale and productivity, then the ratios of the scale and productivity coefficients in these regressions should be the same whether the dependent variable is scale or productivity. It is clear from Table 3 that the matching patterns in the data—who merges with whom—cannot be fully described by models where sorting is based on homogeneous preferences over a single source of firm heterogeneity.

Table 3: Regressions of Firms' Characteristics on those of their Merger Partners.

	scale_Y (1)	scale_Y (2)	TFPr_Y (3)	TFPr_Y (4)
scale_X	0.051*** (0.015)	0.015 (0.015)	0.025*** (0.007)	0.019*** (0.007)
TFPr_X	0.018 (0.034)	-0.016 (0.034)	0.243*** (0.015)	0.199*** (0.015)
Constant	7.278*** (0.164)	8.316*** (0.263)	1.753*** (0.072)	2.130*** (0.115)
Country-Ind FE	No	Yes	No	Yes
Observations	4447	4447	4447	4447
R^2	0.003	0.037	0.057	0.093

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Consider columns (1) and (3). The ratio of the coefficients of scale and productivity is 2.8 in column (1) and 0.1 in column (3). These coefficients do not seem to be consistent with there being a one-dimensional index that summarizes the attractiveness of a merger partner. A χ -squared test rejects equality of the ratios decisively. The equivalent test for the ratios of coefficients in columns (2) and (4) is similarly rejected.¹⁵ We, hence, turn to our model of matching in multiple dimensions to provide richer empirical insight about firms' motives.

3 Empirical Framework

We start by describing the matching model and our approach to identification and estimation as if we observed firms that choose to remain standalone. Subsection 3.4 explains how we adapt our methods to the data we use, which only reports realized mergers.

3.1 Firm value and merger surplus

Each firm i consists of a bundle of observable and unobservable characteristics that are summarized in a triple $\{a, x, \tilde{\epsilon}_i\}$. The first element a is an ordered pair, indicating the country where the firm is located and the industry group of its primary activity. The second element x is a two-dimensional

¹⁵While this does not yield a valid test in general (see Chiappori, Orefice, and Quintana-Domeque (2020)), we show in Appendix A.2 that it does apply to the variant of the Choo and Siow (2006) model that we develop.

vector of continuous variables $x = (z_x, \Omega_x)$, where z_x is the revenue productivity of the firm and Ω_x is a measure of its size. Our main results proxy size by input scale. All market participants and researchers observe a and x . The third element of the triple is a random variable that describes how firm i differs from other firms on dimensions that are observable to other market participants but not to the researchers. The value of the firm $\Pi_a(x, \tilde{\epsilon}_i)$ depends on all three elements of the triple.

When firms i and j merge, the value of the new merged entity depends on the pre-merger characteristics of the two firms. If firm j is described by $\{b, y, \tilde{\epsilon}_j\}$, then the value of the merged firm is $\Pi_{ab}(x, y, \tilde{\epsilon}_i, \tilde{\epsilon}_j)$. The merger gains, denoted $\tilde{\Phi}_{ab}(x, y, \tilde{\epsilon}_i, \tilde{\epsilon}_j)$, are therefore

$$\tilde{\Phi}_{ab}(i, j) = \Pi_{ab}(x, y, \tilde{\epsilon}_i, \tilde{\epsilon}_j) - \Pi_a(x, \tilde{\epsilon}_i) - \Pi_b(y, \tilde{\epsilon}_j). \quad (2)$$

That is, the merger generates positive gains if the new entity is more valuable than the sum of the pre-merger firm values. At this point, we impose some structure on how the unobservable (to us) characteristics of firms i and j , $\tilde{\epsilon}_i, \tilde{\epsilon}_j$ enter the merger surplus function. We maintain the assumption made by [Choo and Siow \(2006\)](#) that the joint surplus function is independent of any interactions between the unobserved characteristics $\tilde{\epsilon}_i$ and $\tilde{\epsilon}_j$ conditional on the observed characteristics $\{a, x\}$ and $\{b, y\}$. This is known as the *separability* assumption. More precisely, let $I(a, x) = \{i \text{ s.t. } a_i = a \text{ and } x_i = x\}$ denote the set of firms with observable characteristics $\{a, x\}$. We assume that there exists a matrix of functions Φ and random vectors of functions ϵ^i such that in a match between firms i and j , the joint surplus is

$$\tilde{\Phi}(i, j) = \Phi_{ab}(x, y) + \epsilon_b^i(y) + \epsilon_a^j(x). \quad (3)$$

We require $\Phi_{ab}(x, y) = \Phi_{ba}(x, y)$, so that $\tilde{\Phi}(i, j) = \tilde{\Phi}(j, i)$. We denote $\tilde{\Phi}(i, 0) = \epsilon_0^i$ for a firm $i \in I(a, x)$ that does not participate in a merger.

As [Galichon and Salanié \(2021\)](#) note, while restrictive, the separability assumption does allow for a limited form of “matching on unobservables” in that it allows surplus to arise from interactions between each firm’s unobservable and the observables of its partner. It only rules out interactions between the unobservable characteristics of the two partners.¹⁶

¹⁶By way of illustration, the education of CEOs is not recorded in our data. The separability assumption allows

3.2 Matching Equilibrium

The equilibrium consists of a feasible, stable matching, and a set of firm payoffs. A *matching* is a specification of who merges with whom; we denote $\mu(i, j)$ the number of matches between i and j . *Feasibility* requires that each firm be matched to one partner at most (some firms may not merge): $\mu(i, 0) + \sum_j \mu(i, j) = 1$. *Stability* requires that (i) no merging firm would prefer to opt out of the merger and (ii) no two firms would both prefer merging with each other to their observed outcome.

In equilibrium, any pair of merging firms must split the joint surplus. Each firm seeks the partner with whom its *net* surplus will be highest: the equilibrium payoff u_i of any firm i solves

$$u_i = \max_j (\tilde{\Phi}(i, j) - u_j),$$

where the maximization runs over all potential partners and the option not to merge (for which $\Phi(i, 0) = 0$ and $u_0 = 0$). Firm i is matched in equilibrium to a firm j that achieves this maximum payoff in the equilibrium matching (or remains standalone if the optimum is achieved at $j = 0$).

We adapt the approach in Galichon and Salanié (2021)'s marriage market setting to fit the roommate problem. We start by substituting equation (3) into the equilibrium payoff of firm $i \in I(a, x)$:

$$\begin{aligned} u_i &= \max_{b, y} \left(\Phi_{ab}(x, y) + \varepsilon_b^i(y) + \max_{j \in I(b, y)} (\varepsilon_a^j(x) - u_j) \right) \\ &= \max_{b, y} \left(\Phi_{ab}(x, y) + \varepsilon_b^i(y) - U_{ba}(y, x) \right), \end{aligned}$$

where we denote $U_{ba}(y, x) = \min_{j \in I(b, y)} (u_j - \varepsilon_a^j(x))$.

Suppose that there are matches between $I(a, x)$ and $I(b, y)$ in equilibrium. For each of these (i, j) matches, the joint surplus must be split between the two firms: $u_i + u_j = \tilde{\Phi}(i, j)$. This yields

$$U_{ab}(x, y) + U_{ba}(y, x) = \Phi_{ab}(x, y). \quad (4)$$

While $\Phi_{ab}(x, y)$ is symmetric, $U_{ba}(y, x)$ will normally differ from $U_{ab}(x, y)$: in any equilibrium

the surplus to be greater for a merger of a firm whose CEO studied at the LSE if the firm merges with any UK services firm. However, it rules out the possibility that the merger surplus is higher, controlling for all observable characteristics, when the CEOs of both merging firms attended the LSE.

match, the joint surplus from observables is split between merging firms according to the value the market assigns to each of their observable characteristics. As always, this reflects both the relative scarcity and the relative desirability of these observables.

A simple extension of the argument in [Shapley and Shubik \(1971\)](#) shows that the equilibrium matching $\mu(i, j)$ maximizes the total joint surplus over the set of feasible matchings $\mathcal{M}$:

$$\mathcal{W} = \max_{\mu \in \mathcal{M}} \left(\sum_{i,j} \mu(i, j) \tilde{\Phi}(i, j) + \sum_i \mu(i, 0) \tilde{\Phi}(i, 0) \right) \quad (5)$$

where feasibility imposes that $\mu \geq 0$; μ is symmetric; and for each i , $\mu(i, 0) + \sum_j \mu(i, j) = 1$. The symmetry constraint is specific to the roommate problem. While it may lead to the non-existence of a stable matching in general, [Chiappori, Galichon, and Salanié \(2019\)](#) show that with large populations, there always is an equilibrium.

Our primary object of interest is the relationship between merger surplus and the observable pre-merger characteristics of the merging firms, as summarized in the function $\Phi_{ab}(x, y)$ in equation (3). We estimate features of this relationship by assuming that the observed matching $\hat{\mu}$ is the equilibrium matching, and finding the functions $U_{ab}(x, y)$ such that the surplus function Φ defined by equation (3) comes closest to generating $\hat{\mu}$ in equilibrium.

We begin by partitioning firms based on observables. Let N_a be the number of firms with discrete characteristics a , and $f(x|a)$ be density of the continuous characteristics of these firms. The number of mergers between firms with discrete characteristics a and b is denoted μ_{ab} , and $\mu(x, y|a, b)$ is the joint density of these merging firms' continuous characteristics. We decompose the equilibrium matching $\mu_{ab}(x, y)$ as

$$\mu_{ab}(x, y) = \mu_{ab} \times \mu(x, y|a, b)$$

where $\sum_b \mu_{ab} = N_a$ and the following symmetry constraints apply: $\mu_{ab} = \mu_{ba}$ and $\mu(x, y|a, b) = \mu(y, x|b, a)$. A matching is feasible iff for all (a, x) :

$$\sum_b \int \mu_{ab}(x, y) dy = N_a f(x|a). \quad (6)$$

We now choose distributions for the error terms in equation (3) in order to write down the choice probabilities for each firm i . Since x and y are continuous variables, we must depart from the standard apparatus of discrete choice models. We follow Dagsvik (1994) and Dupuy and Galichon (2014) and assume that

- each firm i selects a partner $(b, y, \varepsilon_b^i(y))$ from the points of a Poisson process with intensity $(1/M) \times dy \times \exp(-\varepsilon)d\varepsilon$, where M is the number of possible discrete characteristics combinations;
- these Poisson processes are independently and identically distributed across firms i .

The Poisson specification can be seen as an analog of the multinomial logit model for continuous choice. Note that it is standardized, which implicitly sets the scale of surplus and utilities.

Under these assumptions, given a set of functions $U_{ab}(x, y)$, firm $i \in I(a, x)$:

- remains a standalone firm with probability

$$\mu(\emptyset|a, x) = \frac{1}{1 + \sum_{b=1}^M \int \exp(U_{ab}(x, y))dy};$$

- merges with a firm with discrete characteristics b with probability

$$\mu(b|a, x) = \frac{\int \exp(U_{ab}(x, y))dy}{1 + \sum_{d=1}^M \int \exp(U_{ad}(x, y))dy};$$

- and conditional on merging with a firm with discrete characteristics b , firm i merges with a firm that has continuous characteristics y with probability

$$\mu(y|a, x, b) = \frac{\exp(U_{ab}(x, y))}{\int \exp(U_{ab}(x, t))dt};$$

The average expected equilibrium payoff of the firms in $I(a, x)$ is

$$u_a(x) = \log \left(1 + \sum_{b=1}^M \int \exp(U_{ab}(x, y))dy \right).$$

In equilibrium, the matching probabilities must be consistent with the number of potential partners with given observable characteristics: $\mu_{ab}(x, y) = N_a f(x|a) \mu(b, y|a, x)$ must equal $\mu_{ba}(y, x) =$

$N_b f(y|b) \mu(a, x|b, y)$ for all (a, b, x, y) . We substitute the matching probabilities into this equation, yielding the equilibrium condition:

$$\begin{aligned} \mu_{ab}(x, y) &= N_a f(x|a) \frac{\exp(U_{ab}(x, y))}{1 + \sum_d \int \exp(U_{ad}(x, t)) dt} \\ &= N_b f(y|b) \frac{\exp(\Phi_{ab}(x, y) - U_{ab}(x, y))}{1 + \sum_c \int \exp(\Phi_{cb}(z, y) - U_{cb}(z, y)) dz}. \end{aligned} \quad (7)$$

This system of equations in $\mathbf{U} = (U_{ab}(x, y))$ (or equivalently in $\boldsymbol{\mu} = (\mu_{ab}(x, y))$) has a solution, which is unique; the proof follows the logic in Galichon and Salanié (2021) and Dupuy and Galichon (2014). Moreover, it can be computed efficiently using a variant of the Iterative Proportional Fitting Procedure, which is described in Appendix A.3.

