Network-dependent dynamics of innovation and production

Anton Pichler

Wolfson College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

Abstract

Complex networks have been demonstrated to play an important role in many economic systems. This thesis aims at further integrating the science of complex networks with economic theory with a focus on dynamic aspects of two types of economic networks: technology/innovation and production networks.

Technological evolution is often described as a recursive process whereby the recombination of existing components leads to new or improved technological components. While this idea implies a network structure of technological interdependencies, very little has been done to establish empirically that these interrelations help predict future innovation dynamics. We propose a simple model of network-dependent knowledge creation that we calibrate to patent data. We find that simple time series models yield fairly good predictive performance and represent a good benchmark for alternative forecasting models. Incorporating information on the network of patent citations, however, can substantially improve patenting forecasts.

The second part of this thesis studies the dynamics of macroeconomic variables using production network models. We propose a macroeconomic model which we calibrate to the UK economy to quantify sectoral and aggregate impacts of the Covid-19 pandemic. Beside the production function assumption and the estimation of first-order shocks, we find that inventories and input bottlenecks are crucial in economic prediction. To better understand the driving forces behind these predictions, we study simpler dynamic input-output models. We suggest a simple mathematical optimization procedure that allows us to determine lower bounds of shock propagation. We find that the minimal shock propagation can be substantial, but much smaller than when compared to decentralized rationing behaviors of firms. In particular, the estimated economic impact strongly depends on the density of the underlying production network.

Acknowledgements

Doing this DPhil would not have been the same without the support of many people:

Doyne Farmer, it has been an enormous privilege to be your student. I have learned too many things from you to put them into a short paragraph. Thank you for all the support, the open-minded environment and for being a continuous source of inspiration.

Francois Lafond, thank you for a countless number of discussions which always taught me something new. Rarely I have met such a careful thinker who challenged my way of reasoning like you did. This thesis has greatly benefited from it.

My "older" DPhil colleagues and co-authors: Maria del Rio-Chanona and Marco Pangallo. Thank you for all the tutoring, new perspectives and fun.

The whole INET gang including Andrea Bacilieri, Blas Kolic, Torsten Heinrich, Kieran Marray, Penny Mealy, Luca Mungo, Valentina Semenova, Vilhelm Verendel, Rupert Way, Julian Winkler and many more: The frequent chats, cycle trips, football and squash games, the regular breakfast gossip and many more small occasions made my time in Oxford an invaluable experience.

I am grateful to Cameron Hepburn for giving me lots of useful advice and to Christian Diem for discussions that frequently untie knots in my thinking. I am also grateful to Dorothy Nicholas, Sandhya Patel and the whole INET admin team. Without them I would not have survived in the haze of the university jungle.

I acknowledge the financial support from the Oxford Martin School Programme on the Post-Carbon Transition, IARPA, the Theodor Körner Fonds and Partners for a New Economy.

Very special thanks to my parents, Maria and Johann, who taught me self-reliance and enabled me to create my own, individual trail through life that lead to this thesis.

I am mostly indebted to Carina who has gone with me through the best and worst of times, and has become my very favorite Sapiens. Thank you for providing me with unending love and support, and enriching every day of my life. This thesis is dedicated to you.

Contents

Introduction

Complex networks have been demonstrated to play an important role in many economic systems. This thesis aims at further integrating the science of complex networks with economic theory by focusing on dynamic aspects of two types of economic networks: technology/innovation and production networks.

Innovation and technical change have been key drivers of past economic growth (Solow 1956) and structural change (Schumpeter 1939, Dosi 1982) and will also play a major role in the "green transition" of modern economies (Popp 2016, Farmer et al. 2019, Pless et al. 2020). Technological progress is a path-dependent process of recombining existing knowledge, motivating that the evolution of interdependent technologies should be coupled (Usher 1954, Kauffman 1993, Fleming 2001, Arthur 2009). While this idea implies a network structure of technological interdependencies, very little has been done to establish empirically that these interrelations help predict future innovation dynamics. By using patent data, we show that patenting rates of technological classes can be fairly well predicted by conventional time series models. Despite the good predictive performance of time series models, we find that incorporating the network of patent citations can still improve patenting forecasts.

Similarly, complex supply chains are at the center of modern economic systems, having effects on economic growth (McNerney et al. 2018), business cycles (Long & Plosser 1983, Acemoglu et al. 2012, Pangallo 2020) and propagation of economic shocks (Leontief 1936, Bak et al. 1993, Miller & Blair 2009). In the second part of this thesis we focus on the dynamics of macroeconomic variables using production network models. We propose a macroeconomic model which we calibrate to the UK economy to quantify sectoral and aggregate impacts of the Covid-19 pandemic. Beside the production function assumption and the estimation of first-order shocks, we find that inventories and input bottlenecks are crucial in economic prediction. To better understand the driving forces behind these predictions, we study simpler input-output models. We show that existing input-output models do not yield feasible market allocations when calibrated to simultaneous supply and demand shocks as observed in the

ongoing pandemic. We therefore introduce a mathematical optimization procedure, allowing us to determine lower bounds of shock propagation. We find that the minimal shock propagation can be substantial, but much smaller than when compared to decentralized rationing behaviors of firms. Moreover, the estimated economic impact strongly depends on the density of the underlying production network.

The relevance of innovation and production networks for the evolution and stability of economic systems has been a key motivation for this thesis. An original idea was to apply combined innovation and production network models similar to the ones considered here to study the green transition of the economy. The dramatic consequences of the Covid-19 pandemic, however, convinced us to re-focus our efforts on more short-term disaster modeling. In the following chapters we therefore look at these types of networks independently from each other, but recognize that we expect strong interactions between innovation dynamics and economic production, in particular in the long-run. We leave this undertaking for future research.

