Replication package for "Credit supply shocks and prices: Evidence from Danish firms"

Tobias Renkin* and Gabriel Züllig†

February 14, 2024

I. Data

A. Data availability

Data accessed via Statistics Denmark The main datasets used in the paper are based on administrative micro data collected by Danish government agencies and provided by Statistics Denmark (DST in short). All data is merged by DST using firm identifiers provided in the Danish business registry. These datasets can only be accessed by researchers approved by DST. The data are physically stored at DST servers and can be remotely accessed and analyzed by approved researchers, but may not be transferred outside this secure environment.

Researchers interested in obtaining access to administrative microdata at DST must submit a written application to gain approval. Applications can be submitted by researchers who are affiliated with Danish institutions accepted by DST, or by other researchers who collaborate with researchers affiliated with these institutions. The procedure is described at www.dst.dk/en/TilSalg/Forskningsservice.

Access to most datasets used in the paper is covered by the standard data-for-research program "Forskningsservice". We use three additional datasets that are not covered by the

*Danmarks Nationalbank, corresponding author: tobr@nationalbanken.dk; Langelinie Allé 47, 2100 Copenhagen Ø, Denmark

†Swiss National Bank, gabriel.zuellig@snb.ch; Börsenstrasse 15, 8022 Zurich, Switzerland

"Forskningsservice". First, we use the micro data underlying the Danish producer price index (PPI). This dataset is also collected by DST, but access is granted only upon request on a case-by-case basis. Second, we use data on bank balance sheets collected by Finanstilsynet. This dataset is provided by Finanstilsynet to researchers on request. Third, we complement the Finanstilsynet data with data collected in Danmarks Nationalbank's statistics on monetary and financial institutions (MFI). This dataset is not publicly available and can be merged to DST data on request. We only use the MFI data to fill in balance sheets for a handful of banks that are not supervised by Finanstilsynet, but the results of our paper can be replicated with minor differences using just the data available from Finanstilsynet.

Publicly available data We complement the register data with several other datasets (e.g. estimated demand elasticities or aggregate time series) that are publicly available and included in this replication package.

B. Details on each data source

Most datasets that are available through "Forskningsservice" are documented by DST. We refer to the documentation for these datasets on Denmark Statistics' website (https://www.dst.dk/extranet/forskningvariabellister/Oversigt%20over%20registre.html). The links are also included in the full citation in the references at the end of this document.

Data accessed via Statistics Denmark

- URTEVIRK (Statistics Denmark, 2022*f*): This dataset covers loan accounts from Danish financial institutions to Danish firms. The data is collected by the Danish tax agency SKAT and provided by DST Forskningsservice.
- IRTEVIRK (Statistics Denmark, 2022*d*): This dataset covers deposit accounts of Danish firms at Danish financial institutions. The data is collected by the Danish tax agency SKAT and provided by DST Forskningsservice.
- FIRM (Statistics Denmark, 2022*c*): This dataset is based on the Danish firm register and contains firm identifiers and basic firm characteristics. The dataset is created based on various administrative sources by DST Forskningsservice.

- FIKS (Statistics Denmark, 2022*a*): This dataset covers sales and purchases based on firms' VAT declarations. The dataset is created based on data collected by the Danish tax agency SKAT by DST Forskningsservice.

- FIRE (Statistics Denmark, 2022*b*): This dataset covers the Danish accounting statistics with detailed balance sheet information on Danish firms. The dataset is created based on various administrative sources by DST Forskningsservice.

- UHDM (Statistics Denmark, 2022*g*): This dataset is based on Danish customs data and provides values and quantities of exported goods. It covers both intra- and extra-EU trade. The dataset is provided by DST Forskningsservice.

- PPI (Statistics Denmark, 2022*e*): This dataset contains the micro data underlying the Danish producer price index. The data is collected by DST. It is not covered by Forskningsservice and access is granted on request on a case-by-case basis. A description of the survey methodology and a contact for data inquiries is available at https://www.dst.dk/da/Statistik/emner/erhvervsliv/industri/producent-og-importprisindeks-for-varer. We provide a variable list in Table 1 below.

- Annual bank balance sheets from Finanstilsynet and Danmarks Nationalbank MFI statistics (Danmarks Nationalbank, 2022, Finanstilsynet, 2022): These datasets cover the annual balance sheets of most banks active in Denmark. The balance sheets can be linked to account-level loan and deposit data using banks' unique firm identifiers. We use two balance sheet variables: total loans (on the asset side) of the balance sheet, and total deposits (on the liability side) of the balance sheet. Both variables exist in Finanstilsynet and Danmarks Nationalbank MFI data. We use Finanstilsynet as the source for most banks. For four banks that are not covered in Finanstilsynet data, we fill in loans and deposits data from Danmarks Nationalbank MFI statistics. The combined variables are called loans_ft and deposits_ft in the replication code. They are used to construct the loans-to-deposits ratio that we use to measure the exposure of banks to the 2007 financial crisis.

