Assignment

Database and Data Management

Developed by:
201372071

May 25, 2022

1 Introduction

For centuries the shipping industry has played a major role in moving goods around the globe. This is true today and will continue to be as international integration increases (Mohanty et al., 2021). Even with the recent disruptions and resulting integration push back, managers cannot entirely replace the global sourcing, instead they are choosing to reduce their supply chain exposure relative to key components to their companies (McLain, 2021). The dominance of the shipping industry on international trade will persist, as it is the best option to move large quantities at a reduced costs (Park, Seo, and Ha, 2019). This allows firms to reduce costs, but also makes their production output dependent of the global supply chains. These infrastructures are instable and vulnerable to shocks, as recent events, such as, the Covid-19 pandemic or Suez-canal accident have shown us.

Theory dictates that the output of a firm's production function depends on productivity, labour hired and the production inputs. It is on the latter that the supply chains are relevant, as they act as linking points between companies. Moving the required inputs from where they are produced to where they are needed. To study the impact of supply networks on goods production, Elliott, Golub, and Leduc (2020) develop a model, where the inputs are conditional on other firms. They assumed that the supplier's link strength is endogenous because it relies on each company's decisions to investment in improving their present relationship or to build new ones. Their main findings were the following. Improving input's reliability comes with additional costs. In addition, there are reliability spill-overs – this means that a firm' payoff for improving their connections is smaller than the overall benefit for the supply chain, this is known as the reliability externality. The model demonstrates that in equilibrium there is always underinvestment in the supply chains connections. We aim to test if the stock market incentivises a more reliable equilibrium point, by acting as a disciplining mechanism on this market failure.

Finally, previous empirical research on supply chain links and their implications to the financial market, have showed that for supplier-customer pairs, their accounting fundamentals and stock price movement are positively correlated (Cohen and Frazzini, 2008; Menzly and Ozbas, 2010). Also, more recently, Ersahin, Giannetti, and Huang (2022) measure the supply chain risk reported on each firm earnings conference calls to study how companies manage and mitigate this risk. Finally, Barrot and Sauvagnat (2016) analysed how the inputs specificity affects the shocks propagation throughout the production networks and its effects on the company market value. All of these, focus on exploring how the supply chain characteristics impacts the stock market. While our research question flows in the contrary direction. We are testing the effects of the financial markets on the supply chain connections strength, assessing if they can discipline the reliability externality. Taking this in consideration, our work also relates to the Hsu, Li, and Tsou (2020) investigation of how the stock market can help correcting the pollution externality. We share the same economic concept, but on a different topic.

2 Data

In this section, we describe our main data sources and the variables we use and construct on our analysis. We provide in detail the databases, variables definitions and transformations required to replicate this study.

2.1 Shipping cost

Our goal is to estimate the firm's fundamentals sensitivity to the changes on the supply chain pressures. This will allow us to build long-short stock portfolios and test if there is a risk premium for the supply chain reliability. If so, this shows that the stock market automatically can discipline the reliability externality.

As previous studies Hummels (2008), Hummels (2007), and Bradford (2005), we proxy this variable with the *ad valorem* value between the Cost-Insurance-Freight[1] and the Free-on-Board[2] at the product level for US importations. These metrics are compiled by the US Census Bureau from the customs records for transactions with value superior to $2,500 per commodity. Also, these reports denote the mode of transport. Allowing us to select only importations by sea. We retrieve this data from Feenstra (1996) from 1974 to 1988, and the remaining of the sample we get it through Peter Schott's website[3] (Schott, 2007) from 1989 to 2014.