The sectoral World Input-Output Database (Timmer et al. 2015) and the extensive record of US utility patents represent the basic data sources which this thesis is built on. While these are highly valuable datasets, we believe that our research could be improved and extended in the future with more detailed data. Sectoral input-output data comes with the advantage of being fully consistent with national accounts. As our analysis shows, however, many of the results depend on subtleties such as behavioral rules of economic agents or production mode specificities. The sectoral data considered here might be to coarse for uncovering some of these relationships and we expect further insights from future work that is able to study more fine grained firm- or establishment-level production network data.

Similarly, patent data comes with the advantage of being well-curated and of high-quality. However, patents are obviously an incomplete measure of invention and innovation efforts. An ideal dataset for studying innovation dynamics would also feature information on technological subcomponents, prices, research investment and deployment. At present, such data is not available and we thus represent innovation dynamics with the ready available patent data, even if this compromises the conclusions we can draw.

We next summarize the various chapters of the thesis in some detail.

In Chapter 1 we investigate statistical properties of patenting growth time series, to gain a better understanding of innovation efforts in various technological classes. In particular, we ask: How predictable are future rates of patenting in distinct technological categories? We analyze data on over 10 million US utility patents that

have been granted between 1836 and 2017, and their groupings into technological classes. After showing that technology-specific time series exhibit a strong common trend, we find that a geometric random walk with negatively autocorrelated errors fits the data fairly well. Results from an extensive hindcasting analysis show that the geometric random walk predicts patenting dynamics better than we would expect if the underlying data generation process would truly follow a random walk. Building on our observation of strong common trends in the patenting system, we further demonstrate that the predictive performance can be improved by pooling technology-specific parameters into common system parameters. This is particularly true for long forecasting horizons and short estimation windows. We also compare the geometric random walk model with an alternative network model which previously has been proposed in the literature and find that the random walk model vastly outperforms the more complicated model. These results do not prove that patent networks are uninformative for future patenting dynamics. Instead, our results establish a proper null hypothesis, to which more sophisticated models should be benchmarked against.

In Chapter 2 we extend the work of the previous chapter and focus on network effects in the yearly frequency of patents appearing in a given technological class. Here, we combine the patenting activity time series with data on patent citations. The data allows us to construct weighted temporal citation networks, where nodes represent distinct technological categories and weighted links correspond to the number of citations between them. We observe that a technology's patenting growth rate is not independent from its location in the network. Instead, if technologies are neighbors in the network, they tend to exhibit similar growth rates. Motivated by this observation, we propose a simple network model of knowledge creation that can be used for forecasting future patent frequencies in given technological classes. We solve the model for the steady state and calibrate it to the data, using maximum likelihood techniques inspired from spatial autoregressive models. Estimated model parameters suggest strong network effects in the evolution of technological classes. We validate our model by making out-of-sample forecasts and benchmark them against simpler time series models that do not take network effects into account. We find some remarkable improvements in predictive power, providing further evidence on network effects in patenting dynamics.

In Chapter 3 we study the economic effects on the UK economy resulting from the Covid-19 pandemic, using a dynamic sectoral input-output model. The model which is inspired by previous work on the economic response to natural disasters allows us to study upstream and downstream propagation of industry-specific demand and

supply shocks out-of-equilibrium. We argue that standard production functions, at least in their most parsimonious parametrizations, are not adequate to model input substitutability in the context of Covid-19 shocks. We use a survey of industry analysts to evaluate, for each industry, which inputs were absolutely necessary for production over a short time period. We calibrate our model on the UK economy and study the economic effects of the lockdown that was imposed at the end of March 2020 and gradually released in May 2020. We explore the behavior of the model and establish a series of interesting theoretical results. Using simpler scenarios with only demand or supply shocks, we find that popular metrics used to predict a priori the impact of shocks, such as output multipliers, are only mildly useful. We further show that the production network can translate first-order shocks to higher-order impacts in non-trivial ways. For example, we find that overall macroeconomic impacts are higher when only supply shocks are considered compared to the case where both supply and demand shocks are applied. Switching off demand shocks increases demand and supply imbalances in the system, further facilitating coordination failures and thus leading to an overall adverse impact. We establish that input criticality, inventory levels and assumptions regarding the production function play a key role in the estimation of economic impacts. These features, frequently neglected in other macroeconomic studies, similarly play an important role in the economic recovery after unwinding social distancing measures. We also use the model to reproduce aggregate and industrial dynamics of the UK economy during the first wave of the pandemic. We test our model on recent sectoral and aggregate economic data and find that our model is able to predict how supply and demand shocks propagate through the economy very well.

In Chapter 4 we again study the propagation of pandemic shocks through production networks, but use a simpler framework that is closer to conventional input-output models. A central motivation of this paper is that most input-output models cannot deal with supply and demand shocks simultaneously, but only with either supply or demand shocks at a time. Many natural and anthropogenic disasters, however, frequently affect both the supply and demand side of an economy, with the Covid-19 pandemic being a striking example. We first show that conventional input-output models do not yield feasible economic allocations if initialized with the pandemic shocks. Using the first-order shocks as exogenous constraints, we propose a linear optimization approach that yields a lower bound on total shock propagation. We find that even in this best case scenario network effects substantially amplify the initial shocks. To obtain more realistic model predictions, we study the propagation of

shocks bottom-up by imposing different rationing rules on industries if they are not able to satisfy incoming demand. Our results show that economic impacts depend strongly on the emergence of input bottlenecks, making the rationing assumption a key variable in predicting adverse economic impacts. We further establish that the magnitude of initial shocks and network density heavily influence model predictions.