  Both the MFI and the full Finanstilsynet datasets are confidential, but can be used for research projects on a case by case basis. The Danmarks Nationalbank MFI data are described in detail at https://www.nationalbanken.dk/en/news-and-knowledge/data-and-statistics/banking-and-mortgage-lending, which also includes the relevant contact addresses. Details for Finanstilsynet data can be found at https://www.dfsa.dk/Fact-

and-Figures/Key-performance-indicators. It can take some months to negotiate data use agreements and gain access to the data. The authors will assist with obtaining data access from Finanstilsynet and Danmarks Nationalbank for any reasonable replication attempts.

Alternatively, Finanstylnet provides bank-level loan to deposits ratios as part of their "key performance indicators". These data are publicly available at https://www.dfsa.dk/Fact-and-Figures/Key-performance-indicators. While we calculate this ratio from the separate loan and deposit variables, our code can be easily adapted to directly use the ratio that is publicly available. One can also use data from URTE and IRTE registers to compute a very similar measure directly from data available through Forskningsservice without external sources. As we show in Figure B.1 in the paper, the bank-level balance sheet totals can be replicated very well from the more detailed microdata.

Supporting data: Product-level information

- PRODCOM (Eurostat, 2021): This dataset covers the production of industrial output in all EU member countries, including Denmark. The data can be redistributed and is provided in this replication package in /data/5_demand_elasticities/in/PRODCOM/.
- ElasticitiesBrodaWeinstein90-01_SITCRev3_4-digit.xls: Demand elasticities estimated by Broda and Weinstein (2006*a*) for the period 1990-2001 at the 4-digit level of the SITC Rev. 3. The estimates are retrieved from Broda and Weinstein (2006*b*), saved in /data/ 5_demand_elasticities/in/, and provided in the replication package with permission.
- ElasticitiesBrodaWeinstein90-01_HTS-3.csv: Demand elasticities estimated by Broda and Weinstein (2006*a*) for the period 1990-2001 for HS categories. The estimates are retrieved from Broda and Weinstein (2006*b*), saved in /data/5_demand_elasticities/in/, and provided in the replication package with permission.
- Product code conversion tables
  - Combined Nomenclature definitions and correspondence tables, 2001–2016: Available from Eurostat (https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL). The data can be redistributed and is included in the replication package in /data/4_unitvalues/in.

- Combined Nomenclature-SITC correspondence table: Available from United Nations (UN Statistics Division (2022), go to "HS - SITC/BEC", choose "HS 2007 - SITC Rev. 3") and saved as data/5_demand_elasticities/in/CN6_SITC3.txt.
- Prodcode-Combined Nomenclature correspondence table: Available from Eurostat (https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL&StrLanguageCode=EN&IntCurrentPage=3, go to "CN 2010 - PRODCOM 2010") saved as data/5_demand_elasticities/in/PRC 2010 - CN 2010.xls. Additionally, we use from the same source three files for correspondence tables across years ("CN 2007 - CN 2008", "CN 2008 - CN 2009", "CN 2009 - CN 2010"), all provided in this replication package.

Supporting aggregate data

- /data/5_demand_elasticities/in/exchange_rates_9506.txt: Monthly averages of 50 bilateral exchange rates to DKK provided within the replication package. Accessed via Statistikbanken (Statistics Denmark's data portal). To obtain the data directly from the Statistikbanken URL (www.statistikbanken.dk), follow the following steps: 1. Click on the above link, 2. search for "DNVALM" (monthly exchange rates), 3. choose Currency=all, Type=Exchange rates (DKK per 100 units of foreign currency), Methodology=Monthly average, Month= 1995M01-2006M12, then click "show table", 4. Select "pivot table", 5. unselect "incl. footnotes etc.", 6. Choose "Download file as..." by selecting "Text matrix".

- data/5_demand_elasticities/in/exchange_rates_0717.txt: Due to the data size, the monthly exchange rates have to be downloaded in two batches, thereby splitting the sample into 1995-2006 and 2007-17.