This annually product-level data is aggregated by year and Standard Industrial Classification (SIC code). This will enable us to later link the transportation cost to the company's stock market information at the industry-level. After this step, we follow Barrot, Loualiche, and Sauvagnat (2019) approach that focuses on manufacturing industries (SIC between 2000 and 3999), without this assumption our results would be bias, because physical goods require legal fillings of import transactions when going through US customs, while others the same does not apply. Resulting in different compiling processes for the trade statistics. Additionally, it was been showed that those sectors are more expose to the global supply chains (Ersahin, Giannetti, and Huang, 2022). One final transformation, is to calculate the first difference for the shipping cost, using the resulting values as our main variable of interest. This first difference analysis gives more weight to the small variations, helping us to focus on the supply chain pressure build ups, instead of import competition or inputs complexity. Finally, we drop the observations when the shipping cost first difference is missing.

Table I shows the summary statistics for our main variable of interest, the change of shipping cost. Before and after winsorizing at the 1% level. We can see the impact of outliers before this transformation, especially the bias towards the right-hand side of the distribution. After treating the extreme values, the

[1]International shipping nomenclature when it the responsibility of the seller to cover the shipping cost, insurance of the customer's order. The legal responsibility of the supplier ends on the destination port.

[2]International shipping term when it the responsibility of the buyer to pay the transportation costs and assume the losses of any damaged good during the transportation. The legal responsibility of the supplier ends on the departure port.

[3]https://sompks4.github.io/sub_data.html

mean is 0.004 with a significant standard deviation of 0.197, that is reflected on wide range between the first and last percentile (1.481, significant larger considering that is a first difference variable).

Consistent with previous studies Barrot, Loualiche, and Sauvagnat (2019), our analysis confirms the persistence of the shipping cost. Nevertheless, after applying the first difference, this is no longer true. The table II confirms it, as the transition across quintiles is around one fifth. Additionally, the Wooldridge test for autocorrelation (Appendix A) shows that our transformation solves the autocorrelation problem.

Overall, our variable of interest varies within sector and across time. This will aid us to measure each sector's sensitivity to the global supply chain and its risk premium.

2.2 Stock Market

The firm's stock market information is downloaded from the Center for Research in Security Prices (CRSP) monthly file, and we use Compustat for the accounting information. We focus on common-ordinary shares (item *shrcd* equal to 10 or 11) traded on the AMEX, NASDAQ, or NYSE (item *exchcd* equal to 1, 2, or 3) for our sample period.

The four-digits SIC code is the connecting point between the shipping cost and the firm-level information. We compile it, from the two databases mention above. The anchor for this variable is the Compustat SIC code in the previous years (item *sich*), when this one is missing, we use the backfilled Compustat SIC (item *sic*). After these steps, the existing gaps are filled with the CRSP industry code (item *siccd*). To expurgate the impact of micro-cap on our analysis, we exclude firms with the market capitalization below the 10th percentile relative to the NYSE/AMEX universe. For our sample period, 1975 to 2014, we end up with 39,268 yearly observations for 4,214 different companies.

We retrieve firm characteristics from Compustat Fundaments Annual files that is merge with the previous dataset by firm identifiers (item *permno*). We obtain this though linking the CRSP (using item *ncusip*) to the Compustat (using the first 8 digits of the item *cusip*). After, we get the sales (item *sale*) that we divide by the total assets (item *at*) to control for the company size. For the inventory we download the item *invt* (total inventories) also pondered by the firm size proxy. The gross profit margin is calculated by the markup between the gross profit (item *gp*) and the total revenue (item *revt*). Finally, the net profit margin is the ratio between net income (item *ni*) and the total revenue (item *revt*). Following the previous section approach, we apply the first difference for the firm-level characteristics. And all these variables are winsorized at the first and last percentile.