Chapter 1

Persistence and predictability of patenting

This chapter is on the paper Pichler, Lafond & Farmer (2021).

1.1 Introduction

The role of technology as an essential driver of economic and social change has motivated many attempts to forecast technological progress (Ayres 1969, Jantsch 1967, Martino 1993, National Research Council 2009, Porter et al. 2011). Technological evolution depends on many factors, including technical features, institutions, interactions with other technologies, chance and serendipity. This makes predicting individual inventions difficult (Freeman & Soete 2012, Mokyr 1992, McNerney et al. 2011, Fink et al. 2017). Nonetheless, for a given class of technology, such as transistors or the manufacture of ammonia, metrics such as cost per unit of performance can be usefully forecasted at least twenty years into the future based on historical performance (Nagy et al. 2013, Magee et al. 2016, Farmer & Lafond 2016). This works because the performance of a given technology tends to improve roughly exponentially for long periods of time. The evidence suggests that, while the rate of improvement sometimes changes, this is rare, so that a simple trend extrapolation provides a useful prediction. Unfortunately, we still lack historical performance data for most technologies, which limits our ability to make systematic forecasts of technological progress.

Patents provide what is perhaps the best existing record of technological change (Schmookler 1966, Griliches 1990, Hall et al. 2001). The US patent database spans almost two centuries and contains more than 10 million patents. Patents do not necessarily indicate performance improvement (Triulzi et al. 2020), but they are at

least an indicator of human effort, making the rate at which patents are granted in a given domain interesting for its own sake.

Technologies do not develop in isolation from one another. Patents contain multiple technology classifications (see e.g. Breschi et al. (2003) and Youn et al. (2015)) and they cite prior patents that influenced them, which makes it possible to study how technology classes interact (Napolitano et al. 2018).[1] We were inspired to perform the work here by Acemoglu et al. (2016), who constructed an influence network between technology classes based on citations and used this to predict future patenting rates. They predicted the number of future patents in a given technology class, showed that including past patenting rates from related classes improved their results, and argued that this demonstrates the existence of network effects.

Because patenting rates are very persistent, however, these results require careful interpretation. The number of patents that are granted in a given class in a given year tends to be very similar to the number of patents that were granted in that class in previous years. As originally shown by (Yule 1926), any persistent time series is a good predictor of other persistent time series. Thus, to demonstrate the existence of network effects, one must first establish a careful benchmark.

We demonstrate this point in Fig. 1.1, where we compare predictions based on a simple random walk model to those of Acemoglu et al. (2016). They used the NBER subcategory (SCAT) classification (Hall et al. 2001), which divides technologies into 36 classes, and extrapolated the number of patents in a given class in a given year from the number of patents in the preceding ten years. To predict class i they used the patenting rate for class i as well as the patenting rates from any classes that were cited by patents in class i, weighted according to their citation rates. The left panel of Fig. 1.1 compares actual with predicted values obtained from this network model. We describe the network model and prediction procedures in full detail in Appendix A.4.

For comparison, we used a simple geometric random walk with drift, in which the prediction for each year for class i is based on the patenting rate in the previous year for class i alone. As seen in the right panel of Fig. 1.1, this results in much better predictions than the network method of Acemoglu et al. (2016).

[1] Citations are not fully available as early patents contained in-text citations, in contrast to modern bibliographies, and are thus harder to parse. Technological classifications are not historical classifications, but reclassification of old patents into the most recent classification system (Lafond & Kim 2019).