The total surplus $\mathcal{W}$ in equation (5) is the sum of all individual utilities:

$$\mathcal{W}(\mathbf{U}) = \sum_a N_a \int f(x|a) u_a(x) dx = \sum_a N_a \int f(x|a) \log \left(1 + \sum_{b=1}^M \int \exp(U_{ab}(x, y)) dy \right) dx \quad (8)$$

which is a globally convex function of $\mathbf{U}$.

3.3 The Moment Matching Estimator

In our application, we follow the moment matching strategy developed in Galichon and Salanié (2021) and specify $\Phi_{ab}(x, y)$ as a linear combination of known *basis functions*, $\phi_{ab}^k(x, y)$, with unknown weights λ_k :

$$\Phi_{ab}^\lambda(x, y) = \sum_{k=1}^K \lambda^k \phi_{ab}^k(x, y). \quad (9)$$

We will estimate the vector $\boldsymbol{\lambda}$ (whose dimension is equal to K , the number of basis functions) by fitting the observed first moments of the basis functions. To do this, we minimize the function

$$\mathcal{W}(\mathbf{U}^\lambda) - \frac{\lambda}{2} \cdot \sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy \quad (10)$$

over the parameter vector $\boldsymbol{\lambda}$, where $\mathbf{U}^\lambda$ solves the system (7) for $\boldsymbol{\Phi} = \boldsymbol{\Phi}^\lambda$, $\phi_{ab}(x, y)$ is the vector of basis functions $\phi_{ab}^k(x, y)$ for each $\{a, b, x, y\}$, and $\hat{\boldsymbol{\mu}}$ is the observed matching.

The rationale behind this minimization is simple. The first-order conditions with respect to $\boldsymbol{\lambda}$

are

$$\frac{d\mathcal{W}}{d\boldsymbol{\lambda}}(\mathbf{U}^\lambda) = \frac{1}{2} \sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy.$$

While computing the derivative of social welfare with respect to $\boldsymbol{\lambda}$ may seem forbidding, in fact it can be done quite simply. Galichon and Salanié (2021) showed that the social welfare can be rewritten as

$$\mathcal{W}(\mathbf{U}^\lambda) = \max_{\boldsymbol{\mu}} \left(\sum_{a,b=1}^M \int \int \mu_{ab}(x, y) \Phi_{ab}^\lambda(x, y) dx dy - \mathcal{E}(\boldsymbol{\mu}) \right),$$

where $\mathcal{E}(\boldsymbol{\mu})$ is the entropy of the matching $\boldsymbol{\mu}$. Since $\boldsymbol{\Phi} = \boldsymbol{\phi} \cdot \boldsymbol{\lambda}$, applying the envelope theorem to this program shows that the derivative in $\boldsymbol{\lambda}$ is simply the vector $\sum_{a,b=1}^M \mu_{ab}^\lambda(x, y) \phi_{ab}(x, y)$, where $\boldsymbol{\mu}^\lambda$ is the equilibrium matching for $\boldsymbol{\Phi}^\lambda$ in equation (7). This gives (cancelling out the factors 1/2 whose presence is necessary to avoid counting a merger twice):

$$\sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \mu_{ab}^\lambda(x, y) dx dy = \sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy :$$

and we choose $\boldsymbol{\lambda}$ to fit the observed moments of the basis functions.

3.4 Using EU mergers data

So far we have proceeded as if we observed all firms and all mergers in the universe of potential partners for firms in our five European countries (EU5). Our data, however, do not include deals that involve firms outside these five countries, and we only have data on the firms that do merge. Fortunately, the multinomial logit structure of our specification permits surplus estimation from the data we have. As is well-known, this specification incorporates a form of Independence of Irrelevant Alternatives; this is very useful for our purposes, as we can rely on the relative probabilities of the matches we do observe. In fact, we show in Appendix A.4 that we only need to rewrite the equilibrium conditions (7) as

$$\bar{N}_a \bar{f}(x|a) \frac{\exp(U_{ab}(x, y))}{\sum_d \int \exp(U_{ad}(x, t)) dt} = \bar{N}_b \bar{f}(y|b) \frac{\exp(\Phi_{ab}(x, y) - U_{ab}(x, y))}{\sum_c \int \exp(\Phi_{cb}(z, y) - U_{cb}(z, y)) dz}, \quad (11)$$

where $\bar{N}_a$ and $\bar{N}_b$ refer to our sample of merging EU5 firms, and $\bar{f}(x|a)$ and $\bar{f}(y|b)$ are the distributions of continuous characteristics within discrete group for our merging firms. After this

redefinition, all of Section 3.3 remains valid.

Using only mergers data has consequences for the identification of λ , however. Suppose that we add any term $h_b(y)$ to $\Phi_{ab}(x, y)$ in (11). Such a term will factor out of both the denominator and the numerator, as explained in more detail in Appendix A.4. Thus, and by symmetry, we can only identify $\Phi_{ab}(x, y)$ up to a sum of arbitrary terms $h_a(x) + h_b(y)$. Returning to our linear expansion on basis functions, we identify only the weights λ_k of the basis functions that contain interactions between the observable characteristics of both partners. In our application, we only use this subset of basis functions.

Moreover, it is also apparent from equation (11) that multiplying Φ by any positive constant multiplies all values of U by the same constant, and leaves all matching patterns unchanged. Therefore the parameter vector λ is estimated up to a positive multiplicative constant that sets the units in which the surplus function is defined.

To summarize, for any choice of units we are able to estimate our object of interest Φ up to arbitrary terms that depend on the stand-alone firm characteristics, $h_a(x)$ and $h_b(y)$:

$$\bar{\Phi}_{ab}(x, y) = \Phi_{ab}(x, y) + h_a(x) + h_b(y) \quad (12)$$

and up to a positive multiplicative constant. To put it differently, we identify what Chiappori, Salanié, and Weiss (2017) call the *supermodular core*: all cross-differences

$$\bar{\Phi}_{ab}(x, y) + \bar{\Phi}_{a'b'}(x', y') - \bar{\Phi}_{ab'}(x, y') - \bar{\Phi}_{a'b}(x', y).$$

The identified properties of Φ are, hence, the interactions between the observable characteristics of the potential partners that reveal the complementarities that contribute to merger surplus.

4 Results

In this section, we estimate versions of equation (12) with the data about observed mergers described in Section 2. The matching model estimation allows us to use the frequency of matches between firms with given characteristics among all firms involved in mergers to estimate the surplus function that best corresponds to observed matches.

As set out in Section 3, the surplus of the merger between two firms with observable characteristics $\{a, x\}$ and $\{b, y\}$ is $\Phi_{ab}(x, y)$, where a, b refer to the partners' discrete characteristics (country and industry) and x, y are the corresponding two-dimensional vectors of continuous variables, productivity z and input scale Ω . Our baseline estimation proposes a quadratic surplus function of these attributes:

$$\begin{aligned}\Phi_{ab}(x, y) = & \lambda_{ab}^0 + \lambda_{ab}^1 \frac{1}{2} (z_x + z_y) + \lambda_{ab}^2 \frac{1}{2} (\Omega_x + \Omega_y) + \lambda_{ab}^3 \frac{1}{2} (\Omega_x z_x + \Omega_y z_y) \\ & + \lambda_{ab}^4 (z_x z_y) + \lambda_{ab}^5 \frac{1}{2} (\Omega_x z_y + \Omega_y z_x) + \lambda_{ab}^6 (\Omega_x \Omega_y)\end{aligned}\quad (13)$$

the estimated coefficients λ_{ab}^k are allowed to vary depending whether the two merging firms share discrete characteristics. For $k = 0, \dots, 6$:

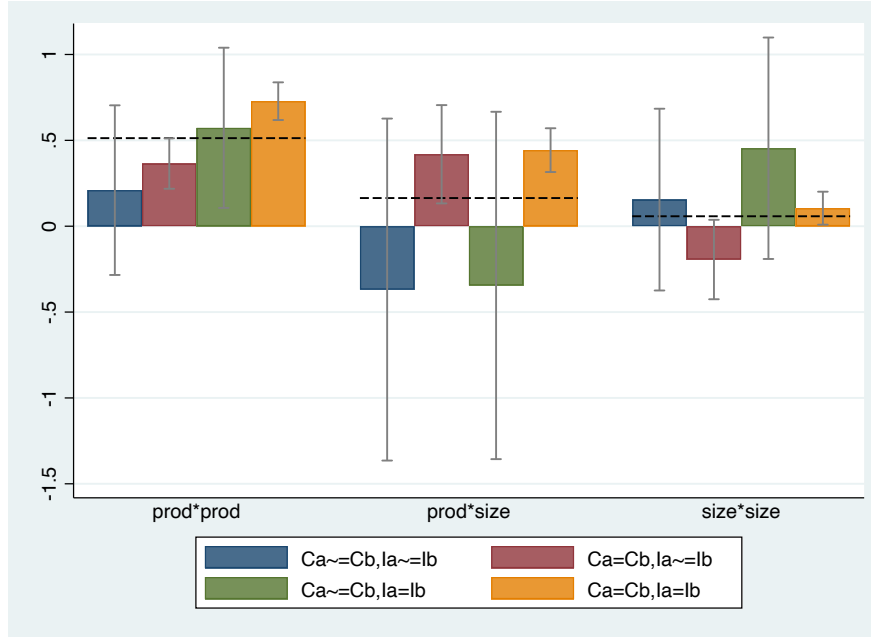
$$\lambda_{ab}^k = \begin{cases} \lambda_k & \text{if } a \neq b \\ \lambda_k + \gamma_k & \text{if } a = b \end{cases} \quad (14)$$

where $a = b$ (resp. $a \neq b$) refers to both firms operating in the same country $C_a = C_b$ (resp. $C_a \neq C_b$) or the same industry group $I_a = I_b$ (resp. $I_a \neq I_b$). Given the identification restriction in equation (12), we are able to identify the coefficients that involve interactions between merging firms' characteristics. This implies that for $k = 0, 1, 2, 3$, only the additive terms γ_k are identified. For $k = 4, 5, 6$, which are identified based on the interaction between both firms' continuous characteristics, our methodology can estimate both λ_k and γ_k .

The main results are shown in Figure 1. We find positive complementarities on average on all dimensions, shown by the black dotted lines in all three sets of histograms. However, there is also significant heterogeneity in the strength of complementarities in different dimensions, shown by the different heights of the black dotted lines, and across types of deals for any one complementarity, shown by the different bars in each set.

First, the complementarity in revenue productivities is the single most important source of synergy between merged firms: The pooled coefficient for all deals in our sample is 0.513 (significant at the 1% level), shown by the black dotted line in the first set of histograms in Figure 1. Positive assortative matching suggests that mergers take place to exploit complementarities between the capabilities of merged firms. This source of complementarity is strongest for within-industry

Figure 1: Quadratic Model. Complementarities



Note: Summary of estimates (mean and 95% intervals) of complementarities between continuous characteristics in equation 13 by type of deals: cross-border and different industry group, domestic and same industry group. Estimation also includes constant, sum of firms' productivities, size (input scale), and their within-firm combination (λ^0 , λ^1 , λ^2 , λ^3) not shown. Dotted black line corresponds to the pooled coefficients.

matches, both domestic and cross-border, which account for 74% of the deals in our data. This finding is consistent with the idea that technologies are not replaceable but, instead, can be adopted by different units within a firm at a cost that depends on the unit’s original productivity. This feature coincides with the finding in [Bilir and Morales \(2020\)](#) for the case of foreign affiliates within a multinational corporation: new technologies in the headquarters increase the productivity of affiliates if the affiliate is also innovative. Note that the definition of within-industry deals in our case is a quite broad set of related industries so that the results suggest a technological complementarity across related mergers in contrast to a much smaller effect for unrelated mergers.