The table III shows the summary statistics for the firm-level variables after the merge with the shipping cost dataset. During our sample period the mean (median) pondered sales growth is 8.3% (0.5%), this is significantly different from the Ersahin, Giannetti, and Huang (2022) empirical study, but could

Table I

Summary statistics

The table shows the summary statistics for the industry-year shipping cost variation for the manufacturing industries traded on the AMEX, NASDAQ, or NYSE from 1974 to 2014. These variables are the change of the ratio between CIF (cost-insurance-freight) to the FOB (free-on-board) at the industry level. The shipping cost are compiled by the US Census Bureau and downloaded from the Peter Schott's website. The second variable is winsorized at the 1% level.

	count	mean	sd	min	p1	p50	p99	max	skewness	kurtosis
ΔSC	16400	422.681	54122.833	-1.000	-0.526	-0.015	0.955	6931104.500	128.051	16398.000
ΔSC Winsorized	16400	0.004	0.197	-0.526	-0.526	-0.015	0.955	0.955	1.548	9.614

Table II

Shipping Cost Variation Persistence

This table shows the persistence of industry-year shipping cost variation by quintiles for the manufacturing industries traded on the AMEX, NASDAQ, or NYSE from 1974 to 2014. This variable is the first difference of the ratio between CIF (cost-insurance-freight) to the FOB (free-on-board) at the industry level. This variable is winsorized at the 1% level.

	from year t-1 to year t						from year t-5 to year t				
	Q_t^1	Q_t^2	Q_t^3	Q_t^4	Q_t^5		Q_t^1	Q_t^2	Q_t^3	Q_t^4	Q_t^5
Q_{t-1}^1	0.240	0.162	0.148	0.163	0.287	Q_{t-5}^1	0.249	0.187	0.148	0.184	0.232
Q_{t-1}^2	0.160	0.198	0.239	0.223	0.180	Q_{t-5}^2	0.181	0.222	0.226	0.208	0.163
Q_{t-1}^3	0.141	0.245	0.236	0.227	0.150	Q_{t-5}^3	0.151	0.217	0.243	0.228	0.161
Q_{t-1}^4	0.162	0.224	0.227	0.228	0.159	Q_{t-5}^4	0.170	0.218	0.224	0.213	0.175
Q_{t-1}^5	0.302	0.174	0.150	0.162	0.212	Q_{t-5}^5	0.243	0.165	0.166	0.178	0.248

Table III
Firm-level Variables Summary statistics

The table shows the summary statistics for the firm-year variables for manufacturing companies traded on the AMEX, NASDAQ, or NYSE from 1974 to 2014. These are retrieved from the Compustat. And the variables are: the growth rate for total sales (item *sale*) divided by total assets (item *at*); the inventories (item *invt*) pondered by total assets (item *at*); the gross profit margin, that is the ratio between gross profit (item *gp*) and the total revenues (item *revt*); and the net profit margin that is net income (item *ni*) divided by the total revenue (item *revt*). All these variables are winsorized the 1% level.

	count	mean	sd	min	p1	p50	p99	max	skewness	kurtosis
Δ Sales	37150	0.083	0.526	-0.770	-0.770	0.005	3.264	3.264	3.848	22.313
Δ Inventories	34326	0.038	0.387	-0.766	-0.766	-0.011	2.057	2.057	2.420	13.103
Δ Gross Profit Margin	37017	0.010	0.492	-1.961	-1.961	0.001	2.582	2.582	1.505	16.835
Δ Net Profit Margin	37017	-0.196	3.769	-19.086	-19.086	-0.094	18.417	18.417	-0.046	17.725

be due to their different sample selection procedure (not exclusively focus on manufacturing). For the inventory divided by the total assets the mean (median) is 3.8% (-1.1%). The growth rate of the gross profit margin has a mean (median) of 1.0% (0.1%). Finally, the mean (median) of the net profit margin first difference is -19.6% (-9.4%).

The firm-level accounting information will allow us to understand the reaction of each firm to the fluctuations on the shipping cost. Showing us the average impact of the supply chain pressure variation at the company level.