- data/5_demand_elasticities/in/country_currency_keys.txt: list of country names (in Danish and English) and currency abbreviation with start and end date if applicable (e.g. Tyskland, Germany, DEM, NA, 1998m12; Tyskland, Germany, EUR, 1999m1, NA); compiled ourselves and provided in the replication package.

- data/5_demand_elasticities/in/DestinationControls.csv: real GDP, unemployment rate, real exports and imports for the same countries compiled from OECD Quarterly National Accounts database (OECD (2021)). Use is permitted (see https://www.oecd.

org/termsandconditions/). The data in the right format is provided in this replication package.

- analysis/7_agg_counterfact/in/PRIS4315.txt: Official manufacturing price index (dashed line in Figure 6 in the paper) and provided within the replication package. Accessed via Statistikbanken (Statistics Denmark's data portal). To obtain the data directly from the Statistikbanken URL (www.statistikbanken.dk), follow the following steps: 1. Click on the above link 2. search for "PRIS4315", 3. choose Industry=Manufacturing, Market=Total Danish production, Unit=Index, Month=2005M01-2020M12.

- analysis/7_agg_counterfact/in/weights.txt: The exact item-level weights used by DST to aggregate the PPI micro data to the headline number are not provided in our dataset. DST infrequently publishes the industry weights. We use the weights published in January 2009, which are provided here: https://www.dst.dk/Site/Dst/SingleFiles/GetArchiveFile.aspx?fi=pris&fo=prodv9--pdf&ext=%7B2%7D.

- analysis/7_agg_counterfact/in/data_q_in.xlsx: 6 aggregate time series on Danish and US GDP, PPI and loans to corporations. US series are compiled from FRED (mnemonics GDPC1, PCUOMFGOMFG and BUSLOANS, see U.S. Bureau of Economic Analysis (2021), U.S. Bureau of Labor Statistics (2021), Board of Governors of the Federal Reserve System (2021)). None of the series have explicit copyright notices and can be re-distributed (see https://fred.stlouisfed.org/legal/#:~:text=FRED%20provides%20data%20and%20data,and%20conditions%20of%20the%20service.) Series on Danish GDP and loans from Statistikbanken (step-by-step instruction for search window provided in "sources" worksheet of file). The Danish PPI is an PPI index computed by aggregating the micro data. All series and sources are provided within this replication package.

C. Details on important intermediate datasets

We merge the raw datasets to several intermediate datasets which we use in alaysis. We describe the most important (needed for the core of the analysis) below:

- sample.dta: Cross-section of manufacturing firms that are a.) active, b.) have at least 10 employees and c.) a bank connection (2,703 obs., see fourth row Table A.1 in paper). All variables describe time-invariante information (e.g. industry) or revenue

statement/balance sheet information for the year 2007. A list of selected variables is provided in Table 2 to faciliate interpreting the codes.

- panel.dta: Panel of the same 2,703 firms for the years 2003-14 containing a.) the same time-invariant variables as sample.dta, and b.) time-varying firm characteristics such as sales, employment, etc. See Table 3 for illustration.

- uv_all.dta: sum of values and quantities for each firm-product category (CN2)-year with corresponding unit value indices based on UHDM. While we use the raw PPI data directly (it is cleaned by DST), the UHDM data requires substantial processing for several reasons. First, firm-products are not unique in the customs data, mostly due to firms exporting to multiple countries. For most analyses, we need the firm-product-destination data to be collapsed to the firm-product level. Second, the original UHDM is at the 8-digit CN level, which is subject to frequent re-classifications. We track re-classifications as well as possible to get time-consistent product classifications. Even so, substantial noise remains, which is why we aggregate exports to the CN2-level.

II. Description of programs/code

In the following, we describe the organization of the codes in our replication package. The entirety of our project codes can be run by "00_main.do", which initializes all settings and environment variables and subsequently calls all codes for data generation and analysis. We divide all codes into two groups of tasks, namely data generation/merging ("data") and "analysis". Each task is contained in a separate subfolder to "data/" or "analysis/" and is self-contained, i.e. can be run separately (after initializing all settings) if the remaining input data has been generated. The subfolder contains one or several scripts, typically Stata .do-files, the header of which describes its output as labeled in the paper, if any. Each task subfolder further contains in most cases 3 own subfolders, namely "in/", "temp/", and "out/". "in/" contains raw supporting data that are unique to this task, "temp/" contains files produced and only used within the respective task, and "out/" contains output files used by other tasks or figures/tables used in the paper with the equivalent labeling.

The replication package has the following high-level structure.

- "/00_main.do" runs the entire project in correct order.