This result is at odds with the assumption in the international trade literature that a headquarters is able to transmit its superior technology to foreign subsidiaries within the firm, either freely or at an exogenous discount that parallels the iceberg trade cost in international trade (see, for example, [Ramondo and Rodriguez-Clare, 2013](#), [McGrattan and Prescott, 2009](#), [Helpman, Melitz, and Yeaple, 2004](#)). When FDI takes the form of foreign acquisition, the technology of the local unit is replaced by the superior one from HQ, and, if low productivity firms have lower standalone values, the surplus of the cross-border merger increases in the difference between acquirer’ and target’s productivities. In all cases, superior foreign technology replaces local unproductive units, either through the general equilibrium effect (the exit of less productive firms) or directly through M&As. As a result, the implied global productivity gains from multinational activity are very large. We find instead that productive foreign firms *cherry-pick* productive local firms with whom to merge, consistent with [Javorcik and Arnold \(2009\)](#) and [Guadalupe, Kuzmina, and Thomas \(2012\)](#). The implication of our finding is that already high-performing local firms become subsidiaries of multinationals. If cross-border mergers result in productivity gains in the host country, it *is not* because formerly unproductive local firms now attain the high productivity of a multinational parent but because the adoption of new technologies is easier for originally productive firms. The expected productivity gains from multinational expansion are hence much reduced compared to the models in the literature. In addition, our findings suggest that the ability of countries to attract FDI is likely limited by the existing capabilities of domestic firms.

Second, the second largest complementarity in our data, on average, is the cross-characteristic complementarity (scale*TFPr). The pooled cross coefficient is positive: 0.164 and statistically significant at the 1% level (see the black dotted line in the second set of histograms in [Figure 1](#)).

But as we can easily see in Figure 1, the average hides very different results for cross-border and within-country mergers, and turns out to be significantly positive only for domestic mergers. In fact, for within country, cross-industry mergers this positive size-productivity complementarity is the largest source of expected complementarity we estimate.

In contrast, for cross-border mergers, the estimates are negative, suggesting substitutability in the characteristics (though with large standard errors). These results inform the models of FDI that explicitly incorporate heterogeneous local firms, as in [Nocke and Yeaple \(2007\)](#), which typically predict positive assortative matching in cross characteristics—i.e., more productive firms are likely to merge with larger foreign counterparts. The intuition is that firms can leverage their productivity by combining it with the relatively immobile assets of their merger partner, be it their tangible inputs, such as capital and labor, or other assets such as their distribution network or marketing capabilities. We do not find that highly productive firms match with foreign counterparts with large input scale in the cross-border deals in our sample on average.

Finally, the smallest source of expected complementarities are those between the input scale of both firms. This is true on average but also for all types of mergers with the exception of cross-border, within industry mergers where they are large (although imprecisely measured). For mergers between firms in the same country-industry market, the complementarity between partners' size is low but significantly different from zero at 5% level.

Next, to further explore the heterogeneity in complementarities across deals, we estimate an extension of equation (13) that incorporates non-linearities between continuous variables. This empirical model enables us to analyze how the complementarities shown in Figure 1 change across firms and it is also closer to theoretical frameworks in which productivity and size enter multiplicatively in the surplus function.¹⁷

$$\begin{aligned} \Phi_{ab}(x, y) = & \lambda_{ab}^0 + \lambda_{ab}^1 \frac{z_x + z_y}{2} + \lambda_{ab}^2 \frac{z_x \Omega_x + z_y \Omega_y}{2} + \lambda_{ab}^3 \frac{\Omega_x + \Omega_y}{2} + \lambda_{ab}^4 z_x z_y + \lambda_{ab}^5 \frac{z_x \Omega_y + z_y \Omega_x}{2} \\ & + \lambda_{ab}^6 \Omega_x \Omega_y + \lambda_{ab}^7 \frac{\Omega_x + \Omega_y}{2} z_x z_y + \lambda_{ab}^8 \frac{z_x + z_y}{2} \Omega_x \Omega_y + \lambda_{ab}^9 z_x z_y \Omega_x \Omega_y. \end{aligned} \quad (15)$$

¹⁷Profits in equation (2) are often modeled as the product of productivity and size $\Pi(x) = \eta z_x \Omega_x$. See [Nocke and Yeaple \(2007, 2008\)](#), where z_x is firm-wide factor, mobile across markets within the corporation, and Ω_x is interpreted as the firm's market-specific dimension; or [Blonigen, Fontagné, Sly, and Toubal \(2014\)](#), which refers to Ω_x as irreversible investment for entry into markets, or the purchase of distribution networks. In [Deneckere and Davidson \(1985\)](#), Ω_x corresponds to the set of heterogeneous goods produced by the firm.

The coefficients λ_{ab}^k for $k = 0, \dots, 9$ are defined as in (14). Under the restrictions in equation (12), our methodology cannot identify $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ but only $\gamma_0, \gamma_1, \gamma_2, \gamma_3$, based on the differential effect of the interaction between firms' discrete characteristics. For $k = 4, \dots, 9$, where the relevant variables involve interactions between the continuous characteristics of merging firms, both λ_k and γ_k can be estimated.

This estimation allows us to analyze how complementarities vary according to the size and productivity of the parties. Differentiating equation (15) with respect to the productivity of each firm, and then also with respect to the scale of each firm, gives the following expressions showing how the synergies between the partner firms' productivities and sizes vary across firms of different characteristics:

$$\frac{\partial^2 \phi_{ab}(x, y)}{\partial z_x \partial z_y} = \lambda_{ab}^4 + \lambda_{ab}^7 \frac{(\Omega_x + \Omega_y)}{2} + \lambda_{ab}^9 (\Omega_x \Omega_y) \quad (16)$$

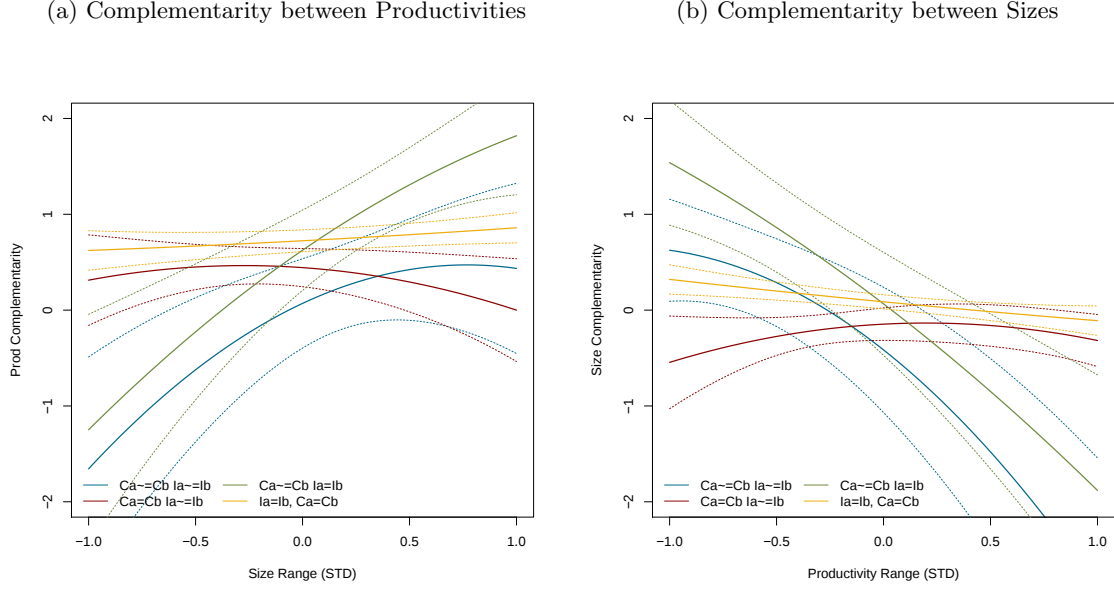
$$\frac{\partial^2 \phi_{ab}(x, y)}{\partial \Omega_x \partial \Omega_y} = \lambda_{ab}^6 + \lambda_{ab}^8 \frac{(z_x + z_y)}{2} + \lambda_{ab}^9 (z_x z_y) \quad (17)$$

The first term in equation (16) corresponds to the baseline complementarity between firms' productivities when $\Omega_x = \Omega_y = 0$, which is the standardized mean size in the full sample. The second term shows how these synergies evolve with the size of the combined entity. And, finally, the third term recovers the joint size-productivity complementarities. Similarly, size complementarity in equation (17) can be decomposed into three corresponding terms. The second and third terms show how the size synergies evolve according to the productivity of the partners.

The results are shown in Figure 2. Figure 2a plots the complementarity between productivities according to the scale of the paired firms ($\Omega_x = \Omega_y$ is shown in the horizontal axis, one standard deviation around its group mean). This decomposition confirms the results in Figure 1: within-industry deals exhibit larger synergies than cross-industry deals, for any firm size. For domestic deals, the coefficient is stable across the size range. In the case of cross-country deals, productivity synergies change substantially with the size of the partner firms, which explains the high standard errors of the estimates in Figure 1.

Figure 2b plots the complementarity between the scales of the merged firms at different levels of joint scale, based on expression (17). In the case of domestic deals, the results in Figure 1 are fairly constant across firms of different productivities (with $z_x = z_y$, is in the horizontal axis).

Figure 2: Complementarities between Merging Firms



Note: The horizontal axis correspond to size (in Figure a) and productivity (in Figure b) of the merging firms (same for both firms). The range goes for -1 and +1 group-specific standard deviation around the group mean. 95% interval in dotted lines. Mean and STD for each group of mergers are in Table 2.

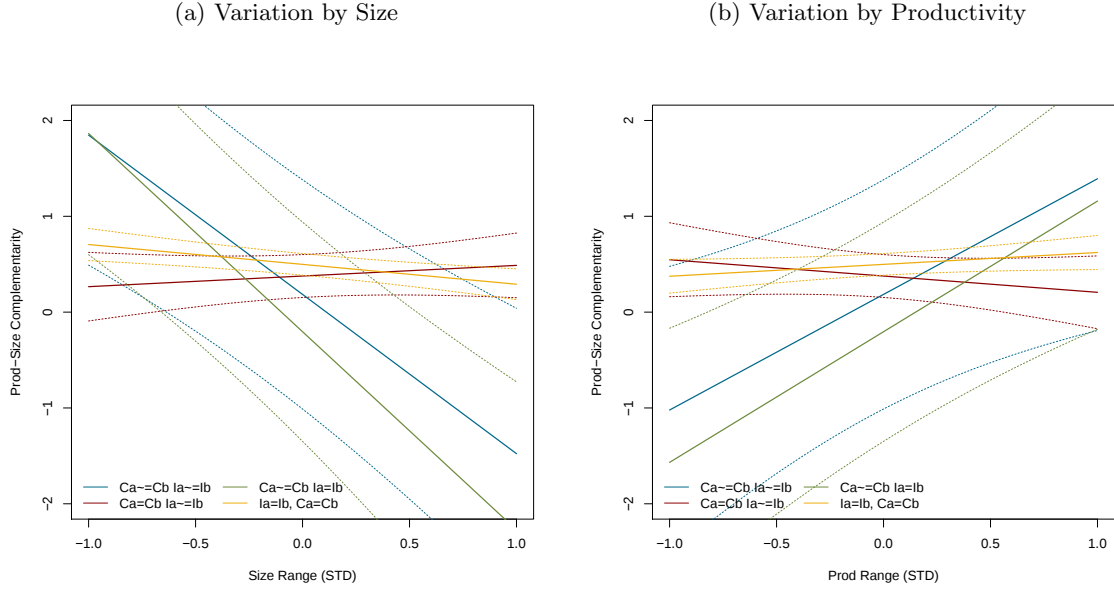
Interestingly, mergers between large firms competing in the same industry-country market, which could be expected to have a market-power motive, show no significant complementarity in scale. We cannot, however, interpret from this result that market power is not a profitable motive for mergers in any general sense. Since our analysis is based on actual deals, our findings may suggest that mergers aimed at exploiting monopoly rents were blocked (or expected to be blocked) by antitrust authorities. For cross-country deals, size synergies are the highest for partners with productivity below the group-average (the group average is larger for cross-country deals, as shown in Table 2); they become significantly negative for deals between large counterparts.

