Introducing the Business Relations Survey

Gábor Békés[a], Miklós Koren[b], Balázs Muraközy[c,*], Álmos Telegdy[d]

[a] *Central European University, CERS and CEPR*

[b] *Central European University, CERS and CEPR*

[c] *University of Liverpool, CERS*

[d] *National Bank of Hungary and Corvinus University of Budapest*

Abstract

We introduce the *Business Relations Survey* that collected information about buyer and supplier portfolios of over 1500 firms in three European countries.

keywords: SMEs, global value chain, supplier-buyer relationship, Business Relations Survey, relationship strength, innovation, Hungary, Romania, Slovakia

JEL-codes: C83, D22, D23

1. Measuring the nature of buyer-supplier relationships: The Business Relations Survey

This section describes the The Business Relations Survey (earlier called Supplier Survey), which aimed at measuring the nature of supplier-buyer relationships including

*Corresponding author: Balázs Muraközy, University of Liverpool Management School, Chatham St, Liverpool L69 7ZH United Kingdom

**The authors gratefully acknowledge funding from the European Research Council (ERC Starting Grant agreement number 313164) and the European Union's Horizon 2020 research and innovation programme (under grant agreement No 822390, the Microprod project). We thank Kinga Ritter, Imola Csóka, Gergely Kiss, Krisztián Fekete and András Vereckei for excellent research assistance. The views expressed in this paper are those of the authors and do not necessary reflect the official view of the National Bank of Hungary.

their number, share and strength. In Subsection 1.1 we discuss the main aim and approach of the survey, in Subsection 3 we introduce some of the key questions in the survey and in Subsection 3.1 we characterize our sample.

1.1. Key aims and approach

A limitation of the reviewed literature is that relatively little is known about the strength of linkages between firms. Administrative data typically do not include such information and most of the surveys we have reviewed focus on other questions.

Our aim was to design a method which would be informative along this dimension. The literature on supplier-buyer relationships suggests that both firm fundamentals (productivity, size, etc), type of the product (specific or not) and strategic choices matter. Therefore, qualitative data on links and administrative balance sheet data strongly complement each other. In order to access both types of information, we have designed a procedure that allows us to link survey answers to key financial, industry and ownership information at the firm level.

Another key decision was to conduct a multi-country survey. One motivation was to strengthen the external validity of the analysis with multi-country research. A more practical motivation was that given the typical response rate for such surveys – 10-15% – it was unlikely that we will be able to collect a large enough sample from a single small-sized country.

When choosing the group of countries, an important decision was whether to conduct the survey in relatively similar or rather different countries. We opted to choose similar countries to maximize the power of the statistical analysis. Our choice of the countries was mainly motivated by the fact that participation in Global Supply Chains is prevalent in Central and Eastern Europe (**?**). Many multinational firms have manufacturing affiliates in these countries that are assembling parts or making consumer products from white goods to cars. Many local firms produce parts used in assembly lines all over Europe and beyond. These countries are fully integrated not only to EU value chains but are also part of many global operations.

In terms of practical details, we used a computer assisted personal interview, assuming that this provides a good balance between depth and response rate. We trained surveyors, run pilots and had also conducted meetings with company managers to make sure questions are clear and comparable across countries. As a compromise between depth and length, we asked some basic questions about all partners and asked a set of detailed questions only about a limited set of *key partners*, defined as having at least 10% share in sales/material costs.

2. The survey process

2.1. Design

In terms of practical details, we used a computer assisted personal interview, assuming that this provides a good balance between depth and response rate. We trained surveyors, run pilots and had also conducted meetings with company managers to make sure questions are clear and comparable across countries. We limited questions to ensure that the survey should last no longer than 45 minutes as a key practical constraint was to make sure that managers are willing to answer our questions.

As a compromise between depth and length, we asked some basic questions about all partners and asked a set of detailed questions only about a limited set of key partners. Indeed, the most substantial trade-off proved to be between the number of suppliers and buyers we learn about and the depth of the information about each of them. As a compromise, we asked a few questions about all suppliers and buyers and asked the details only about *key partners*, defined as having at least 10% share in sales/material costs.[1] This definition worked very well and firms could easily identify these partners. We also allowed the managers to name a special supplier/buyer, even if that partner was not among the largest ones. Such special partner can be the oldest partner, a firm that is a good reference, or a foreign partner.

The most significant question regarding key partners was the identity of these firms. To maximize the accuracy of data we asked the firms to provide the EU VAT number of their partners. We needed this to merge the information with the Amadeus dataset for further analysis involving financial data. As this information is usually not available on the spot, we asked managers to supply this information after the survey. Unfortunately, few managers were ready to supply the EU VAT number either during or after the interview. Consequently, we got very sparse coverage here. We will hence have to complete name-based matching of partner firms.