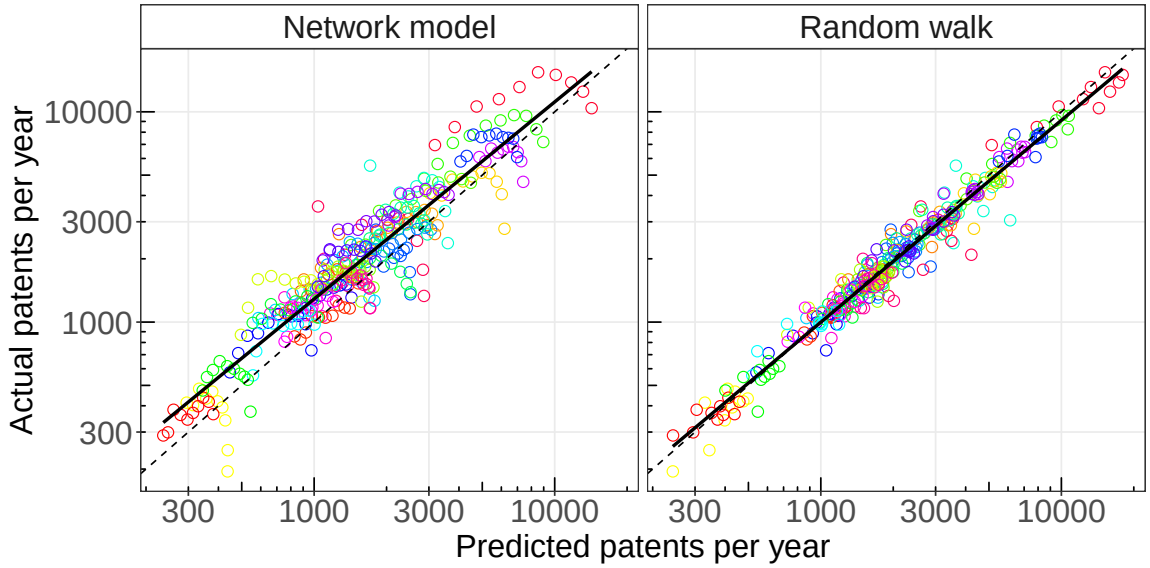

Figure 1.1: **Predicted vs. actual annual patenting activity.** The left panel shows predictions obtained from the network model introduced in Acemoglu et al. (2016); see Eq. (A.4) in Appendix A.4. The right panel shows predictions from a geometric random walk; Eq. (1.2). Colors correspond to the 36 SCAT technology classes. The clustering of colors hints at persistence in patenting activity. The dashed line in each panel is the 45 degree line, corresponding to a perfect prediction, and the solid lines are regression lines obtained from a panel model. Details on the network model and its predictions are discussed in Appendix A.4.

The results in Fig. 1.1 indicate that the number of patents filed in a given class in a given year, which we will call *patenting activity*, is very persistent. This persistence is interesting in and of itself. In this paper we develop a simple univariate time series model for predicting patenting activity and use it to characterize the persistence and heterogeneity of patenting activity in different technology classes. This provides a good benchmark against which the predictive power of network models can be measured. In Appendix A.4 we perform several quantitative tests, such as forecast encompassing and network randomization, on the results of Acemoglu et al. (2016). Our results suggest that predictive performance in Acemoglu et al. (2016) is driven mostly by persistence in patenting activity, rather than by network effects.

The data we use for this study is from the US Patent Office (USPTO) from 1836–2017. We use the Cooperative Patent Classification (CPC), which at the 4-digit level divides the data into 631 categories. Patent activity tends to increase with time. To minimize problems due to nonstationarity, we difference the time series, and base our analysis on logarithmic annual rates of change in patenting activity. As our baseline model we choose a simple geometric random walk with drift and with autocorrelated error terms (ARIMA(0,1,1)) and study its ability to predict future patenting activity on different time horizons. One of the advantages of this model is that it is possible

to characterize its forecasting performance *a priori* for any time horizon (Farmer & Lafond 2016). While the agreement with the data is not perfect, it provides a reasonable approximation.

In addition to providing a useful forecasting benchmark, this model provides us with an instrument to study the heterogeneity of patenting activity across different technology classes. The average growth rate in patenting activity, which corresponds to the drift term of our time series model, is of particular interest. A high average growth rate signals that a technology domain has been increasing as a focus of innovation. Naïve empirical estimates of the drift terms yield a broad distribution of rates ranging from -1.4 to 24.8%. Its distribution is strongly right skewed with a heavy tail, indicating a small number of classes with exceptional rates of growth. However, analysis using the time series model demonstrates that most of this variation can be explained by statistical estimation error. The exception is the tail, which is mostly occupied by "high-tech" technologies such as 3D printers. These have growth rates larger than 10% and lie outside of the range of statistical variation with respect to our simple null model.

As another test of heterogeneity vs. estimation error we compare univariate models to a pooled model. In the univariate models the estimated drift term for class i is based on data from class i alone, whereas in the pooled model the estimated drift term for class i is the cross-sectional mean of the estimated drift terms for all classes. There is a trade-off between heterogeneity, which favors univariate models, and estimation error, where pooling has an advantage due to more data (see e.g. Wang et al. (2019)). For a time horizon of one year the prediction accuracy of the two models is similar, but on longer time horizons the pooled model always performs better. When we use an historical window of 10 years to estimate the drift term, the performance of the pooled model is substantially better, but with longer historical windows this advantage erodes. Thus estimation error always dominates heterogeneity.

We also find that patenting activity is remarkably stationary. We test for this by varying the length m of the sampling window that we use to estimate the drift parameter. In a nonstationary world where the drift parameter varies substantially, there is a trade-off between shorter windows, which can cope with nonstationarity, and longer windows, which provide superior statistical estimation. We find that we get the best results when we use very long estimation windows (40-50 years). This indicates that the growth rate of patenting activity is quite persistent. There is a strong negative temporal autocorrelation in patenting activity growth rates, with an

average value of -0.36, indicating mean reversion in annual patent activity, but this does not affect the long term persistence.

We now do some exploratory analysis of the data and then present the results of the model. The model we develop here provides a benchmark for other predictive models to beat. In Chapter 2 we will show that it is indeed possible to beat this model using network effects.

1.2 Results

We use virtually all US patents ever granted prior to 2017 and classify patents according to the Cooperative Patent Classification (CPC) scheme.[2] We describe the data in more detail in Appendix A.1. There are over 10 million utility patents, each of which is associated with at least one CPC code. We denote the number of published patents in technological category i in year t by $p_{i,t}$.

The upper panel of Fig. 1.2 shows patent activity in log-scale over time for 1-digit classification levels (which are also called sections). In the early part of the 19^{th} century all of the sections show rapid growth, but then the overall growth rate slows down. The figure makes it clear that all the sections tend to move in tandem. In fact, as we show in Appendix A.2, at the 1-digit level, 77% of the variance in the growth rates $g_{i,t} = \log(p_{i,t}/p_{i,t-1})$ is explained by a common factor, which is very close to overall patenting activity. This is much less so, however, at the level of 631 4-digit technology categories (also called subclasses) where overall patenting activity only explains 13% of the variance. To our surprise, the explanation of variance across levels is not nested, i.e. in most cases the patenting activity for any given category explains very little of the patenting activity in the sub-categories contained within it, beyond what is already explained by the overall patenting rate. See the discussion in Appendix A.2.