The cross-characteristic complementarity between revenue productivity and input scale is shown in Figure 3. It changes with the size and productivity of the involved firms according to the following expression:

$$\frac{\partial^2 \phi_{ab}(x, y)}{\partial z_x \partial \Omega_y} + \frac{\partial^2 \phi_{ab}(x, y)}{\partial z_y \partial \Omega_x} = \lambda_{ab}^5 + \lambda_{ab}^7 \frac{(z_y + z_x)}{2} + \lambda_{ab}^8 \frac{(\Omega_x + \Omega_y)}{2} + \lambda_{ab}^9 (z_y \Omega_x + z_x \Omega_y) \quad (18)$$

Again, the estimates describing the synergies between domestic firms are stable across size and

Figure 3: Cross-Characteristic Complementarity



Note: In Figure (a), productivity of matched firms fixed at the group mean and size (equal for both firms) varies between -1 and +1 standard deviation around the mean. In Figure (b), size of matched firms fixed at the group mean and productivity (equal for both firms) varies between -1 and +1 group-specific standard deviation around the mean. 90% confidence in dotted lines. Mean and STD for each group of deals are in Table 2.

productivity of the partners.

The estimates from equation (15) for cross-country mergers, on the other hand, show very heterogeneous and noisy patterns. Figures 2 and 3 show huge variation in the estimates of complementarities over variation in merged-entity productivity or size. This is why the estimates in Figure 1 are imprecisely estimated for cross-border mergers. Overall, the product-size complementarity tends to decrease with the size of the partners and increase with their productivity for these mergers.

5 Comparing Predicted and Observed Merger Gains

We next explore whether post-merger performance is related to the estimates of expected merger surplus from Section 4. If the parameters of the matching model are informative about firms' merger motives and if firms are able to realize the anticipated gains from the merger then we expect to see that post-merger firm performance is positively correlated with expected merger surplus. However, no such relationship will likely be discernible in the data if either (or both) of these conditions fails

to hold.¹⁸

In this investigation, we are interested in the difference between the actual post-merger performance and the hypothetical counterfactual outcomes had the firms not merged, which is the empirical equivalent of the merger surplus defined in equation (2). To construct estimates of both the actual and counterfactual components of this surplus measure, we will assume the value of firm i in year t reflects a constant returns to scale Cobb-Douglas production function that can be written :

$$\Pi_{it} = \eta R_{it} = \eta [z_{it} L_{it}^{\alpha} K_{it}^{1-\alpha}] \quad (19)$$

where z_{it} is revenue productivity, and L_{it} and K_{it} are labor and capital. R_{it} is firm revenues and η is a constant. Before turning to make some further assumptions about the hypothetical counterfactual, we first discuss how we measure actual post-merger performance.

5.1 Data about post-merger outcomes

We have 8,894 firm-level observations for the 4,447 bilateral mergers studied in this paper. The firm identifiers in Zephyr allow us to find firm-level performance after the merger in ORBIS.¹⁹ Historical ORBIS (HO) is country-specific and contains information about the firms that continued to file separate accounts post merger. Out of the 628, 2,029, 3,312, and 1,011 firm entries in the sample for Germany, Spain, the UK, and Italy, respectively, 351, 1,399, 2,804, and 446 firms have post-merger data in HO for at least the first year after the merger.²⁰ Where they are available, we gather the data on total assets, number of employees, operating revenue, and profit before tax for these firms. Unfortunately, we were unable to evaluate post-merger performance for mergers including French firms as none of the 1,914 French merging firm ids were in HO.

To measure the performance of the merged entity in the years after the merger, we need data on both firms in the pair. We are able to construct measures for 1,375 deals in the first year after

¹⁸The management literature details various challenges for successful post-merger integration that could limit any merger gains. Birkinshaw, Bresman, and Håkanson (2000) argue that human resource management and task integration are critical, and Reus, Lamont, and Ellis (2016) discuss how attempts to transfer knowledge within firm boundaries can destabilize power structures and hamper performance.

¹⁹We note that some firm ids are repeated because they take part in multiple deals in the sample.

²⁰An additional mapping of the old to the new ids issued by BvD to some firms in 2016 yielded the majority of the matches for the Italian firms.

the merger. The number of deals included falls in each year as one or both firms leave the data, leaving 841 deals five years after the year of the merger. The combined profit data is available for fewer firms.²¹

Consider two firms i and j that merge in year $t = 0$. All available data on these firms after $t = 0$ reflect any merger gains that are realized. The data from year t allow us to find the joint revenues of the actual merged firm, m , in year t , $R_{mt} = R_{it} + R_{jt}$. Similarly, we sum the profits of the two merged firms to give us the joint profits of the merged firm m in year t , $\pi = \pi_{it} + \pi_{jt}$. From equation (19), the actual revenue productivity and input scale of the merged firm m in years $t > 0$ can be computed by combining data on observed revenues, employment, and total assets of firms i and j in year t . Defining $L_{mt} = L_{it} + L_{jt}$, and $K_{mt} = K_{it} + K_{jt}$, we estimate the revenue productivity for the merged firm in year $t > 0$ as:

$$z_{mt} = \ln(R_{mt}) - \alpha \ln(L_{mt}) - (1 - \alpha) \ln(K_{mt}) = r_{mt} - \alpha l_{mt} - (1 - \alpha) k_{mt}, \quad (20)$$

and $\Omega_{mt} = \alpha \ln(L_{mt}) + (1 - \alpha) \ln(K_{mt}) = \alpha l_{mt} + (1 - \alpha) k_{mt}$ is the estimate for the input scale of the merged firm in year $t > 0$. The α estimates are computed at the two-digit NACE Rev. 2 industry level, following the procedure in Section 2.²²

Table 4 presents the percentage growth in revenues, profit, input scale and revenue productivity in each of the first five years after merging. All the variables have been winsorized at 1% and 99%. The values come with some caveats. First, the HO data we have on revenues and total assets is in USD, whereas the accounts are likely reported in EUR or GBP using the exchange rate on the date corresponding to the end of each firm's financial year. Both the Euro and GBP appreciated against the USD during the 2001 to 2013 period, corresponding to one to five years after the mergers between 2000 and 2008, amplifying the reported average growth rates. Second, while the HO data appear to be consistently reported in year $t = 0$ and in later years, it is possible that any change in firms' organizational structures that accompanied the merger may have affected internal accounting

²¹ Attrition from post-merger data does not necessarily imply firm death, but could reflect relabeling of the firm identifier after the merger, a selection bias where the sign of the bias is not obvious.

²² We use the estimates produced using the [Akerberg, Caves, and Frazer \(2015\)](#) approach at the two-digit industry level, using the acquirer firm's industry code, without imposing constant returns to scale as in equation (19). For the combined firms in the analysis, the mean estimated $\alpha + \beta$ is 1.076, and the standard deviation is 0.032, suggesting that the constant returns assumption in equation (19) is, however, reasonable. The smallest value is 0.096, and the largest is 1.211, which is for the 2-digit industry code 56 (food and beverage service activities) and is an outlier.

Table 4: Mean percentage growth between year 0 and year t , for $t = 1, \dots, 5$.

Years	1	2	3	4	5
Revenues	17.51	20.46	22.77	18.98	19.59
Std dev	(39.79)	(52.04)	(62.88)	(81.85)	(88.00)
Std dev within country-industry	(13.87)	(18.60)	(23.41)	(27.91)	(31.33)
Profits	16.03	16.07	19.46	18.08	16.79
Std dev	(82.17)	(102.90)	(110.54)	(118.36)	(121.35)
Std dev within country-industry	(22.01)	(34.94)	(41.31)	(43.22)	(45.31)
Input scale	12.36	17.05	18.81	15.94	15.96
Std dev	(30.70)	(41.92)	(52.49)	(64.16)	(72.06)
Std dev within country-industry	(7.53)	(13.49)	(16.79)	(18.79)	(22.83)
TFPr	4.70	3.26	4.35	2.76	2.87
Std dev	(32.14)	(34.38)	(38.54)	(44.59)	(45.51)
Std dev within country-industry	(8.57)	(9.71)	(11.67)	(13.67)	(12.86)
Num. mergers, for revenues, scale, and productivity	1,375	1,202	1,051	939	841
Num. mergers, for profit variable	941	789	676	588	522

Notes: The table shows winsorized values at 1% and 99%.

processes.

Bearing these two caveats in mind, the data show that the first three years after merging saw large increases in revenues, profits and input scale, on average, and small increases in revenue productivity. The standard errors are very large, although they are smaller once average variation by the country and industry group of each firm in the merger has been removed. By years four and five, some of the early gains had been lost, but all variables are positive, and mean revenue, profit, and input scale growth remain large.

5.2 The difference between actual and counterfactual post-merger outcomes

To evaluate whether the gains shown in Table 4 are informative about merger surplus, we need to compare them to an estimate of the equivalent growth rates had the firms not merged.²³ To do this, we return to HO. We use data on the revenue, profits, employment, and total asset growth rates of

²³There is a growing literature investigating merger effects on firm performance that adopts various ways to address the endogeneity of merger decisions that is the main focus of this paper. For example, [Blonigen and Pierce \(2016\)](#) find evidence of increases in markups and little productivity gain using differences in differences methods comparing to matched firms that do not merge, to firms in deals that are not completed, and to firms that will be acquired in the future. They also instrument for successful merger completion.

firms in the same industry in the same country in the years since $t = 0$. Specifically, for revenues, we assume that the revenues of each firm would have grown at the same rate as the median growth rate of all firms in HO between years t and $t - 1$ that are in the same country and NACE Rev 2 two-digit industry. Let i' indicate firm i under the counterfactual that it did not merge with firm j , then the counterfactual revenues in each year $t > 0$ after the merger are:

$$R_{i't} = R_{i',t-1}(1 + g_{rt})$$

where g_{rt} is the median revenue growth rate for firms in the same country-industry group between years t and $t - 1$, and $R_{i'0} = R_{i0}$. That is, the counterfactual and actual growth are both measured relative to the same value in year 0. Counterfactual profits, employees, and total assets are estimated in a similar way.

We next construct a synthetic firm m' that consists of the two counterfactual stand-alone units i' and j' . The expected revenues for the synthetic m' in the years after the merger are assumed to simply be the sum of the counterfactual estimated revenues of firms i' and j' :

$$E_0[R_{m't}] = R_{i't} + R_{j't}, \quad (21)$$

and the counterfactual expected profits, number of employees and total assets are found analogously, $E_0[\pi_{m't}] = \pi_{i't} + \pi_{j't}$, $E_0[L_{m't}] = L_{i't} + L_{j't}$, and $E_0[K_{m't}] = K_{i't} + K_{j't}$.

Under the constant returns to scale assumption in equation (19), the expected revenue productivity of the synthetic counterfactual merged firm m' in years $t > 0$ is computed by combining data on estimated counterfactual revenues, employment, and total assets of firms i' and j' in year t . The estimate for the revenue productivity for the counterfactual entity in year $t > 0$ is:

$$E_0[z_{m't}] = \ln(E_0[R_{m't}]) - \alpha \ln(E_0[L_{m't}]) - (1 - \alpha) \ln(E_0[K_{m't}]), \quad (22)$$

and the estimate of the counterfactual entity's input scale in year $t > 0$ is:

$$E_0[\Omega_{m't}] = \alpha \ln(E_0[L_{m't}]) + (1 - \alpha) \ln(E_0[K_{m't}]), \quad (23)$$

using the same α estimates as in equation (20).

For each of revenues, profits, input scale, and productivity, we next compute the percentage difference between actual and counterfactual growth by subtracting the log of expected counterfactual from the log of the actual. So, this difference for revenues is $\Delta r_t = \ln(R_{mt}) - \ln(E_0[R_{m't}])$, and similarly for profits, while input scale and productivity are already in logs, and so are given by $\Delta \Omega_t = \Omega_{mt} - E_0[\Omega_{m't}]$ and $\Delta z_t = z_{mt} - E_0[z_{m't}]$, respectively.

Table 5 shows the percentage growth in revenues, profits, input scale, and revenue productivity compared to the counterfactual in each of the first five years after merging. These variables are winsorized at 1% and 99% after subtracting the counterfactual growth. Since both the actual and counterfactual measures are given in USD, the first of the caveats attached to Table 4 does not cast any doubt on the large magnitudes in Table 5. However, because the counterfactual measure does not account for any organizational changes, some of the large gains in revenues, profits, and input scale may reflect changes in the reporting basis for the firm-level accounts.