- "/init.do" sets environment variables such as paths and global settings for figures and tables.
- "/data" contains scripts to conduct tasks such as data preparation, cleaning and merging for firms' bank loans (1), banks' balance sheets (2), firm registers (3), export unit values (4) and product-level demand elasticities (5).
- "/analysis" contains scripts for all tasks needed to replicate the results in the paper, including all figures and tables: sample descriptives (1), first-stage regressions/loan market outcomes (2), reduced-form results using the PPI data (3), reduced-form results using the export unit values (4), instrumental variable regressions (5), other firm outcomes (6) and the counterfactual time series exercise (7).
- "/ado" contains various stata .ado files that are used in several places in the analysis.

We provide details on each task as follows, including – if applicable – the raw input data and major dependent data files used as well as the output (figures and tables) used in the paper.

Besides the data files that are included in the package, the programs assume that confidential DST datasets are located in the paths outlined in Table 4.

A. /data/

- **1_combine_urtevirk_irtevirk/**
  - **1_gen_combine.do**: load and clean account-level loan and deposit data, augment with bank-level information and collapse to firm-bank and firm-level. Raw input data: URTEVIRK, IRTEVIRK, Bank balance sheets.
- **2_bank_data/**
  - **1_gen_bank_panel.do**: generate panel of bank-level characteristics based on balance sheet as well as totals from account-level micro data. Raw input data: FIRM, Bank balance sheets.
- **3_sample/**

- **1_gen_exposure.do**: generates the firm-level exposure variable used throughout the paper ("tt_uvall", the share of loans from wholesale-funded bank for each firm). Raw input data: Bank balance sheets. Major intermediate data input: firm_panel.dta; firm_bank_panel.dta (both generated in data/1_combine_urtevirk_irtevirk/1_gen_combine.do).
- **2_gen_sample_2007.do**: collects the relevant sample of firms and their characteristics in 2007. Raw input data: FIRM; FIKS; FIRE; IRTEVIRK; URTEVIRK. Major intermediate input data: exposure.dta (generated in previous script). Output: sample.dta. Output files referenced in the paper: Table A.1.
- **3_marketsizeEU.R**: computes EU market sizes for each industry. Raw input data: PRODCOM.
- **4_gen_panel.do**: collects firm-level variables for the firm sample during years 2003–2012. Raw input data: FIRM; FIKS; FIRE. Major intermediate input data: firm_panel.dta (generated in data/1_combine_urtevirk_irtevirk/1_gen_combine.do). Output: panel.dta.

- **4_unitvalues/**
  - **0_gen_CN_panel.do**: Generate a panel of valid CN codes in every year based on CN definition files. Raw input data: CN definitions
  - **0_gen_CN_transitions.do**: Generate mappings between CN codes in different years based on CN correspondence tables. Raw input data: CN correspondence tables
  - **1_gen_unit_value_indices.do**: Creates unit value indices at the firm-CN2 and firm-CN2-country level. Raw input data: UHDM. Output: uv_all.dta.

- **5_demand_elasticities/**
  - **1_DE_distrib_PRODCOM.R**: compiles distribution of demand elasticities in the EU and Denmark. We run this script on an open computer, i.e. not a Statistics Denmark server, and then transfer the resulting file temp/cdf_prodcom.dta to the Statistics Denmark computer. Raw input data: PRODCOM; in/ElasticitiesBrodaWeinstein90-01_SITCRev3_4-digit.xls; product code conversion tables. Output files referenced in the paper: Table C.1.

- **2_DE_prepare_broda_weinstein.do**: prepares demand elasticities at various levels of disaggregation and product definitions to be merged onto our price datasets. Raw input data: in/ElasticitiesBrodaWeinstein90-01_SITCRev3_4-digit.xls (saved as .txt); product codes conversion tables (only in/CN6_SITC3.txt); in/ElasticitiesBrodaWeinstein90-01_HTS-3.csv; PPI; Major dependent input data: exportdata.dta (byproduct of data/4_unitvalues/1_gen_unit_value_indices.do).
- **3_DE_aggregated.do**: aggregates product-level demand elasticities to firm-level (weighted) averages for the firms in our data. Raw input data: PPI; Major dependent input data: cn6_byyear.dta (generated in previous script); exportdata.dta (byproduct of data/4_unitvalues/1_gen_unit_value_indices.do); sample.dta.
- **4_DE_distributions.do**: plots cumulative distributions of demand elasticities across datasets and samples. Major dependent input data: panel.dta; firmlvl_DE_allcvrnr.dta (generated in previous script); Output files referenced in the paper: Figure 4c.
- **5_SC_compile_exchange_rates.do**: compiles exchange rates for pass-through regressions. Raw input data: in/exchange_rates_9506.txt; in/exchange_rates_0717.txt; in/country_currency_keys.txt.
- **5_SC_compile_trade_data.do**: compiles independent control variables (aggregate data) and merges on export unit values for pass-through regressions. Raw input data: in/DestinationControls.csv; Major dependent input data: exportdata_countries.dta (generated in data/4_unitvalues/1_gen_unit_value_indices.do).
- **6_SC_regressions.do**: runs exchange rate pass-through regression on data generated in previous script. Output files referenced in the paper: Table C.2.
- **7_SC_combine_results_by_CN2.do**: collects pass-through estimate by CN2 category from regression output in previous script, to be merged on price datasets later.