To get a better feeling for how the importance of the sections has changed through time, in the bottom panel of Fig. 1.2 we show the relative levels of patenting activity. Electricity has had the fastest growth, starting with only a few patents per year, but becoming the most active section. It now roughly ties with Physics, which has also grown rapidly. Textiles, in contrast, have declined in relative terms by more than a factor of ten, and Fixed Constructions has also declined substantially.

[2]See `https://www.cooperativepatentclassification.org/index` for further information on technological classification.

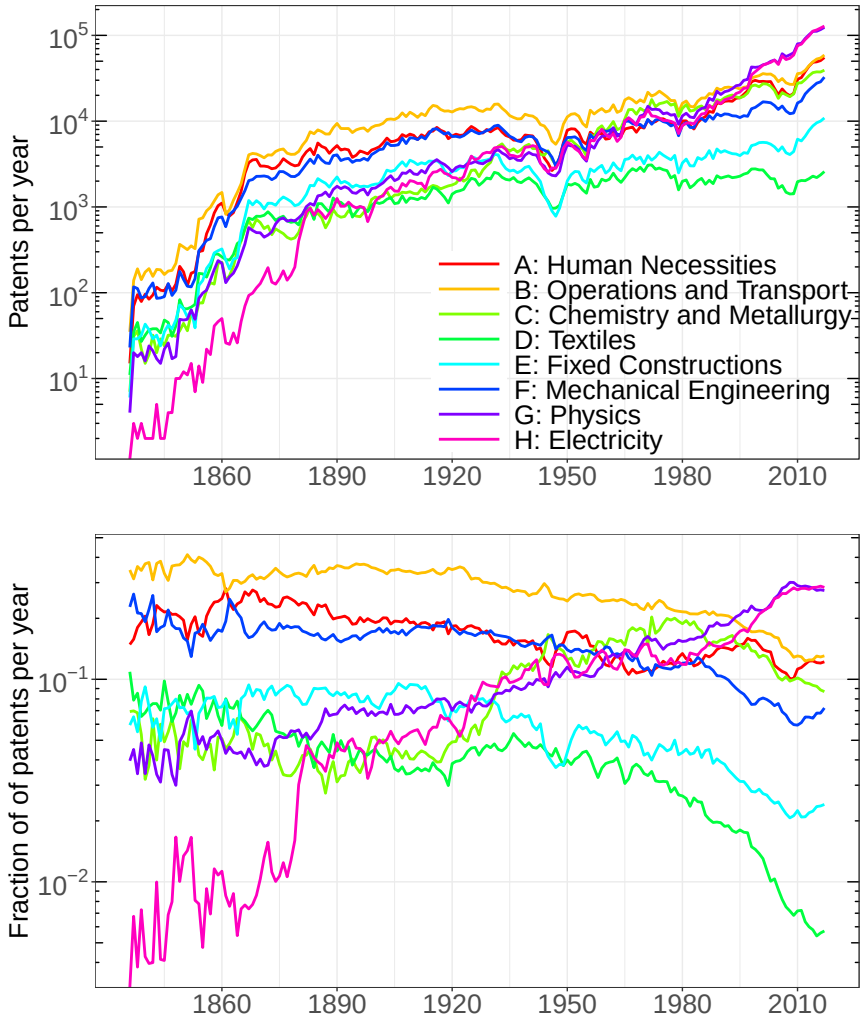

Figure 1.2: Upper panel: Number of US utility patents granted per year in each 1-digit CPC class (also called section). Lower panel: Relative share of each section, i.e. $p_{i,t}/\sum_i p_{i,t}$. The y-axes are on log-scale. For data source see Appendix A.1.

1.2.1 Time series modeling

We now investigate the predictability of patenting activity using a simple univariate time series model. We do this for the most granular 4-digit CPC subclass level, which has 631 categories. After some exploratory data analysis, we settled on the geometric random walk with autocorrelated error terms, i.e. an ARIMA(0,1,1) or equivalently IMA(1,1) model. Letting the logarithm of patenting activity be $y_{i,t} = \log(p_{i,t})$, the model is

$$y_{i,t} = y_{i,t-1} + \mu_i + \epsilon_{i,t} + \theta_i \epsilon_{i,t-1}, \tag{1.1}$$

where μ_i is the drift parameter, θ_i is the moving average parameter and the noise is normally distributed and centered around zero, i.e. $\epsilon_{i,t} \sim N(0, s_i^2)$. The variance is

$\mathrm{Var}(y_{i,t} - y_{i,t-1}) \equiv \sigma_i^2 = (1+\theta_i^2)s_i^2$. If there is no autocorrelation, i.e. $\theta_i = 0$, Eq. (1.1) simplifies to an (uncorrelated) geometric random walk with drift

$$y_{i,t} = y_{i,t-1} + \mu_i + \epsilon_{i,t}, \tag{1.2}$$

and $\mathrm{Var}(y_{i,t} - y_{i,t-1}) = \mathrm{Var}(\epsilon_i) = s_i^2$.