Subject to this remaining caveat, some interesting features emerge from Table 5. Revenue growth and input scale growth post-merger remain large when compared to the counterfactuals, at around 10%.²⁴ Profit growth compared to the counterfactual is actually much larger than actual profit growth because counterfactual profit growth tends to be negative in this time period. Much of the post-merger data relates to the years of the financial crisis in Europe after 2007, which means that the positive actual profit growth observed in merging firms, shown in Table 4, is more striking. The standard deviations in all measures of post-merger performance given in Table 5 remain large even when controlling for country-industry averages, but Table 5 offers some evidence that the owners and managers of the merging firms were able to achieve better performance than their non-merging country-industry peers.

The post-merger revenue productivity growth relative to the counterfactual is less impressive. While it is positive for the first three years after the merger, the magnitudes are relatively small.²⁵ This finding is somewhat puzzling given the large role that complementarity in productivity plays in

²⁴Interestingly, [Geurts and Van Biesebroeck \(2019\)](#) study the long-term employment effects of mergers between Belgian firms between 2005 and 2012. They find that mergers likely to be motivated by market power show a permanent employment reduction of 14%, whereas those motivated by efficiency gains lead to employment expansions of 10%.

²⁵[Ashenfelter, Hosken, and Weinberg \(2015\)](#) and [Braguinsky et al. \(2015\)](#) show industry-specific sources of productivity gains after merging relating to plant-level economies of scope and increased leverage of managerial talent, respectively.

Table 5: Mean percentage growth between year 0 and year t , for $t = 1, \dots, 5$, relative to counterfactual.

Years	1	2	3	4	5
Revenues	13.28	15.22	16.01	10.66	10.92
Std dev	(38.18)	(49.03)	(59.76)	(78.51)	(83.74)
Std dev within country-industry	(7.19)	(6.14)	(11.51)	(13.11)	(17.19)
Profits	36.52	62.01	88.73	112.83	136.54
Std dev	(81.28)	(99.08)	(103.98)	(114.71)	(119.17)
Std dev within country-industry	(16.78)	(20.42)	(24.89)	(32.29)	(37.46)
Input scale	10.20	13.81	14.31	10.01	9.45
Std dev	(29.99)	(40.93)	(51.53)	(62.65)	(71.01)
Std dev within country-industry	(5.46)	(9.46)	(13.00)	(13.63)	(18.50)
TFPr	2.89	1.32	2.22	0.25	0.74
Std dev	(31.51)	(34.21)	(37.40)	(43.90)	(45.82)
Std dev within country-industry	(5.95)	(8.19)	(8.95)	(11.24)	(12.10)
Num. mergers, for revenues, scale, and productivity	1,375	1,202	1,051	939	841
Num. mergers, for profit variable	941	789	676	588	522

Notes: The table shows winsorized values at 1% and 99%.

motivating mergers according to the surplus function estimates in Section 4. Worth remembering, though, is that the anticipated gains relating to revenue productivity complementarity need not come in the form of revenue productivity growth, but might play some role in increasing the scale of the merged entity relative to the sum of the standalone scales.

5.3 Relationship between estimated surplus and post-merger performance

We now investigate whether the data summarized in Table 5 provide any support of the surplus function we generated in Section 4. Our goal is to estimate versions of the following equation, with post-merger revenues relative to the hypothetical counterfactual on the left hand side, and the estimated merger surplus predicted by the matching model on the right hand side:

$$\Delta r_t = \ln(R_{mt}) - \ln(E_0[R_{m't}]) = f(\Phi_{ab}(z_x, \Omega_x, z_y, \Omega_y)) \quad (24)$$

as well as the analogous relationships with profits, input scale, and revenue productivity replacing revenues on the left hand side.

We first simply regress performance relative to the counterfactual on the sum of the identified complementarity terms in the predicted merger surplus $\Phi_{ab}(z_x, \Omega_x, z_y, \Omega_y)$ shown in Section 4. We take the pooled λ estimates from the dotted lines in Figure 1 together with the measures of firm revenues and profits, and estimated input scale and revenue productivity, from the year $t = 0$ using the HO data and denote this term $\hat{\Phi}_m$. Table 6 presents the estimated coefficients for each year $t = 1, \dots, 5$ after the merger for revenues, profits, input scale, and productivity. While this regression fails to account for post-merger performance attributable to the unidentified terms in the merger surplus function, that is, the contributions to surplus coming from firm-level characteristics, the estimated coefficients for revenue productivity growth are striking. In each year, the coefficient is positive and significant, increasing from 5% to 8% over five years.

Table 6: Regression coefficients, growth between year 0 and year t relative to counterfactual regressed on the identified merger surplus terms.

Years	(1)	(2)	(3)	(4)	(5)
Revenues on $\hat{\Phi}$	0.048*** (0.017)	0.033 (0.024)	0.043 (0.033)	0.065 (0.044)	0.118** (0.051)
Profits on $\hat{\Phi}$	0.028 (0.051)	0.044 (0.068)	-0.030 (0.082)	-0.111 (0.096)	0.044 (0.102)
Input Scale on $\hat{\Phi}$	-0.015 (0.014)	-0.021 (0.020)	-0.025 (0.029)	0.002 (0.035)	0.011 (0.043)
TFPr on $\hat{\Phi}$	0.047*** (0.014)	0.043** (0.017)	0.053** (0.021)	0.052** (0.025)	0.083*** (0.028)
Observations for Revenues, Input Scale, and TFPr	1375	1202	1051	939	841
Observations for Profits	940	789	676	589	521

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We next turn to a more stringent set of tests of the relationship between revenue productivity growth and estimated merger surplus by including control functions on the right hand side of equation (24) to account for the contribution to merger surplus made by non-interacting firms characteristics (measured by λ_k for $k = 0, 1, 2, 3$ in equation (13)). To proxy for these unidentified terms, we include country- and industry-specific linear control functions of each firm's two continuous characteristics at the time of the merger. So, for firm i , we interact z_{i0} and Ω_{i0} , and their interaction, $z_{i0} \times \Omega_{i0}$, with indicator variables for the firm i 's country and industry group. We do

the same for firm j .²⁶

We then regress Δz_t on the identified terms $\hat{\Phi}_m$ as well as the control functions of each firm's observable pre-merger discrete and continuous characteristics and merger year fixed effects for each year $t > 0$. We also include a set of fixed effects for the year between 2000 and 2008, inclusive, indicating when the merger took place:

$$\Delta z_t = \beta_t \hat{\Phi}_m + \mathbf{x}_0' \mathbf{\Gamma}_t + \mathbf{y}_0' \mathbf{\Gamma}_t + Y_m + \nu_{mt}. \quad (25)$$

where $\mathbf{x}_0$ and $\mathbf{y}_0$ denote the matrices of both firms' characteristics at the time of the merger, $\mathbf{\Gamma}_t$ is the vector of estimated coefficients on the control functions, and Y_m is year-of-the-merger fixed effect.²⁷

Figure 4 plots the $\hat{\beta}_t$ coefficients from estimating equation (25) for each of the five years after the merger. The coefficients in blue correspond to the $\hat{\beta}_t$ coefficients when using the predicted merger surplus estimates across all mergers (i.e. the horizontal lines shown in each set of histograms in Figure 1 on the right-hand side of equation (25)). In green are the $\hat{\beta}_t$ coefficients when using the $\hat{\Phi}$ estimates that vary by type of mergers in equation (25), and as presented in the histogram bars in Figure 1. The regression output is in Appendix A.5.

The coefficients remain positive and significant in most years, showing a positive association between post-merger revenue productivity gains and the merger surplus estimated in the matching model. The magnitudes are similar to those shown in the final row of Table 6, although smaller in later years.²⁸

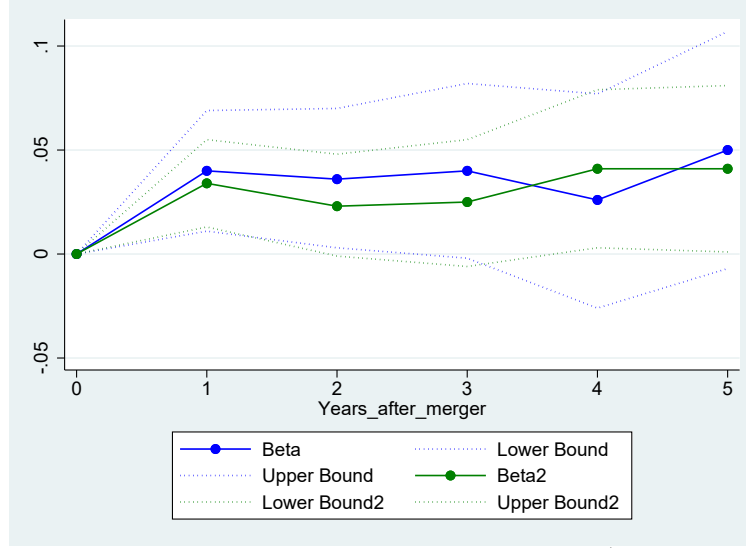
The mean level of $\hat{\Phi}$ is 0.2375 and the standard deviation is 0.6954. Hence, an estimated coefficient of 0.040 for one year after the merger implies that a one standard deviation increase in $\hat{\Phi}$ corresponds to 0.028, or 2.86%, higher predicted revenue productivity growth relative to the coun-

²⁶This gives 30 control functions per firm. The interaction of each firm's country and industry group add another 25 functions per firm. Including these 110 functions in the regressions is quite restrictive and limits our focus to variation in post-merger performance over and above the variation related to each firm's initial revenue productivity, input scale, country and industry group.

²⁷The $\mathbf{\Gamma}_t$ coefficients are constrained to reflect the fact that the merging firms enter the surplus function symmetrically, so that if firms i and j are from the same country and in the same industry group, the firm-specific control functions are constrained to have equal coefficients for the initial characteristics of the two firms.

²⁸Equivalent specifications including revenue, profits, and input scale on the left hand side of equation (25) show no systematic relationship with these variables and estimated surplus, similar to Table 6.

Figure 4: The correlation between revenue productivity gains and predicted merger surplus, including control functions



Note: The x-axis is the number of years post merger. The y-axis shows the $\hat{\beta}$ coefficients from equation (25). In blue are the coefficients based on the quadratic specification of $\hat{\Phi}$. In green, $\hat{\Phi}$ varies with discrete characteristics. The dotted lines are the corresponding 95% confidence intervals.

terfactual. This increase appears large when compared to the mean relative revenue productivity growth relative after one year of 2.89% shown in column 1 of Table 5. However, the standard deviation in relative revenue productivity growth is also very high and the predicted change of 2.86% is equivalent to around 10% of one standard deviation in this variable. By five years after the merger, a one standard deviation in $\hat{\Phi}$ is associated with a 3.48% higher revenue productivity growth compared to the counterfactual.

These regressions offer some evidence that two conditions do indeed hold: the estimates from the matching model are able to capture some part of the firms' merger motives, and also that firms were able to realize some of their anticipated merger gains in the first few years after the merger in the form of revenue productivity gains, although not significantly in the form of revenue, profit, or scale gains. The first of these conditions suggests the results in the matching model are informative about why firms merge. Overall, the mergers predicted to generate more surplus in the matching model experience higher revenue productivity growth immediately after the merger.

6 Conclusion

We study selection into the choice of merger partner among merging firms in five Western European countries between 2000 and 2008. We find that merger partner selection depends on both the firm’s and its merger partner’s pre-merger characteristics. There is positive assortative matching on revenue productivity and also on revenue productivity and scale, suggesting that merging firms anticipate that complementarities between these pre-merger characteristics will generate surplus post merger. Our estimates of merger surplus are positively correlated with post-merger revenue productivity growth in the five years after merging.

The results allow an evaluation of the prominent view that mergers enable productivity growth by reallocating corporate assets to more effective management ([Jovanovic and Rousseau \(2002\)](#), [Nocke and Yeaple \(2008\)](#)). Such theories typically involve a post-merger production function complementarity between the assets of one firm and the superior productivity level of the other. Also implicit in these theories is a redundancy, in that the management skill of the less productive firm is replaced by the superior contribution made by its merger partner. By estimating the signs and relative magnitudes of various complementarities in anticipated merger surplus, we explore how well these theories explain the mergers in our data.