Given the time constraints and managers' reluctance to supply financial information (**?**), we included very few questions on firm finances, especially if that was available from other sources. One of the exceptions was asking about total and export revenue. We compared this information to data available from the balance sheet information to double check whether the interviewee belongs to the firm we think.

Another challenge we faced was the different jargon – and to some extent logic – used by researchers and managers. We had to make sure that the interviewees understood the

[1] In case the firm did not have three partners above the 10% threshold, we asked for the three with the highest shares.

questions the same way as we did. We started with a few longer open-ended interviews. These preliminary interviews played an essential role in designing the first draft of the survey. After the draft was ready, we conducted pilot interviews with a diverse set of firms in two rounds, testing two versions of the questionnaire. Pilots were essential in the design of the final version of the questions. Based on feedback from the pilot, we simplified some questions. For example, managers typically were able to name the main industry of the partner, but could not provide much detail in this respect. We also added explanations to questions and answers to make sure respondents understand the them equally well. This was especially important for questions on innovation and learning.

A specific concern with the multi-country study was making sure that the questions are similarly interpreted by the managers in different countries. It was essential to provide a chance for the managers to answer in their native language. In order to do so, we relied on translators with an extensive knowledge of this type of work and their work was reviewed both by the GfK and by other academics. We also conducted the interviews in English when that was the main language of the manager.

2.2. Data handling and cleaning

This section explains the basic structure of the dataset we created from the survey. First it describes our measures to guarantee data security followed by the cleaning process we implemented.

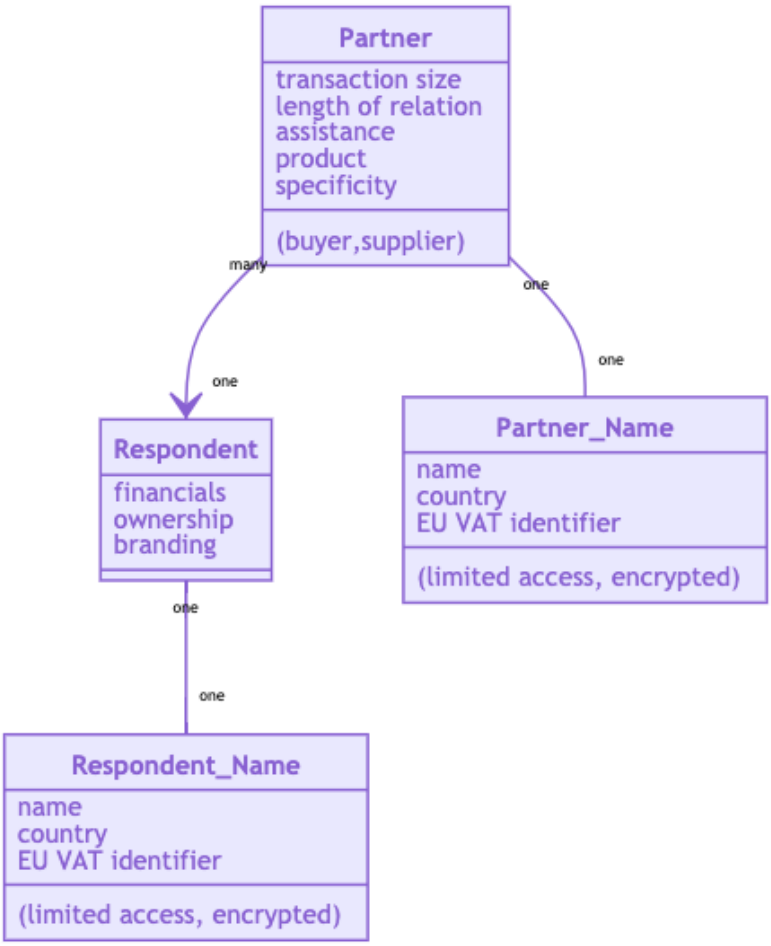
Owing to the many levels of questions, the Survey yields a somewhat complex data structure. Most importantly, we have one questionnaire for each respondent, but each respondent could name a number of suppliers and buyers, also generating information at the respondent-partner pair level. To increase efficiency, we separated these two types of information into separate files. For security reasons, we also store company identifiers separately from our main datasets. Figure 1 illustrates the logical model of our dataset with an Entity-Relation Diagram.

The first two files include respondent-level information. The "Respondent" file contains the anonymized respondent-level information from the questionnaire and the balance sheet data, while the "Respondent name" file includes the name and the identifier of the respondent. The latter file is limited access and encrypted.