To get an understanding of the behavior of individual technologies, we estimate the parameters μ_i, s_i and θ_i for Eq. (1.1) using maximum likelihood. We do this individually for each of the technology subclasses, and show the parameter distributions in Fig. 1.3 (blue histograms).[3]

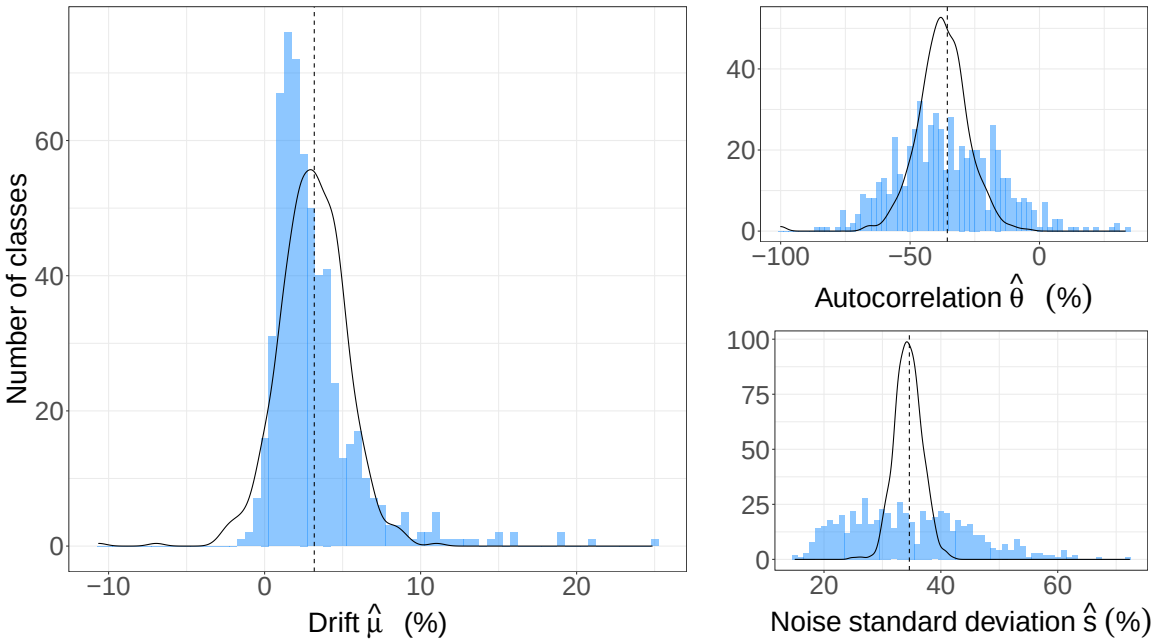

Figure 1.3: **Histograms of autocorrelated random walk parameters.** The parameters are obtained from estimating the model in Eq. (1.1) on the time series ranging from 1867 to 2017 for all subclasses with at least 20 years of data. The blue histograms represent the empirical parameter estimates and the dashed vertical lines indicate sample means. Most technologies exhibit positive drift parameters, but negatively autocorrelated errors. The kernel density plots (solid lines) represent the non-heterogeneous null model and are obtained from estimating the parameters on surrogate data. The data is is created by simulating the model with common parameters $\bar{\mu} = \sum_i \hat{\mu}_i/N$, $\bar{\theta} = \sum_i \hat{\theta}_i/N$ and $\bar{s} = \sum_i \hat{s}_i/N$ for all i and matching the length of the original time series.

The parameter estimates for the drift term $\hat{\mu}_i$ are almost all positive. The distribution is strongly right skewed and has a long tail. Because there is substantial statistical estimation error, it is not clear how much of the heterogeneity shown in

[3]We exclude all data points prior to 1867 due to the obvious break in the time series observed in Fig. 1.2. We include a time series only if there are at least 20 data points available and use only the data since the subclass came into existence; the mean existence time is 130 years – see Appendix A.3 for a histogram.

Fig. 1.3 is real. To better understand this we compare the empirical distributions to that of a simple null model with homogeneous parameters. We do this by simulating Eq. (1.1) to generate surrogate time series. We match the length of each of the time series and use the same parameters,

$$\bar{\mu} = \sum_{i=1}^{N} \hat{\mu}_i / N = 3.2\%, \quad \bar{s} = \sum_{i=1}^{N} \hat{s}_i / N = 34.6\%, \quad \bar{\theta} = \sum_{i=1}^{N} \hat{\theta}_i / N = -35.6\%,$$

corresponding to cross-sectional means, in all of the simulations. We then re-estimate the parameters of Eq. (1.1) for each surrogate time series. The overlaid solid lines indicate the estimated distribution of the parameter estimates for the surrogate time series. Since these have no heterogeneity by construction, they provide a useful benchmark. This suggests that most of the apparent heterogeneity in $\hat{\mu}_i$ is due to statistical estimation error. Nonetheless, the tail is clearly outside of the surrogate distribution, suggesting that the rate of growth of the twenty or so technologies in the tail is truly exceptional.

The estimates for s_i and θ_i shown in the right panels of Fig. 1.3 also naïvely suggest substantial heterogeneity, but as for μ_i, a comparison to the null distribution indicates that much of this is due to estimation error. Almost all of the estimates of the autocorrelation parameter θ_i are negative, but most of the variability is consistent with estimation error. The standard deviation parameter s_i ranges from roughly 15% to 70%, and seems to exhibit substantial heterogeneity beyond that caused by estimation error. However, as we show in Appendix A.3, Fig. A.3, there is a strong anticorrelation between $\hat{s}_i$ and average patenting activity. This is to be expected: if patent activity followed a Poisson process, for example, one would expect lower average patenting activity should imply larger year to year fluctuations in relative terms, so some of the apparent heterogeneity may be caused by this effect.[4]

Table 1.1 zooms into the tail of the distribution of μ_i in Fig. 1.3, listing the 20 technologies with the largest growth rates. They are all "high tech" domains such as 3D printing, bioinformatics, photovoltaic modules and nanotechnology. Many of these subclasses are very young, with little or no patenting activity until the 1980s. As we show in Appendix A.3, Fig. A.4, there is a negative relationship between the drift and the age of a technology more generally. However, the list also includes a few technologies such as electrical digital data processing and semiconductor devices for storage that have existed for more than 70 years. Eleven of these twenty technologies

[4]As shown in Appendix A.3, Fig. A.2, there is a negative correlation between autocorrelation and volatility parameters $\hat{\theta}_i$ and $\hat{s}_i$ (Pearson corr. of -0.53), but the other two possible parameter pairs show no discernable correlation.

are in the Physics section. Surprisingly, there is only one Electricity technology (photovoltaic modules).