B. /analysis/

- **1_sample_descriptives/**
  - **1_sample_descriptives.do**: generates tables on descriptives statistics by sample and exposure. Output files referenced in the paper: Table 1, Figure A.1a-b, Figure A.3, Figure A.4, Table A.2, Figure A.5.

- **2_piecharts_industry.do**: generates pie charts on the industry composition of the samples. Output files referenced in the paper: Figure C.1.
- **3_loans_micro_vs_balance_sheets.do**: compares bank lending when measured from banks' balance sheets directly to when aggregate from the account-level data. Output files referenced in the paper: Figure A.2a-b.
- **4_uhdm_vs_ppi_comparison.do**: compares changes in unit values to changes in prices reported in the PPI for firms that are included in both datasets. Output files referenced in the paper: Figure B.1.

- **2_firststage/**
  - **1_firststage_banklevel.do**: shows bank-level outcomes of wholesale-funded vs. deposit-funded banks. Major dependent input data: banks_panel.dta (generated in data/2_bank_data/1_gen_bank_panel.do). Output files referenced in the paper: Figure 1a-b.
  - **2_firststage_firmlevel.do**: shows firm-level credit outcomes by exposure to wholesale-funded banks ("tt_uvall"). Major dependent input data: panel.dta. Output files referenced in the paper: Figure 2a-f, Table A.3, Table A.4, Table A.5.

- **3_did_ppi/**
  - **1_gen_panel.do**: generates the good-level panel of prices, augmented with firm-level (time-invariant) characteristics and (time-invariant) good characteristics used in the following two scripts. Raw input data: PPI; Major dependent input data: sample.dta; cn6.dta (generated in data/5_demand_elasticities/3_DE_aggregated.do).
  - **2_did_ppi.do**: performs reduced-form regression for baseline result (Equation 4 in the paper) and robustness checks. Output files referenced in the paper: Figure 3a, Table A.6, Table A.7
  - **3_did_ppi_heterogeneity.do**: performs reduced-form regression with various interaction terms (Equation 6 in the paper). Output files referenced in the paper: Figure 4a, Table 3 (columns 1-3), Table 4 (columns 1-3).

- **4_did_uhdm/**
  - **1_gen_panel.do**: generates the good category-level panel of unit values, augmented with firm-level (time-invariant) characteristics and (time-invariant) good

characteristics at the appropriate level of disaggregation (see data/5_demand_elasticities/3_DE_aggregated.do) used in the following two scripts. Major dependent input data: uv_all.dta; sample.dta; firmCN2lvl_DE_UHDMfirms.dta (generated in data/5_demand_elasticities/3_DE_aggregated.do).

  - **2_did_uhdm.do**: performs reduced-form FGLS regression for baseline result and robustness checks. Output files referenced in the paper: Figure 3b, Table A.8, Table A.9, Figure A.6a-b.
  - **3_did_uhdm_heterogeneity.do**: performs reduced-form FGLS regression with various interaction terms. Output files referenced in the paper: Figure 4b, Table 3 (columns 4-6), Table 4 (columns 4-6).

- **5_iv/**
  - **1_gen_ppi_panel.do**: generates the good-level panel of prices, augmented with firm-level (time-invariante and time-varying) characteristics used in the following script. Raw input data: PPI; Major dependent input data: sample.dta; panel.dta.
  - **2_iv_baseline.do**: runs IV regressions on both price datasets. Major dependent input data: data merges created in previous script (for PPI) and analysis/4_did_uhdm/1_gen_panel.do (for UHDM); panel.dta. Output files referenced in the paper: Table 2.