The other two files include information at the respondent-partner level. The file named "Partner" includes the respondent-partner specific answers from the questionnaire and "Partner name" includes the name and identifier of the partner.[2]

[2]Note that one partner can belong to several respondents in principle. This information is critical

Figure 1: Logical Model of the Business Relations Survey Dataset



Notes: This figure shows the Entity-Relation Diagram of our analysis dataset. Each respondent can have multiple partners reported. The names of respondents and partners have been separated out in tables and handled separately to preserve anonymity. The graph gives a sample of information known about each entity, not a full list of variables.

2.3. Harmonization

Thanks to the efforts described in Subsections **??** and **??**, questionnaires were remarkably consistent across countries and answers were clear so relatively little data cleaning was necessary. The main steps were the following:

---

when drawing the network. We are currently working on identifying firms which were reported by multiple respondents as partners.

1. All values were given in national currency or a currency the company used (e.g., USD). To harmonize this information across countries and firms, all such variables were converted to 1,000 Euros.

2. We corrected partner index numbering to ensure that the largest customer is indeed the one with the largest share (as some respondents mentioned their second or third largest customer first). We created a new variable named "share rank" containing the actual ordering for partners for which the share was available. We added a flag when the share was missing. This change affected about 3-4% of rankings (both for buyers and sellers).

3. We checked if shares satisfied basic algebraic constraints, such as being less than 100% and summing up to less than 100%. There were only few such mistakes and they did not appear to be systematic.

4. We corrected every type of duplication in the list of partners. Most importantly, in a number of cases the "special" buyer/supplier mentioned by the respondent had already been mentioned among the most important partners. We cleaned this information by deleting the duplicates. We used both the name and share information to identify these instances, mainly in an automated way, but also checked some suspicious cases manually. This affected less than 5% of observations.

2.4. Survey quality test

A key concern with these kind of surveys is whether some surveyors make less effort to gather reliable answers. surveyor features affect the results. As important as it is, such an effect is actually not easy to verify. As regions and firms in these countries vary, replies to questions may also vary across respondents. Furthermore, surveyors are different in experience and hence response time may vary for completely honest reasons, too.

Nevertheless, we conducted several checks to investigate this and related issues. We looked for correlations between the quality of the response data and metadata. A simple measure of quality is item non-response: how many questions remain unanswered in the questionnaire. Survey metadata includes an anonymous identifier of the surveyor and the precise time stamp of starting and finishing the survey. These were recorded automatically in the interview software.

We identified several surveyors who were very different from the vast majority in terms of the number of interviews conducted (too many), the time gap between surveys (too short), or the length of the interview (too short). However, interviews conducted by these surveyors did not differ significantly in terms of data quality indicators such as the prevalence of item nonresponse. We could not rule out that these interviews

were, in fact, conducted in the same fashion as others, with data entry into the survey application happening at a later stage.

3. Key questions on supplier-buyer relationships

In this section, we present some key questions from the questionnaire to illustrate the type of information available from the survey.

The questionnaire starts with questions about the firm itself. Table 1 includes a number of examples. These questions were typically answered by most firms, with some of them unable or unwilling to provide basic financial information. Questions here include basic information about the partners, including the number of buyers and suppliers.

Table 1: Firm specific questions

Question	Resp.rate	Question	Resp.rate
Number of employees	100%	Owner nationality	100%
Total sales in 2015	75%	High-value machines/equipment in the production process?	100%
How many buyers did your company have in 2015?	100%	Who decides typically on a new, large contract with a buyer?	100%

The table displays a sample of questions with their respective response rates. *Source: Business Relations Survey.*

After asking this basic information and learning about all partners, the questionnaire focused on *key partners*. Some of these questions are reported in Table 2. The basic questions, such as the location and industry of the partner and the age of the relationship were symmetric for buyers and suppliers. As the table shows, response rates were very high for these basic information questions. We measured the importance of suppliers relative to total purchase costs and that of buyers relative to the sales of the respondent. To make the questions less sensitive, we asked about ratios rather than the absolute values of sales and purchases. The response rate for this question was somewhat lower than for the basic questions, but still above 85%. If the manager was willing to specify a partner, in most cases she was also willing to supply us with this basic information about the relationship.

Table 2: Basic buyer/supplier questions

Buyer	Resp.rate	Supplier	Resp.rate
Where is the headquarters of the buyer located?	100%	Where is the headquarters of the supplier located?	100%
What is the buyer's main business activity?	100%	What is the supplier's main business activity?	100%
Number of years selling products to this buyer?	96%	How long have you been making purchases from this supplier?	95%
What share of your sales comes from this buyer?	89%	Share of overall purchase costs goes to this supplier?	87%

The table displays a sample of questions with their respective response rates. *Source: Business Relations Survey.*

When asking more detailed questions, in line with the typology discussed in Subsection **??**, we focused on the nature of collaboration and types of products they trade. A few examples are shown in Table 3. Again, the response rate for describing the product and the substitutability of the product was high, with a somewhat lower response rate about cooperation and whether the product is critical. The lower response rate may partly be due to the sensitivity of these questions, but most likely the interviewees did not have the necessary information to answer these questions.