The empirical methods we apply are adapted from the large literature on matching in marriage markets based on [Becker \(1973\)](#) and are well suited for identifying different complementarities. Even when multiple dimensions of firm heterogeneity are available in the data, prior studies of mergers tend to evaluate matching on each dimension separately. For example, [David \(2020\)](#) establishes negative assortative matching on profitability, and then also on productivity, for merging firms in US mergers. [Braguinsky et al. \(2015\)](#) find that merging firms tend to be similarly productive but differ in profitability in their study of the Japanese spinning industry. Our data illustrate that looking at correlations in matching patterns in one dimension can be misleading when characteristics are correlated within firm. For example, in our data, firm scale and revenue productivity are negatively correlated. Given that the data show a positive correlation on productivities within mergers, the extent of matching between scale and productivity within merger would be underestimated in a framework that didn’t simultaneously account for the joint marginal distributions of firm-level characteristics.

While our findings show that complementarity in revenue productivities is the single most

important source of anticipated merger surplus, our understanding of the mechanism explaining this complementarity is limited because we cannot distinguish between efficiency-based TFP and the impact of prices on revenue productivity. We, hence, cannot tell whether merger gains arise from increasing efficiency or from increasing prices. However, several features of our results suggest that the ability to raise prices is unlikely to be the main explanation. The complementarity in revenue productivities is relatively stable for mergers of different total scales, and there is no evidence that this complementarity is greater when merging firms are closer competitors.

It is interesting that the complementarity in revenue productivities is largest for within-industry group mergers, both domestic and cross-border. We tentatively suggest that the results are consistent with expected knowledge exchange rather than knowledge transfer post merger. For domestic mergers across industry groups, however, the complementarity between revenue productivity and input scale is the single largest source of anticipated surplus. This finding implies that gains from reallocation within firm boundaries do play an important role in these types of mergers. While our industry groupings are too coarse to separate horizontal and vertical mergers, those between firms in different broad sectors are likely vertical in nature. One avenue for future work is to relate observed heterogeneity in sources of merger surplus to the business models of the larger organizations.

We also relate our results to some assumptions that are typically made about the sources of gains from FDI via M&A. Models in this literature often assume that the productivity of the most productive partner acts as a public good within enlarged firm boundaries. The strongest of our results suggests the opposite: that the capabilities of both merging firms play a role in anticipated surplus. This finding is in line with the fact that multinational firms from developed economies are more likely to undertake FDI via M&A rather than greenfield investment in other developed economies (Nocke and Yeaple (2008), where the characteristics of potential merger partners are likely to be more complementary with their own.

The value of the M&A transactions in the global market for corporate control regularly exceeds 4 trillion USD annually (imaa-institute.org). Our findings reveal that there is selection into merging based on the contributions of various firm characteristics, and that this selection is related to the future merger gains. Further work is needed to dig down into the microfoundations for the observed complementarity in revenue productivities that is separately identified in our results and shown to be the largest source of anticipated merger gains.

References

- Akerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica* 83 (6):2411–2451.
- Ashenfelter, Orley C, Daniel S Hosken, and Matthew C Weinberg. 2015. "Efficiencies brewed: pricing and consolidation in the US beer industry." *The RAND Journal of Economics* 46 (2):328–361.
- Asker, John and Volker Nocke. 2021. "Collusion, Mergers, and Related Antitrust Issues." Tech. rep., working paper, UCLA, 13 January 2021 (forthcoming, Handbook of Industrial .
- Becker, Gary S. 1973. "A theory of marriage: Part I." *Journal of Political economy* 81 (4):813–846.
- Bilir, L. Kamran and Eduardo Morales. 2020. "Innovation in the Global Firm." *Journal of Political Economy* 128 (4):1566–1625.
- Birkinshaw, Julian, Henrik Bresman, and Lars Håkanson. 2000. "Managing the post-acquisition integration process: How the human iintegration and task integration processes interact to foster value creation." *Journal of management studies* 37 (3):395–425.
- Blonigen, Bruce A, Lionel Fontagné, Nicholas Sly, and Farid Toubal. 2014. "Cherries for sale: The incidence and timing of cross-border M&A." *Journal of International Economics* 94 (2):341–357.
- Blonigen, Bruce A and Justin R Pierce. 2016. "Evidence for the effects of mergers on market power and efficiency." Tech. rep., National Bureau of Economic Research.
- Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson. 2015. "Acquisitions, productivity, and profitability: evidence from the Japanese cotton spinning industry." *American Economic Review* 105 (7):2086–2119.
- Chiappori, P.-A., B. Salanié, and Y. Weiss. 2017. "Partner Choice, Investment in Children, and the Marital College Premium." *American Economic Review* 107:2109–67.
- Chiappori, Pierre-Andre, Alfred Galichon, and Bernard Salanié. 2019. "On Human Capital and Team Stability." *Journal of Human Capital* 13:236–259.

- Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2012. “Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market.” *Journal of Political Economy* 120:659–695.
- Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2020. “Erratum: Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market.” *Journal of Political Economy* 128 (12):4673–4675.
- Choo, Eugene and Aloysius Siow. 2006. “Who marries whom and why.” *Journal of political Economy* 114 (1):175–201.
- Ciscato, Edoardo, Alfred Galichon, and Marion Goussé. 2020. “Like attract like? A structural comparison of homogamy across same-sex and different-sex households.” *Journal of Political Economy* 128 (2):740–781.
- Cohen, Wesley M and Daniel A Levinthal. 1990. “Absorptive capacity: A new perspective on learning and innovation.” *Administrative science quarterly* :128–152.
- Dagsvik, J.K. 1994. “Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes.” *Econometrica* 62 (5):1179–1205.
- David, Joel. 2020. “The aggregate implications of mergers and acquisitions.” *Review of Economic Studies* forthcoming.
- Deneckere, Raymond and Carl Davidson. 1985. “Incentives to Form Coalitions with Bertrand Competition.” *The RAND Journal of Economics* 16 (4):473–486.
- Dupuy, Arnaud and Alfred Galichon. 2014. “Personality traits and the marriage market.” *Journal of Political Economy* 122 (6):1271–1319.
- Farrell, Joseph and Carl Shapiro. 1990. “Horizontal Mergers: An Equilibrium Analysis.” *American Economic Review* 80 (1):107–126.
- Fox, Jeremy T. 2018. “Estimating matching games with transfers.” *Quantitative Economics* 9 (1):1–38.

- Gale, David and Lloyd S Shapley. 1962. "College admissions and the stability of marriage." *The American Mathematical Monthly* 69 (1):9–15.
- Galichon, Alfred and Bernard Salanié. 2019. "Labeling Dependence in Separable Matching Markets." *mimeo, Columbia* .
- . 2021. "Cupids invisible hand: Social surplus and identification in matching models." Forthcoming in the *Review of Economic Studies*.
- Geurts, Karen and Johannes Van Biesebroeck. 2019. "Employment growth following takeovers." *The RAND Journal of Economics* 50 (4):916–950.
- Guadalupe, Maria, Olga Kuzmina, and Catherine Thomas. 2012. "Innovation and foreign ownership." *American Economic Review* 102 (7):3594–3627.
- Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. 2004. "Export Versus FDI with Heterogeneous Firms." *American Economic Review* 94 (1):300–316.
- Javorcik, Beata and Jens Arnold. 2009. "Gifted Kids Or Pushy Parents? Foreign Acquisitions And Plant Performance In Indonesia." *Journal of International Economics* 79 (1):42–53.
- Jensen, Michael. 1986. "The Agency Cost of Free Cash Flow: Corporate Finance and Takeovers." *American Economic Review* 76:323–329.
- Jovanovic, Boyan and Peter L Rousseau. 2002. "The Q-theory of mergers." *American Economic Review* 92 (2):198–204.
- McGrattan, Ellen and Edward Prescott. 2009. "Openness, technology capital, and development." *Journal of Economic Theory* 144 (6):2454–2476.
- Mindruta, Denisa, Mahka Moeen, and Rajshree Agarwal. 2016. "A two-sided matching approach for partner selection and assessing complementarities in partners' attributes in inter-firm alliances." *Strategic Management Journal* 37 (1):206–231.
- Nocke, Volker and Stephen Yeaple. 2007. "Cross-border mergers and acquisitions vs. greenfield foreign direct investment: The role of firm heterogeneity." *Journal of International Economics* 72 (2):336–365.

- . 2008. “An assignment theory of foreign direct investment.” *The Review of Economic Studies* 75 (2):529–557.
- Ozcan, Yasin. 2015. “Innovation and acquisition: two-sided matching in M&A markets.” *Northwestern University* .
- Ramondo, Natalia and Andres Rodriguez-Clare. 2013. “Trade, Multinational Production, and the Gains from Openness.” *Journal of Political Economy* 121 (2).
- Reus, Taco H, Bruce T Lamont, and Kimberly M Ellis. 2016. “A darker side of knowledge transfer following international acquisitions.” *Strategic Management Journal* 37 (5):932–944.
- Rhodes-Kropf, Matthew and David T Robinson. 2008. “The market for mergers and the boundaries of the firm.” *The Journal of Finance* 63 (3):1169–1211.
- Roll, Richard. 1986. “The Hubris Hypothesis of Corporate Takeovers.” *Journal of Business* 59:197–216.
- Shapley, Lloyd S and Martin Shubik. 1971. “The assignment game I: The core.” *International Journal of game theory* 1 (1):111–130.
- Shleifer, Andrei and Robert W. Vishny. 1989. “Managerial Entrenchment: The Case of Manager-Specific Investments.” *Journal of Financial Economics* 25 (1):123–139.
- Stiebale, Joel. 2016. “Cross-border M&As and innovative activity of acquiring and target firms.” *Journal of International Economics* 99:1–15.

Appendix

A.1 Appendix: The Firms in the sample

Table 7 presents the cross tabulation of sample firms by country and industry group. The German and Italian firms are relatively more likely to be in secondary manufacturing and less likely to be in non-high-tech services while the reverse is true for UK firms. France has relatively more high-tech services and Spain has relatively few services firms in the sample.

Table 7: Joint Distribution of Firms by Country and Industry Group

	Total	Primary Manuf.	Sophisticated Manuf.	Other	High-Tech Serv.	Non High-Tech Serv.
Total	8,894	828	1,839	2,321	1,177	2,729
Germany	628	42	213	122	105	146
Spain	2,029	287	433	721	129	459
France	1,914	180	376	446	371	541
UK	3,312	212	516	789	426	1,369
Italy	1,011	107	301	243	146	214

The average firm has revenues of 531 million euros (in year-2000 euros), but the median is just under 26 million euros. We are interested in the heterogeneity in firm revenues at the time of merging because the components of this heterogeneity form the arguments of our surplus function. Table 8 shows the descriptive statistics of the firms in our sample.²⁹

German firms are 7% of the sample and are the largest, on average, with median revenues of 66 million euro. Italian firms, 11% of the sample, have median revenues of 44 million. UK firms make up 37% of the sample and have median revenues of 25 million. Spanish firms (23%) are at 22 million, and French firms (22%) have the smallest median revenues, at 19 million.

Non-high-tech services, which are all services other than telecommunications and information service activities, are the largest group, making up 31% of all firms, with median revenues of 17.5 million euros. The next largest industry group includes construction, utilities, wholesale, and retail firms, which are 26% of the sample, and we call this group other activities. These firms tend to be large, with median revenues of 38.1 million euros. The third largest group is what we call sophisticated manufacturing, at 21% of total with median revenues of 31.0 million. The two smallest groups are high-tech services, 13% of firms with 16.6 million euros in revenues, and

²⁹To be included in our sample, firms need to report revenues, assets and employment pre-merger. Therefore, the deals dropped due to incomplete information often correspond to those that involve small firms. Around 13% percent of firms take part in more than one bilateral merger in the data. We treat these cases as if they were different firms.