- **6_did_firm_outcomes/**
  - **1_did_firmoutcomes.do**: runs reduced-form regressions of several firm-level variables on "tt_uvall" interacted with annual dummies. Major dependent input data: panel.dta; firmlevel_DE_allcvrnr.dta (generated in data/5_demand_elasticities). Output files referenced in the paper: Figure 5a-b, Table A.10, Figure A.7a-b.
  - **2_average_working_capital.do**: shows moments of working capital for the relevant sample of firms and imputes the corresponding increase in marginal cost following Equation 7 in the paper. Output files referenced in the paper: Table 5.

- **7_agg_counterfact/**
  - **1_prep_dstseries.do**: prepares DST-provided information to replicate the official PPI using the micro data (e.g. industry weight approximations). Raw input data: PRIS4315.txt; weights.txt.

- **2_estimate_cf.do**: computes counterfactual prices in the PPI based on reduced-form estimates (see Equation 8 in the paper). Raw input data: PPI; Major dependent input data: sample.dta, panel.dta.
- **3_construct_index.do**: aggregates counterfactual prices to firm-, industry-, and aggregate averages using estimates and weights prepared in previous two scripts. Output files referenced in the paper: Figure 6.
- **m4_runRemainderAggregate.m**: runs a 2-variable VAR to show that prices during the financial crisis increased/fell less compared to a conditional forecast with ex-post observed GDP. Since Matlab is not installed on DST servers, this script was run on an other computer after transferring the relevant data series. The script will call m5_prepare_data_q.m, m6_var.m, m7_counterfactual.m and several functions to estimate vector autoregressions all provided in the subfolder var_tbx/. Raw input data: in/data_q_in.xlsx. Output files referenced in the paper: Figure D.1a-b; Table D.1.

C. /ado/

This directory contains several custom ado files:

- **rifhdreg.ado, _grifvar.ado**: Implements RIF-based median regressions. The methodology and code are documented in Rios-Avila (2020). The package is available is in SSC, but was manually installed since it is not available in the DST environment.
- **group_toway.ado**: group_twoway (Mation and Maciente, 2014) groups observations by the connected components of the parent and child variables. The package is available is in SSC, but was manually installed since it is not available in the DST environment.
- **ivreghdfe_wfix.ado**: ivreghdfe (Correia, 2018) is a popular package for IV regressions with high-dimensional fixed effects. Prior to version 1.2, there is a bug where weights are ignored in IV regressions (see https://github.com/sergiocorreia/ivreghdfe/issues/38). Since the DST environment has this older version installed, we added an updated version where this bug is fixed and use this version for weighted IV regressions.
- **rlasso_fe.ado**: Implements PSDLASSO with fixed effects: all fixed effects are absorbed from all candidate variables before the PDSLASSO methodology is applied.

- **eststo_wrap.ado, eststo_ivwrap.ado**: Short convenience scripts that add several statistics to models before the model is saved.

III. Computational requirements

- STATA: The main analysis was conducted with Stata (version 17) on the servers of Statistics Denmark. We use the following STATA packages that are installed from SSC: reghdfe, ivreghdfe, gtools, distinct, winsor2, coefplot, binscatter, grstyle, catplot, addplot, pdslasso, estout
- R: Selected scripts are provided in R and Matlab. We used R version 4.2.0 and R Studio 2022.12.0. Additionally, the following libraries are installed from CRAN: gdata, readxl, data.table, foreign, stringr, dplyr, spatstat, haven.
- Matlab: We used a Student version of Matlab R2022a and no further packages. Used only for the time series analysis in Appendix D (analysis/7_agg_counterfact/).

Approximate time needed to reproduce the analyses: about 20 hours on DST servers.

IV. Instructions to Replicators

1. Apply for access to Danmarks Statistics Forskningsservice, including all adminstrative datasets described in this document.
2. Transfer the replication package to the DST server. Make sure globals in init.do are set to the right data locations.
3. All programs described above are executed in correct order by running "00_main.do".

References

Board of Governors of the Federal Reserve System. 2021. "Commercial and Industrial Loans, All Commercial Banks [BUSLOANS], retrieved from FRED, Federal Reserve Bank of St. Louis." *https://fred.stlouisfed.org/series/BUSLOANS*, Accessed: January 2021.

Broda, Christian, and David E. Weinstein. 2006*a*. "Globalization and the Gains From Variety." *Quarterly Journal of Economics*, 121(2): 541–585.

Broda, Christian, and David E. Weinstein. 2006*b*. "Globalization and the Gains From Variety: Dataset." *http://www.columbia.edu/∼dew35/TradeElasticities/TradeElasticities.html*, Accessed: 10 April 2021.

Correia, Sergio. 2018. "IVREGHDFE: Stata Module for Extended Instrumental Variable Regressions with Multiple Levels of Fixed Effects." *Statistical Software Components*.