Table 3: Detailed questions on the relationship

Buyer	Resp.rate	Supplier	Resp.rate
Name the most important product sold to this buyer?	100%	Name the most important product bought from the supplier?	100%
Buyer performed regular checking in the last two years?	88%	Product of the sup. critical in your production process?	79%
Easy to sell an almost identical product to buyer?	100%	Easy to buy an almost identicial product from an. supp.?	100%

The table displays a sample of questions with their respective response rates. *Source: Business Relations Survey.*

Beyond buyers and suppliers, a part of the questionnaire was devoted to understanding high valued machinery use as shown in Table 4. Again, the response rates were very high, with the exception of the age of the machine. Most likely the reason for not answering was the lack of information rather than the sensitivity of this question.

Table 4: Questions on high value machinery

Question	Resp.rate	Question	Resp.rate
Where is the headquarters of the manufacturer located?	100%	Quality of the machine on a 5 point scale	100%
The name of the machine	94%	Train existing emp. to successf. operate this machine?	99%
When was the machine produced?	81%	Received any assistance from the producer to operate this machine?	100%

The table displays a sample of questions with their respective response rates. *Source: Business Relations Survey.*

3.1. Target population and sampling frame and additional data

We constructed a sampling frame based on the AMADEUS database (**?**) to make sure that we can link survey responses to financial information. The target population included manufacturing firms in the industries with a NACE code between 20 and 30 (Revision 2) with at least 10 employees in the 3 countries.

We selected firms with non-missing basic financial information for 2012 and 2013. Based on the estimate of our surveyor partner, GfK, we anticipated a response rate of 15%, therefore our sampling frame included 6-7 times more firms than the actual sample. We used stratified sampling within countries based on two dimensions: size (10-50, 50-250 and over 250 employees) and ownership. In terms of ownership, we distinguished between foreign- and domestically-owned firms. All companies with at least one non-domestic owner were considered foreign. Both the size and the ownership variables were extracted from the Amadeus dataset. We created inverse probability weights to restore the representativity of our sample with respect to the target population.

Our survey partner received basic identifying information for each firm in our sampling frame (name, industry, address of the headquarters), randomly ordered within each stratum. We also assigned target sample sizes to reach within each size-ownership stratum. These targets implied oversampling of smaller hard-to-reach groups like large foreign firms. The country-level targets were proportional to the number of firms in each country: it was 500 firms in Slovakia, 600 in Hungary and 700 in Romania.

Given our sampling frame, we could link our survey data with Amadeus and extract financial information such as sales or assets. At this stage, all financial variables refer to 2013. Eventually we have financial information (such as total sales in 2013) for 82% of firms in the sample.

3.2. Sample

A first view of the data is presented by Table 5, which shows the number of observations along several dimensions. There are 1,535 firms in the final sample: 556 from Hungary, 584 from Romania and 395 from Slovakia. The majority of firms are small, with almost two thirds having less than 50 employees. This distribution is similar across countries, with a somewhat greater share of large firms in Slovakia.

The vast majority (72%) of the firms in our sample are domestically owned and the remaining 28% are foreign owned. These shares are also similar in the three countries, with a somewhat higher share of foreign-owned firms (37%) in Slovakia. In terms of industry, the largest share of firms operate in the fabricated metals industry, followed by rubber/plastic and machinery.

Table 5: Number of firms by number of employees, ownership and industry

Country	Hungary	Romania	Slovakia
Number of employees			
Less than 20 (38.7%)	203	213	179
Between 21 and 50 (25.7%)	134	167	93
Between 51 and 250 (28.1%)	184	167	80
More than 250 (7.5%)	35	37	43
Ownership			
Domestic (71.7%)	409	443	249
Foreign (28.3%)	147	141	146
Manufacturing sector			
Chemicals (4%)	19	25	17
Pharmaceuticals (0.8%)	3	6	4
Rubber and plastic (12.8%)	67	79	50
Non-metallic mineral (9.1%)	37	68	35
Basic metals (2.5%)	13	19	7
Fabricated metals (39.4%)	249	235	121
Computer, electronic and optical (4.8%)	24	23	27
Electrical equipment (6.9%)	36	28	42
Machinery (11.9%)	77	60	46
Motor vehicles (4.6%)	26	24	26
Other transport equip. (2.7%)	5	17	20
Total (100%)	556	584	395

Manufacturing sectors are NACE rev2, 2-digit categories. Ownership is based on majority owner. Number of employees is based on survey response.