Table 8: Firms. Descriptive Statistics

	N firms	$\ln(\textit{Revenues})$		Input Scale		TFPr		$\beta(\text{Scale on TFPr})$
		Mean	s.d.	Mean	s.d.	Mean	s.d.	
Panel A: by Country								
Total	8,894	10.27	2.29	7.69	2.30	2.58	1.04	-0.11
Germany	628	11.26	2.22	8.68	2.23	2.58	1.06	-0.14
Spain	2,029	10.08	2.29	7.44	2.08	2.63	1.04	-0.07
France	1,914	9.94	2.09	7.30	2.16	2.63	0.93	-0.13
UK	3,312	10.26	2.44	7.74	2.48	2.51	1.05	-0.11
Italy	1,011	10.70	2.18	8.10	2.18	2.57	1.18	-0.17
Panel B: by Industry Group								
Total	8,894	10.27	2.29	7.69	2.30	2.58	1.04	-0.11
Primary Manuf.	828	10.42	2.11	7.48	1.95	2.91	1.03	-0.08
Sophisticated Manuf.	1,839	10.47	2.06	7.78	2.01	2.70	0.84	-0.07
Other Activities	2,321	10.74	2.34	7.93	2.42	2.80	1.07	-0.13
High-Tech Services	1,177	9.81	2.47	7.51	2.60	2.30	0.89	-0.11
Non-High-Tech Serv.	2,729	9.88	2.28	7.56	2.33	2.32	1.11	-0.14

primary manufacturing, at 9% of firms with 32.5 million euros.³⁰

Despite the large differences across country-average and across industry-group-average revenues, country and industry group explain very little of the overall variation in firm revenues in the sample. A regression of the log of firm revenues on the interaction of country and industry group has an R-squared of only 0.02. Figure 5 compares the histogram of the distribution of the log of firm revenues demeaned by the sample average, by the relevant country mean, and by the relevant industry-group mean, and shows that most of the variation in revenues is within country-industry groups.

A.2 Appendix: the single-index test

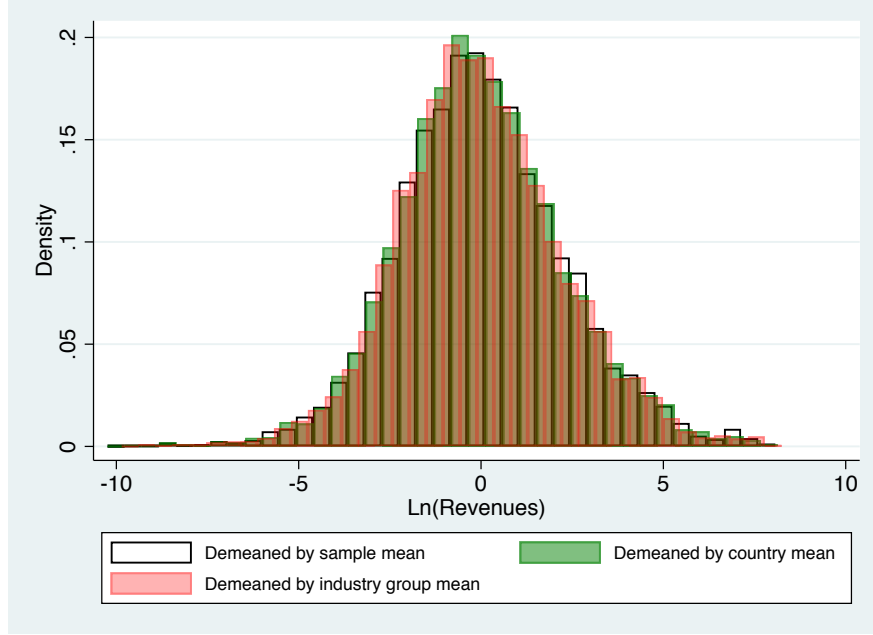
In Section 2, we ran a test of whether the two continuous characteristics z and Ω can be aggregated into a single combined characteristic—an index. We now show that while this test is in general invalid, it applies in our setting.

For notational simplicity, we condition on discrete characteristics on both sides and we omit them from the notation. Suppose that there exist *scalar* indices $I(x)$ and $J(y)$ such that

$$\Phi(x, y) = F(I(x), J(y))$$

³⁰The NACE code groupings are, respectively: all NACE codes between 54 and 99, excluding those between 61 and 63; NACE codes 34 to 53, inclusive; NACE codes 18 to 33, inclusive; NACE codes 61 to 63, inclusive; and two-digit NACE codes below 17.

Figure 5: Distribution of the log of Firm Revenues, demeaned by sample mean, by country mean, and by industry group mean



for some function F . Proposition 2 of [Chiappori, Orefice, and Quintana-Domeque \(2012\)](#) implies that, for any stable marriage market matching, the conditional distribution of $I(x)$ given Y only depends on the index $J(y)$. However, this property may not extend to the characteristics X themselves: one can find examples for which the conditional distribution of some (or all) of the components of the vector x , given y , does not depend only on $J(y)$ ([Chiappori, Orefice, and Quintana-Domeque \(2020\)](#)).

Nonetheless, the underlying multinomial logit structure of the model developed in Section 3 makes the test valid in our case. We know that in the logit specification,

$$\mu(x, y)^2 = \mu(x, 0)\mu(0, y) \exp(\Phi(x, y))$$

which gives, for any component of x (in our case, $k = 1, 2$)

$$\begin{aligned} E(x_k|y) &= \frac{\int x_k \mu(x, y) dx}{\int \mu(x, y) dx} \\ &= \frac{\int x_k \sqrt{\mu(x, 0)} \exp(F(x, J(y))/2) dx}{\int \sqrt{\mu(x, 0)} \exp(F(x, J(y))/2) dx}, \end{aligned} \tag{A.1}$$

where we canceled out the terms $\sqrt{\mu(0, y)}$ in the numerator and the denominator. The fraction in

equation (A.1) is a function of $J(y)$ only; therefore

$$E(x_k|y) = G_k(J(y))$$

for some function G_k . As a consequence, if we run flexible regressions of x_k and x_l on y the same $J(y)$ should show up in both. This proof obviously carries over to the distribution of y conditional on x . On the other hand, it relies heavily on the multiplicative formula for $\mu(x, y)$, which fails even in simple extensions of the Choo and Siow (2006) model (for instance if the error terms have standard errors that depend on the types x and y).

A.3 Appendix: Numerical Methods

This appendix gives more detail on our implementation of the methods presented in the main text.

A.3.1 Change of Variables

As explained in the text, we transformed each of our continuous variables to a standard normal. This is done simply by matching quantiles. If a scalar variable T has a cdf F_T , we associate to it the variable t defined by

$$F_n(t) = F_T(T)$$

where F_n is the cdf of $N(0, 1)$. By construction, the transformed variable t is distributed as $N(0, 1)$.

We apply this transformation within each country-industry cell a , replacing F_T with the empirical cdf of a component of x in a cell a . For the productivity Z for instance, we define

$$F_n(z) = \hat{F}_{Z|a}(Z|a).$$

In the rest of this Appendix we will use $(Z, \bar{\Omega})$ to denote the original variables, and (z, Ω) for the transformed variables.

A.3.2 Numerical Integration

We need to approximate two types of integrals:

1. $\mathcal{I} = \sum_{a,b} \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy$
2. $\mathcal{J}_a = \int g(a, x) dx$

where $\hat{\mu}_{ab}(x, y)$ is estimated from the data and the notation x refers to the pair of continuous variables (Ω, z) that result from applying the transformation above to the size and productivity variables within a cell.

The sum of four-dimensional integrals $\mathcal{I}$ seems forbidding, but we can simply approximate it with its empirical analog. Averaging over all M observed mergers and dividing by two to avoid double-counting gives us

$$\hat{\mathcal{I}} = \frac{1}{2M} \sum_{i=1}^M \phi_{a_i, b_i}(x_i, y_i).$$

To evaluate the double integral $\mathcal{J}_a$, we start with the simpler, one-dimensional integral over the transformed productivity variable: $\mathcal{K}_a = \int p(a, z) dz$. We first note that since $F_n(z) = F_{Z|a}(Z|a)$, taking derivatives gives us $f_n(z) dz = f_{Z|a}(Z|a) dZ$, where the f notation refers to the pdfs associated with the distributions. Denoting $P(a, Z) = p(a, z)$, we can rewrite the integral as

$$\mathcal{K}_a = \int p(a, z) dz = \int P(a, Z) dZ = \int \frac{P(a, Z)}{f_{Z|a}(Z|a)} f_n(z) dz,$$

which suggests using Gauss-Hermite quadrature. To do so, we choose a number of nodes m (we used eight nodes and we checked that going to 16 did not change our results). Let the associated nodes and weights be (n_k, w_k) for $k = 1, \dots, m$. Define $q_k \equiv \sqrt{2}n_k$. The Gauss-Hermite quadrature formula can be written as

$$\int h(a, z) f_n(z) dz \simeq \frac{1}{\sqrt{\pi}} \sum_{k=1}^m w_k h(a, q_k)$$

for any function $h(a, z)$. Changing variables³¹, we define nodes Z_k^a by

$$F_{Z|a}(Z_k^a|a) = F_n(q_k)$$

and we approximate the integral with

$$\hat{\mathcal{K}}_a = \frac{1}{\sqrt{\pi}} \sum_{k=1}^m w_k \frac{P(a, Z_k^a)}{f_{Z|a}(Z_k^a|a)} = \frac{1}{\sqrt{\pi}} \sum_{k=1}^m w_k \frac{p(a, q_k)}{f_{Z|a}(Z_k^a|a)}.$$

To approximate $\mathcal{J}_a$, we only need to use this approximation scheme on both continuous dimensions:

$$\hat{\mathcal{J}}_a = \frac{1}{\pi} \sum_{k, k'=1}^m w_k w_{k'} \frac{g(a, q_k, q_{k'})}{f_{\bar{\Omega}}(\bar{\Omega}_k^a|a) f_Z(Z_{k'}^a|a)}.$$

Note for further reference that this is also

$$\hat{\mathcal{J}}_a = \sum_{k, k'=1}^m W_{kk'}^a G(a, \bar{\Omega}_k^a, Z_{k'}^a) \tag{A.2}$$

³¹This is simply reversing the transform in A.3.1.

where the weights $W_{kk'}^a$ do not depend on the integrand g and can be computed and stored in advance.

A.3.3 Estimating Densities

To apply the formulæ in Section A.3.2, we need estimators of the conditional cdfs and pdfs of $\bar{\Omega}$ and Z within each cell a . In addition, we will need an estimator of their joint conditional density to run the IPFP algorithm.

While our largest cell has 1,369 firms (GB, non-high-tech services), the smallest one has only 42 (DE, basic manufacturing); this is too little for a nonparametric estimator. We only used a kernel density estimator in a cells of size larger than 200. In smaller cells we fitted a normal (univariate or bivariate) distribution. This is not such a bad approximation as the variables Ω and z are logs of productivity and size.

In the following we denote $\bar{f}(x|a)$ the estimated joint density of the transformed variables (Ω, z) within cell a .

A.3.4 The IPFP Algorithm

To solve for the $U_{ab}(x, y)$ functions, we apply the Iterative Proportional Fitting Procedure. To do this, we first note that, taking logs on equation (11), we get:

$$\log \mu_{ab}(x, y) = U_{ab}(x, y) + k_a(x) = \Phi_{ab} - U_{ab}(x, y) + k_b(y),$$

which we can rewrite as:

$$2U_{ab}(x, y) = \Phi_{ab}(x, y) + k_b(y) - k_a(x)$$

and

$$\mu_{ab}(x, y) = \exp((\Phi_{ab}(x, y) + k_a(x) + k_b(y))/2).$$

Denoting $H_{ab}(x, y) = \exp(\Phi_{ab}(x, y)/2)$ and $K_a(x) = \exp(k_a(x)/2)$, the feasibility constraints become

$$\bar{N}_a \bar{f}(x|a) = \sum_{b \in EU5} \int \mu_{ab}(x, y) dy = K_a(x) \sum_{b \in EU5} \int H_{ab}(x, y) K_b(y) dy. \quad (\text{A.3})$$

Solving for the equilibrium $U_{ab}(x, y)$ is equivalent to solving (A.3) for the functions K_a . To do so, we convert the equation into an iterative process: at step p , we approximate

$$K_a^{(p)}(x) = \frac{\bar{N}_a \bar{f}(x|a)}{\sum_{b \in EU5} \int H_{ab}(x, y) K_b^{(p-1)}(y) dy}$$

by using Gaussian quadrature at the nodes, using (A.2):

$$K_a^{(p)}(k, k') = \frac{\bar{N}_a \bar{f}(q_k, q_{k'} | a)}{\sum_{b \in EU5} \sum_{l, l'=1}^m W_{ll'}^b H_{ab}(\bar{\Omega}_k^a, Z_{k'}^a, \bar{\Omega}_l^b, Z_{l'}^b) K_b^{(p-1)}(l, l')}.$$

This looks more complicated than it is: all terms $W_{ll'}^b H_{ab}(\bar{\Omega}_k^a, Z_{k'}^a, \bar{\Omega}_l^b, Z_{l'}^b)$ can be precomputed, and each iteration only involves one matrix multiplication and one element-by-element division.