Danmarks Nationalbank. 2022. "Balance sheet information on Monetary and Financial Institutions [database]." Accessed: 2022.

Eurostat. 2021. "Prodcom - Production of manufactured goods [database]." *https://ec.europa.eu/eurostat/web/prodcom*, Accessed: 3 February 2021.

Finanstilsynet. 2022. "Balance sheet information about Danish banks [database]." Accessed: 2022.

Mation, Lucas Ferreira, and Aguinaldo Nogueira Maciente. 2014. "GROUP_TWOWAY: Stata Module to Group Observations by the Connected Components of Two Variables." *Statistical Software Components*.

OECD. 2021. "Quarterly National Accounts [database]." *https://stats.oecd.org/Index.aspx?DataSetCode=QNA#*, Accessed: 2021.

Rios-Avila, Fernando. 2020. "Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition." *The Stata Journal*, 20(1): 51–94.

Statistics Denmark. 2022*a*. "FIKS - Firmaernes køb og salg [database]." *https://www.dst.dk/extranet/ForskningVariabellister/FIKS%20-%20Firmaernes%20k%C3%B8b%20og%20salg%20(m%C3%A5neder).html*, Accessed: 3 January 2022.

Statistics Denmark. 2022*b*. "FIRE - Regnskabsstatistikken [database]." *https://www.dst.dk/extranet/ForskningVariabellister/FIRE%20-%20Regnskabsstatistikken.html*, Accessed: 3 January 2022.

Statistics Denmark. 2022*c*. "FIRM - Generel firmastatistik [database]." *https://www.dst.dk/extranet/ForskningVariabellister/FIRM%20-%20Generel%20firmastatistik.html*, Accessed: 3 January 2022.

Statistics Denmark. 2022*d*. "IRTEVIRK - Indlånsrenter - Kontospecifikke data for virksomheder [database]." *https://www.dst.dk/extranet/ForskningVariabellister/IRTEVIRK%20-%20Indl%C3%A5nsrenter%20-%20Kontospecifikke%20data%20for%20virksomheder.html*, Accessed: 3 January 2022.

Statistics Denmark. 2022*e*. "Producer Price Index Survey Microdata [database]." *https://www.dst.dk/da/Statistik/emner/erhvervsliv/industri/producent-og-importprisindeks-for-varer*, Accessed: 3 January 2022.

Statistics Denmark. 2022*f*. "Udlånsrenter - Kontospecifikke data for virksomheder [database]." *https://www.dst.dk/extranet/ForskningVariabellister/URTEVIRK%20-%20Udl%C3%A5nsrenter%20-%20Kontospecifikke%20data%20for%20virksomheder.html*, Accessed: 3 January 2022.

Statistics Denmark. 2022*g*. "UHDM - Udenrigshandel diskretioneret (måneder) [database]." *https://www.dst.dk/extranet/ForskningVariabellister/UHDM%20-%20Udenrigshandel%20diskretioneret%20(m%C3%A5neder).html*, Accessed: 3 January 2022.

UN Statistics Division. 2022. "UN Comtrade Conversion table HS2007 to SITCRev3 [database]." *https://unstats.un.org/unsd/classifications/Econ*, Accessed: 2022.

U.S. Bureau of Economic Analysis. 2021. "Real Gross Domestic Product [GDPC1], retrieved from FRED, Federal Reserve Bank of St. Louis." *https://fred.stlouisfed.org/series/GDPC1*, Accessed: January 2021.

U.S. Bureau of Labor Statistics. 2021. "Producer Price Index by Industry: Total Manufacturing Industries [PCUOMFGOMFG], retrieved from FRED, Federal Reserve Bank of St. Louis." *https://fred.stlouisfed.org/series/PCUOMFGOMFG*, Accessed: January 2021.

Table 1: PPI: complete variable list

Variable	Raw*	Description
cvrnr	yes	Anonymized firm identifier
hs	yes	8-digit code in the Harmonized System Nomenclature
lbnr	yes	"Løbenummer", index in case number of reported products per firm and hs-code >1
beskrivelse	yes	Description of product
aar	yes	Year of survey
maaned	yes	Month of survey (conducted mid-month)
quarter	no	We aggregate the original dataset, which is at the monthly frequency, to the quarterly frequency by keeping the first price in each quarter.
tq	no	recoded variables "aar" and "quarter" in Stata datetime format
nationtype	yes	2=export, 3=domestic
pristype	yes	flag for transfer prices (ca. 3.5%)
pris	yes	Price reported in original currency
valuta	yes	Currency
prisdkk	yes	Price in Danish kroner
korrpris	yes	If the properties of the product (e.g. quality) are adjusted, the firm reports the hypothetical price of the adjusted product in the previous period.
Dp2	no	=log(pris) - log(L1.korrpris), winsorized at +/- 1
prisadj2	no	Index of quality-adjusted price built using Dp2
hs6	no	=substring(hs, 1, 6)
panid	no	=group(cvrnr hs lbnr nationtype)

* "yes" indicates that variables were taken from the raw data provided by Statistics Denmark; "no" indicates that the variable was derived from the raw variables, in which case the exact transformations are provided.