3 Regressions

Our goal is to estimate the sensitivity between the shipping cost variation and the firm characteristics. We estimate the following OLS model between the different firm-level variables $Y_{i,j,t}$ for each period against the shipping cost first difference for the same period, $\Delta SC_{j,t}$:

$$Y_{i,j,t} = \beta_0 + \beta_1 \Delta SC_{j,t} + u_{i,j,t} \tag{1}$$

Where $Y_{i,j,t}$ is the changes of sales to total assets, the inventory to total assets, the gross profit margin, or the net profit margin. On the right hand side, we have our coefficient of interest the β_1 estimating the sensitivity of our dependent variable to the $\Delta SC_{j,t}$.

The columns 1, 4, 7, and 10 on the Table IV report our baseline. From the four dependent variables, we only obtain statistically significant coefficient for our variable of interest when regressed against the gross profit margin. The coefficient of 0.018 means that 1% increase in the shipping cost leads to an increase of 0.018% of the gross profit margin.

However, the specification in the Equation 1 makes it difficult to correctly measure β_1. First, the dependent variable will differ from company to company, for example, will depend on their operations system, managing guidelines and culture. The condition $E[u_t|\Delta SC_t] = 0$ will not be satisfied, because the error term includes unobserved firm-level characteristics. Second, this regression may suffer from the omitted variable bias. As both sides of the Equation 1 could be affected by the fluctuations on the economy, the business cycle, inflating the actual impact that our main variable has on the dependent ones.

To solve the first problem identified on the previous paragraph, we run panel regression with firm fixed effects. This will control the unobserved firm specific time-invariant characteristics. Relative to the second, we propose to add the GDP growth to absorb the effects of the business cycles. This led us to run the following regression:

$$Y_{i,j,t} = \beta_0 + \beta_1 \Delta SC_{j,t} + \beta_2 \Delta GDP_t + \alpha_i + u_{i,j,t} \tag{2}$$

The specification is similar to the Equation 1, the difference lays on the added firm-level analysis that enable us to implement the firm fixed effects and the GDP growth to overcome the omitted variable bias.

Table IV

Firm-level variables sensitivity to the ΔSC

The table presents results of firm-year regressions of the firm's fundamental on the shipping cost variation. We use as dependent variable: the sales divided by total assets (columns (1) to (3)); the inventories pondered by total assets (columns (4) to (6)); the ratio between gross profit and total revenue (gross profit margin) in the columns (7) to (9); and net income normalized by total revenue, columns (10) to (12). The shipping cost first difference is reported at the industry-year level. Some regressions include controls for unobserved time-invariant firm characteristics and/or business cycle. The standard errors are reported in parenthesis with the significance level at the 10% (*), 5% (**), and 1% (***). All the variables are winsorized at the first and last percentile. The sample period is 1974 to 2014.

Dep. Var	Sales			Inventory			Gross Profit Margin			Net Profit Margin		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ΔSC	-0.016	-0.003	-0.003	-0.010	-0.011	-0.011	0.018**	0.024**	0.024**	-0.041	-0.042	-0.037
	(0.010)	(0.010)	(0.010)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.070)	(0.072)	(0.072)
ΔGDP			-0.127			-0.207**			0.111			3.623***
			(0.119)			(0.094)			(0.117)			(0.906)
Firm FE	NO	YES	YES	NO	YES	YES	NO	YES	YES	NO	YES	YES
Observations	37150	37150	37150	34326	34326	34326	37017	37017	37017	37017	37017	37017
R^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table IV columns 2, 5, 8 and 11 presents the results for the regressions with firm fixed effects. The coefficients size changes with the new setting, confirming the bias on our initial estimations. It is worth mention that the betas signs did not change, it seems that our baseline is correctly estimating the impact direction. However, we still only get significance for the gross profit margin regression. Also, in the remaining columns we add the control for business cycle. This exclusively affects the sensitivity estimation for the net profit margin. Again, we only get statistical significance for the gross profit margin regression. The coefficient of 0.024 implies that after controlling for unobserved firm specific characteristics and business cycle, an 1% increase in shipping cost will cause an 0.024% increase of the gross profit margin.