Subclass	Title	Drift	Yearly number of patents		
		$\hat{\mu}_i$	1945-80	1981-2000	2001-17
G06N	Computer systems based on computational models	24.8	2	96	600
B33Y	Additive manufacturing (3D printing)	20.9	0	35	336
G16B	Bioinformatics	18.9	0	6	193
B81B	Microstructural devices/systems	18.8	0	10	254
G06T	Image data processing or generation	16.2	10	548	3,785
B81C	Processes/apparatus for microstructural devices/systems	16.1	0	18	353
G16C	Computational chemistry and material science	15.0	0	4	48
A01H	New plants or processes for obtaining them	14.8	2	82	652
G16H	Healthcare informatics	14.5	1	50	675
B82B	Nanostructures formed by manipulating atoms/molecules	13.6	0	2	25
G06Q	Data processing systems/methods	13.2	27	393	6,366
G01Q	Scanning-probe microscopes and techniques	12.6	0	44	121
H02S	Photovoltaic modules	11.9	3	23	200
C40B	Combinatorial chemistry	11.4	0	51	228
G11C	Semiconductor devices for storage	11.0	221	986	3,942
G02F	Manipulating optical properties (e.g. intensity, color, phase)	11.0	84	640	2,787
B82Y	Applications/measurement/analysis of nanostructures	11.0	18	393	2,364
C12Y	Enzymes	10.9	5	124	514
G06F	Electrical digital data processing	10.8	277	3,135	23,954
G09G	Control of indicating devices	10.6	58	399	2,404
Average of all subclasses		3.2	109	231	571

Table 1.1: **Top 20 technological classes regarding their drift estimates $\hat{\mu}_i$.** The column showing the drift parameter estimates (in %) corresponds to zooming into the right tail of the distribution in Fig. 1.3. The three columns to the right show average numbers of patents per year during the time horizons indicated. Note that the CPC 4-digit titles are self-made shortcuts. For full titles we refer to the link in Footnote 2. See Table 1.2 for the names of the section corresponding to the first letter in the subclass.

Table 1.2 presents some summary statistics for the parameter values of the subclasses, aggregated into sections. The subclasses for Physics and Electricity have almost the same average drift parameters, but the standard deviation for Physics is about twice as high as it is for Electricity. The average drift parameters for Physics and Electricity are almost four times larger than for Textiles.

1.2.2 Properties of the forecasting errors of the model

To test the adequacy of the time series models we perform an extensive hindcasting analysis on the whole patenting system. This means that we split every time series into two parts, an estimation and a prediction window. We then fit the model to the data from the estimation window and forecast patenting rates of the prediction horizon (out of sample). The forecast for horizon τ is given by

$$\hat{y}_{i,t+\tau} = y_{i,t} + \hat{\mu}_i \tau + \hat{\theta}_i \hat{e}_{i,t}, \tag{1.3}$$

Section	Subclasses	Patents	Years	AV($\hat{\mu}$)	SD($\hat{\mu}$)	AV($\hat{\theta}$)	SD($\hat{\theta}$)	AV($\hat{s}$)	SD($\hat{s}$)
A: Human Necessities	82	1.69	149	2.1	1.9	-33.5	16.5	31.1	8.9
B: Operations and Transport	166	2.49	138	2.7	3.1	-34.1	20.4	33.2	10.7
C: Chemistry and Metallurgy	86	1.39	111	3.7	2.3	-42.2	19.3	38.8	8.9
D: Textiles	38	0.26	121	1.4	1.2	-48.9	12.2	38.1	10.0
E: Fixed Constructions	30	0.50	158	1.6	0.7	-29.0	16.2	29.0	8.0
F: Mechanical Engineering	96	1.28	141	2.5	1.5	-37.4	16.2	34.6	11.7
G: Physics	83	2.38	107	5.5	4.6	-36.4	23.6	36.0	10.2
H: Electricity	50	2.25	117	5.3	2.1	-24.6	18.6	37.4	9.8
Total	631	10.14	130	3.2	3.0	-35.6	19.6	34.6	10.4

Table 1.2: **General patent statistics.** *Subclasses* refers to number of distinct 4-digit classes within a section (1-digit category). *Patents* is the total number of unique patents in millions per section and *Years* represents the average time series length in years. AV($\hat{\mu}$), AV($\hat{\theta}$) and AV($\hat{s}$) denote averages of the estimated drift, autocorrelation and volatility parameters of the model in Eq. (1.1) using all Subclasses from 1867 onward with at least 20 years of non-zero patenting. SD($\hat{\mu}$), SD($\hat{\theta}$) and SD($\hat{s}$) are the corresponding standard deviations of the estimates. Numbers in the six rightmost columns are expressed in percentages.

where $\hat{e}_{i,t}$ is the t^{th} residual obtained from the model estimation.