Table 2: sample.dta: selected variable list

Variable	Raw*	Description
cvrnr	all	Anonymized firm identifier
nace	FIRM	4-digit NACE code for industry (original name: GF_NACE_DB)
emp07	FIRM	Number of employees in November 2007 (original name: GF_ANSATTE)
revenue07	FIRM	Sales (original name: GF_OMS)
loans_uv07	URTEVIRK	Amount of total loans outstanding on 31 December 2007 according to URTEVIRK
loans_to_revenue07	dep.	= loans_uv07 / revenue07
deposits07	IRTEVIRK	Amount of total deposits on 31 December 2007 according to IRTEVIRK
deposits_to_revenue07	dep.	= deposits07 / revenue07
stshare07	dep.	Share of loans explicitly marked as short-term (maturity $<$ 1 year in URTEVIRK)
interest_rate07	dep.	Interest paid in 2007 according to URTEVIRK / 0.5*(Sum of total loans outstanding on 31 December 2006 and 2007 according to URTEVIRK)
tt_uvall	dep.	Share of loans on 31 December 2007 from banks with a 2007 ratio of total loans to total deposits ratio above the median.
uvbt07	FIRE	total inventories, end-of-period (original name: UVBT)
tgt07	FIRE	accounts receivable under current assets (orignal name: TGT)
kgl07	FIRE	accounts payable under short-term liabilities (original name: KGL)
lgl07	FIRE	accounts payable under long-term liabilities (original name: LGL)
wk_to_revenue07	dep.	= (uvbt07 + tgt07 - kgl07 - lgl07) / revenue07
inPPI_20*	dep.	Dummy for whether firm has observations in PPI data in respective year
inUHDM_20*	dep.	Dummy for whether firm has observations in UHDM data in respective year

* describes the original data register, with "dep." indicating that it is a variable dependent on these data sources.

Table 3: panel.dta: selected variable list

Variable	Raw*	Description
cvrnr	all	Anonymized firm identifier
All variables from sample.dta (see Table 2) are included in this dataset.		
revenue_fiks	FIKS	Sales relevant for VAT (original name: SALG_IALT)
lgag_firm	FIRM	Labor expenses (original name: GF_LGAG)
purchases_fiks	FIKS	Purchases of material and other input cost relevant for VAT (original name: KOB_IALT)
interest_uv	URTEVIRK	Sum of interest payments to lenders over the course of a year
msfiksEU	dep.	= revenue_fiks / Sum of industrial production values in same 4-digit industry in EU28
fte_firm	FIRM	Full-time equivalent employment over the course of a year

* describes the original data register, with "dep." indicating that it is a variable dependent on these data sources.

Table 4: Expected data location for data not included in replication package

Dataset	Files	Notes
URTEVIRK	$cleandatapath/URTEVIRK/out/URTEVIRK_accounts.dta	URTEVIRK 2003–2014
IRTEVIRK	$cleandatapath/IRTEVIRK/out/IRTEVIRK_accounts.dta	IRTEVIRK 2003–2014
Bank balance sheets	$projectpath/data/mfi/out/mfi_panel.dta	Panel 2003–2014
Bank balance sheets	$projectpath/data/mfi/out/mfi_data_2007.dta	Data for 2007 only
FIRM	$cleandatapath/FIRM/out/firm_nyvar_YYYY.dta	Yearly files from 2003–2014
FIRE	$cleandatapath/FIRE/out/fire.dta	FIRE from 2003–2014
FIKS	$cleandatapath/FIKS/out/FIKS_annual.dta	FIKS annual sums 2003–2014
UHDM	$cleandatapath/UHDM/out/UHDM_exports.dta	UHDM dataset (only exports), 2003–2014
PPI	$cleandatapath/PPI/out/PPI_domestic.dta	PPI dataset (only domestic prices), 2003–2014

$cleandatapath and $projectpath are global variables set in init.do. Make sure you set them to the correct paths for the replication files to run.