Comparing the coefficients between different specification we can see that their signals do not change, and the coefficient of interest is significant for all 3 models relatively to the gross profit margin. This could reveal that our model seems to capture the right impact direction, but when we look at the significance and explanatory power of our regression, we conclude that the right framework it stills yet to come.

From this initial empirical analysis, we cannot draw solid conclusions, as the R^2 is small for all the regressions and statistically significant coefficients are only obtained when using the gross profit margin. On the other hand, the sign consistent among different framework, may start to lead us to some conclusions relative to the effect direction.

References

Barrot, J.-N. and Sauvagnat, J. (Aug. 2016) Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks*. en, *The Quarterly Journal of Economics* **131**, 1543–1592.

Barrot, J.-N., Loualiche, E., and Sauvagnat, J. (Oct. 2019) The Globalization Risk Premium. en, *The Journal of Finance* **74**, 2391–2439.

Bradford, S. (2005) The extent and impact of food non-tariff barriers in rich countries. In: *Annual Meeting Theme Day.* Citeseer.

Cohen, L. and Frazzini, A. (Aug. 2008) Economic Links and Predictable Returns. en, *The Journal of Finance* **63**, 1977–2011.

Elliott, M., Golub, B., and Leduc, M. V. (2020) Supply Network Formation and Fragility. en, *SSRN Electronic Journal.*

Ersahin, N., Giannetti, M., and Huang, R. (2022) Supply Chain Risk: Changes in Supplier Composition and Vertical Integration. en, *SSRN Electronic Journal.*

Feenstra, R. C. (1996) *US imports, 1972-1994: Data and concordances.*

Hsu, P.-H., Li, K., and Tsou, C.-Y. (2020) The Pollution Premium. en, *SSRN Electronic Journal.*

Hummels, D. (2007) Transportation costs and international trade in the second era of globalization, *Journal of Economic perspectives* **21**, 131–154.

Hummels, D. (2008) Global trends in trade and transportation, *Benefitting from globalisation. Transport sector contribution and policy challenges. OECD/ITF*, 15–36.

McLain, S. (2021) Auto Makers Retreat from 50 Years of 'Just in Time'Manufacturing, *The Wall Street Journal, May* **3**.

Menzly, L. and Ozbas, O. (Aug. 2010) Market Segmentation and Cross-predictability of Returns. en, *The Journal of Finance* **65**, 1555–1580.

Mohanty, S. K., Aadland, R., Westgaard, S., Frydenberg, S., Lillienskiold, H. E., and Kristensen, C. (2021) Modelling Stock Returns and Risk Management in the Shipping Industry, *Journal of Risk and Financial Management* **14**. Publisher: Multidisciplinary Digital Publishing Institute, 171.

Park, J. S., Seo, Y.-J., and Ha, M.-H. (2019) The role of maritime, land, and air transportation in economic growth: Panel evidence from OECD and non-OECD countries, *Research in Transportation Economics* **78**. Publisher: Elsevier, 100765.

Schott, P. K. (Dec. 2007) The relative sophistication of Chinese exports: CHINESE EXPORTS. en, *Economic Policy* **23**, 5–49.

A Appendix

Table I

Wooldridge test for autocorrelation in panel data

The table shows the Wooldridge test for autocorrelation in panel data for the industry-year shipping cost and its variation relative to the manufacturing firms traded on the AMEX, NASDAQ, or NYSE from 1974 to 2014. These variables are the change of the ratio between CIF (cost-insurance-freight) to the FOB (free-on-board) at the industry level. The shipping cost are compiled by the US Census Bureau and downloaded from the Peter Schott's website. Both variables are winsorized at the 1% level.

H_0: no first-order autocorrelation vs H_1: first-order autocorrelation

	$F(1,496)$	$Prob > F$
SC Winsorized	178.116	0.0000
ΔSC Winsorized	0.015	0.9040