The convergence of IPFP in this case is a consequence of more general results in Galichon and Salanié (2021). After convergence, we take $k_a(q_l, q_{l'}) = 2 \log K_a(l, l')$.

A.3.5 The Moment Matching Estimator

To estimate the parameter vector λ , we need to minimize the function

$$\mathcal{M}(\lambda) = \mathcal{W}(U^\lambda) - \frac{\lambda}{2} \cdot \sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy$$

as explained in Section 3.3. The value of $\mathcal{W}(U^\lambda)$ is the sum of the expected utilities of the different types, which we wrote in the text as

$$\sum_{a \in EU5} \bar{N}_a \int \log \left(\sum_{b \in EU5} \int \exp(U_{ab}^\lambda(x, y)) dy \right) \bar{f}(x|a) dx.$$

The log term is simply $u_a(x)$; using the formulæ in Section A.3.4, we can replace it with

$$\log \left(\sum_{b \in EU5} \int H_{ab}^\lambda(x, y) \frac{K_b^\lambda(y)}{K_a^\lambda(x)} dy \right)$$

which, given (A.3), is simply $\log(\bar{N}_a \bar{f}(x|a)) - \log K_a^\lambda(x) = \log(\bar{N}_a \bar{f}(x|a)) - k_a^\lambda(x)/2$. Collecting terms gives

$$\begin{aligned} \mathcal{M}(\lambda) &= \sum_{a \in EU5} \bar{N}_a \int \bar{f}(x|a) \log(\bar{N}_a \bar{f}(x|a)) dx \\ &\quad - \sum_{a \in EU5} \bar{N}_a \int \bar{f}(x|a) \frac{k_a^\lambda(x)}{2} dx \\ &\quad - \frac{\lambda}{2} \cdot \sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy, \end{aligned}$$

so that we only need to choose λ to *maximize*

$$\mathcal{G}(\lambda) \equiv \sum_{a \in EU5} \bar{N}_a \int \bar{f}(x|a) k_a^\lambda(x) dx + \lambda \cdot \sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy.$$

To take it to the data, we replace all unknown quantities with estimators. The sum of the double integrals can be approximated by an average over the observed mergers, as we did with $\hat{\mathcal{T}}$ in Section A.3.2:

$$\sum_{a,b=1}^M \int \int \phi_{ab}(x, y) \hat{\mu}_{ab}(x, y) dx dy \simeq \sum_{i=1}^M \phi_{a_i b_i}(x_i, y_i).$$

The IPFP algorithm of Section A.3.4 gives us the value of $k_a^\lambda = 2 \log K_a^\lambda$ at the quadrature nodes and we approximate the first term of $\mathcal{G}(\lambda)$ using (A.2):

$$\int \bar{f}(x|a) k_a^\lambda(x) dx \simeq \sum_{l,l'=1}^m W_{ll'}^a \hat{f}(q_l, q_{l'}|a) k_a^\lambda(q_l, q_{l'}).$$

This gives

$$\hat{\mathcal{G}}(\lambda) = \sum_{a \in EU5} \hat{N}_a \sum_{l,l'=1}^m W_{ll'}^a \hat{f}(q_l, q_{l'}|a) k_a^\lambda(q_l, q_{l'}) + \lambda \cdot \sum_{i=1}^M \phi_{a_i b_i}(x_i, y_i).$$

We obtain the standard errors by the usual device: we write the first-order conditions

$$\begin{aligned} 0 &= \frac{1}{\sqrt{M}} \hat{\mathcal{G}}'(\hat{\lambda}) \\ &\simeq \frac{1}{\sqrt{M}} \hat{\mathcal{G}}'(\lambda_0) \\ &\quad + \frac{1}{M} \hat{\mathcal{G}}''(\lambda_0) \times \sqrt{M}(\hat{\lambda} - \lambda_0) \end{aligned}$$

which gives an asymptotic distribution

$$\sqrt{M}(\hat{\lambda} - \lambda_0) \simeq N(0, \mathbf{J}^{-1} \mathbf{I} \mathbf{J}^{-1})$$

with

$$\begin{aligned} \mathbf{I} &= \lim_{M \rightarrow \infty} \frac{V \hat{\mathcal{G}}'(\lambda_0)}{M} \\ \mathbf{J} &= \text{plim}_{M \rightarrow \infty} \frac{\hat{\mathcal{G}}''(\lambda_0)}{M}. \end{aligned}$$

The matrix $\hat{\mathcal{G}}'(\boldsymbol{\lambda}_0)/M$ consists of two terms. Its variance comes from the second term,

$$\hat{E}\boldsymbol{\phi} = \frac{1}{M} \sum_{i=1}^M \boldsymbol{\phi}_{a_i b_i}(x_i, y_i)$$

whose variance can be estimated as

$$\frac{1}{M} \sum_{i=1}^M (\boldsymbol{\phi}_{a_i b_i}(x_i, y_i) - \hat{E}\boldsymbol{\phi})(\boldsymbol{\phi}_{a_i b_i}(x_i, y_i) - \hat{E}\boldsymbol{\phi})'.$$

The matrix $\mathbf{J}$ can simply be estimated by evaluating the Hessian of $\hat{\mathcal{G}}$ at $\hat{\boldsymbol{\lambda}}$.

A.4 Appendix: Equilibrium conditions using only observed matches

Our presentation of the model in Section 3 described a world market for mergers based on a representative sample of firms. Our data, however, only has observed mergers between firms in five EU countries (EU5); it excludes firms from other countries, as well as EU-5 firms that remain standalone.

As a consequence, we observe neither N_a nor $f(x|a)$. Instead, we observe $\bar{N}_a$ and $\bar{f}(x|a)$ when $a \in EU5$: the number of firms in within-EU5 mergers and the distribution of their continuous characteristics. Since we observe all mergers between EU5 firms ($a, b \in EU5$), our data does give us $\mu_{ab}(x, y)$, the product of the number of mergers μ_{ab} between firms of discrete characteristics a and b and the conditional joint pdf of their continuous characteristics $\mu(x, y|a, b)$.

For $a \in EU5$, the relationship between these latent and observed distributions of characteristics is given by subtracting the number of firms that do not participate in mergers and those that merge with a non-EU5 firm:

$$\bar{N}_a \bar{f}(x, a) = N_a f(x|a)(1 - \mu(\emptyset|a, x)) - \sum_{b \notin EU5} \int \mu_{ab}(x, y) dy.$$

Now take a and b both in EU5. Denoting $I_{ad}(x) \equiv \int \exp(U_{ad}(x, y)) dy$, we can write

$$\mu(b, y|a, x) = \frac{\exp(U_{ab}(x, y))}{1 + \sum_{d \notin EU5} I_{ad}(x) + \sum_{d \in EU5} I_{ad}(x)}.$$

Therefore

$$\begin{aligned}
\mu_{ab}(x, y) &= N_a f(x|a) \times \mu(b, y|a, x) \\
&= \frac{\exp(U_{ab}(x, y))}{\sum_{d \in EU5} I_{ad}(x)} \times N_a f(x|a) \times \frac{\sum_{d \in EU5} I_{ad}(x)}{1 + \sum_{d \notin EU5} I_{ad}(x) + \sum_{d \in EU5} I_{ad}(x)} \\
&= \frac{\exp(U_{ab}(x, y))}{\sum_{d \in EU5} I_{ad}(x)} \times N_a f(x|a) \times \mu(EU5|a, x) \\
&= \frac{\exp(U_{ab}(x, y))}{\sum_{d \in EU5} I_{ad}(x)} \times \bar{N}_a \bar{f}(x|a).
\end{aligned}$$

Taking logarithms gives an expression of the form $\log \mu_{ab}(x, y) = U_{ab}(x, y) + k_a(x)$. Since $\mu_{ab}(x, y) = \mu_{ba}(y, x)$ and $U_{ab}(x, y) + U_{ba}(y, x) = \Phi_{ab}(x, y)$, we obtain

$$2 \log \mu_{ab}(x, y) = \Phi_{ab}(x, y) + k_a(x) + k_b(y). \quad (\text{A.4})$$

While we cannot identify the functions k_a , (A.4) tells us that (given enough data) we can identify the functions $U_{ab}(x, y)$ nonparametrically up to additive terms. For instance,

$$\Phi_{ab}(x, y) - \Phi_{ab'}(x, y') - \Phi_{a'b}(x', y) + \Phi_{a'b'}(x', y') = 2 \log \frac{\mu_{ab}(x, y) \mu_{a'b'}(x', y')}{\mu_{ab'}(x, y') \mu_{a'b}(x', y)}$$

is a simple function of the data. With our parametric specification $\Phi_{ab}(x, y) = \sum_{k=1}^K \lambda_k \phi_{ab}^k(x, y)$, we will only be able to identify a parameter λ_k if the double difference

$$\phi_{ab}^k(x, y) - \phi_{ab'}^k(x, y') - \phi_{a'b}^k(x', y) + \phi_{a'b'}^k(x', y')$$

is not constant. This only rules out basis functions of the form $\phi_a(x)$.

The set of equilibrium conditions (7) also needs to be modified: it becomes

$$\bar{N}_a \bar{f}(x|a) \frac{\exp(U_{ab}(x, y))}{\sum_{d \in EU5} \int \exp(U_{ad}(x, t)) dt} = \bar{N}_b \bar{f}(y|b) \frac{\exp(\Phi_{ab}(x, y) - U_{ab}(x, y))}{\sum_{c \in EU5} \int \exp(\Phi_{cb}(z, y) - U_{cb}(z, y)) dz}. \quad (\text{A.5})$$

Finally, the moment matching estimator is still valid, with the only difference that the welfare function becomes

$$\bar{\mathcal{W}}(U) \equiv \sum_a \bar{N}_a \bar{f}(x|a) \log \left(\sum_{b \in EU5} \exp(U_{ab}(x, y)) dy \right) :$$

the expected utility of each firm only includes its share of the surplus from mergers with EU5 firms.

A.5 Appendix: Post-Merger Performance Analysis

Table 9 presents the estimation results plotted in Figure 4. The first set of $\hat{\beta}$ estimates is the quadratic specification with no interactions by discrete characteristic group. The second set calculates $\hat{\Phi}$ using the discrete characteristics interactions allowing the coefficients determining predicted merger surplus to vary with whether the merging firms are in the same country or not and by whether or not they are in the same industry.

Table 9: Regression coefficients, growth between year 0 and year t relative to counterfactual, regressed on the identified merger surplus terms and control functions.

Years	(1)	(2)	(3)	(4)	(5)
Quadratic specification $\hat{\Phi}$, as per Figure 1	0.040*** (0.015)	0.036** (0.017)	0.040* (0.021)	0.026 (0.026)	0.050* (0.029)
Constant	0.086 (0.143)	0.343** (0.158)	-0.247 (0.258)	0.051 (0.309)	-0.674** (0.320)
Observations	1375	1202	1051	939	841
R^2	0.156	0.171	0.188	0.176	0.192
F	4.601	4.482	4.349	3.572	3.520
p	0.000	0.000	0.000	0.000	0.000
Quadratic specification $\hat{\Phi}$, as per Figure 1	0.034*** (0.011)	0.023* (0.012)	0.025 (0.015)	0.041** (0.019)	0.041** (0.020)
Constant	0.105 (0.143)	0.354** (0.159)	-0.245 (0.258)	0.064 (0.308)	-0.681** (0.320)
Observations	1375	1202	1051	939	841
R^2	0.157	0.171	0.187	0.180	0.193
F	4.658	4.457	4.328	3.655	3.544
p	0.000	0.000	0.000	0.000	0.000

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$