We compute the forecast errors from these predictions and repeat this for every possible estimation and forecast window, as well as for every technology. Previous work from Sampson (1991), Clements & Hendry (2001), Farmer & Lafond (2016) and Lafond et al. (2018) has shown that the forecast errors of the random walk model can be normalized so that their distribution is independent of the model's parameters, and only depends on the estimation window length m and the forecasting horizon length τ. This makes it possible to pool the results from testing the model with many different technologies, forecast horizons and estimation windows. It also makes it possible to predict the distribution of forecast errors *a priori*.

We denote the forecast τ steps ahead as $\hat{y}_{i,t+\tau}$ and the corresponding forecast error as

$$\varepsilon_{i,t+\tau} = y_{i,t+\tau} - \hat{y}_{i,t+\tau}. \tag{1.4}$$

If the actual process is consistent with the model, the mean squared normalized forecast error (MSNE) can be approximated as

$$\Xi_i(m,\tau,\theta_i) \equiv \mathbb{E}\left[\left(\frac{\varepsilon_{i,t+\tau}}{\hat{\sigma}_i}\right)^2\right] \approx \frac{m-1}{m-3}\frac{\mathcal{A}(\theta_i,\tau,m)}{(1+\hat{\theta}_i^2)}, \tag{1.5}$$

where $m > 3$ denotes the number of data points in the estimation window and

$$\mathcal{A}(\theta_i,\tau,m) = -2\hat{\theta}_i + \left(1 + \frac{2(m-1)\hat{\theta}_i}{m} + \hat{\theta}_i^2\right)\left(\tau + \frac{\tau^2}{m}\right). \tag{1.6}$$

The circumflex in $\hat{\sigma}_i$ makes clear that forecast errors are normalized with the *estimated*, not the *true*, standard deviation. The fact that we need to estimate the

variance leads to the prefactor $(m-1)/(m-3)$. Note that Eq. (1.5) is only approximate, but has been shown to work well for sufficiently large estimation windows m (Farmer & Lafond 2016). We use Eq. (1.5), but for reasonably large m and τ, it simplifies to

$$\Xi_i(m, \tau, \theta_i) \approx \frac{(1+\hat{\theta}_i)^2}{1+\hat{\theta}_i^2}\left(\tau + \frac{\tau^2}{m}\right).$$

This expression makes it clear that for short time horizons the squared error grows approximately linearly due to diffusion, and for long time horizons it grows quadratically due to estimation error. The squared error is equal to zero when the autocorrelation parameter $\theta_i = -1$ and is an increasing function of θ_i, so that at $\theta_i = 1$ the error is twice as large as it is at $\theta_i = 0$.

To pool experiments we can use the fully normalized forecast error

$$\kappa_i(m, \tau, \theta_i) \equiv \sqrt{\frac{(1+\theta_i^2)}{\mathcal{A}(\theta_i, \tau, m)}}\left(\frac{\varepsilon_{i,t+\tau}}{\hat{\sigma}_i}\right) \sim t(m-1). \tag{1.7}$$

If the assumptions of the model are satisfied, the data will collapse onto the Student t-distribution. Note that in contrast to the mean normalized squared error Ξ, the factor $\mathcal{A}$ is in the denominator rather than the numerator. This means that, all else equal, the fully normalized error is now a decreasing function of θ. The fully normalized error κ is useful for statistical testing as it is possible to collapse results for different values of τ onto the same distribution.

1.2.3 Hindcasting as a test of the model

To assess the model's goodness of fit to the data we perform hindcasting using a fixed rolling window. We then compare the forecast errors obtained using hindcasting to the forecast errors expected under the null that the data follows the model. For the sake of data quality we restrict the sample to the time horizon ranging from 1945 to 2017 and only include technologies with at least one patent every year, leaving us with 510 distinct technological classes.[5] Here we use the simple geometric random walk model (Eq. (1.2)) with $\theta = 0$ when making predictions, but compare its predictive performance with non-zero autocorrelation in the subsequent chapter.

For the hindcasting experiments we split every time series into an estimation window of length m and predict all remaining future time points. To obtain a large number of forecasts, we incrementally roll the estimation window over the whole time

[5]We have also experimented with alternative sampling choices and found that results presented here are very robust with respect to sampling details.

series to make essentially all possible forecasts for a given m. Thus, for a time series of length T we fit the random walk $T - m - 1$ times and make as many forecasts for varying time horizons as we can. For example, for a patenting time series ranging from 1945 to 2017 and $m = 10$, we fit the model to all possible windows of eleven data points, i.e. {1945–55, 1946–56, 1947–57, ..., 2006–16}[6], and make forecasts of patenting activity for the time horizons {1956–2017, 1957–2017, 1958–2017, ..., 2017}. We repeat this process for every technology. We perform hindcasting experiments on the empirical data for the estimation windows $m \in \{10, 20, 30, 40, 50\}$ and all possible forecasting horizons τ, resulting in more than 2.5 million point forecasts.[7]

Fig. 1.4 illustrates how forecast errors, measured as MSNE, decrease as a function of the estimation window m. Holding the prediction horizon τ fixed, up to $m = 40$ the forecast errors reduce as we use more data points for estimation. The fact that forecasting accuracy generally improves with longer estimation windows is a strong indication of substantial persistence of growth rates.

We next compare how well forecast errors are explained by our model. Since our hindcasting procedure makes use of rolling windows, it does not have independent forecast errors, and we cannot use Eqs. (1.5)-(1.7) directly. Thus, to properly test empirical forecasts against the null, we repeat the hindcasting exercise using surrogate data to generate the distribution of forecasts under the null. For each time series we fit the parameters of the model and then create the surrogate dataset by simulating the model under these parameters with the same number of time periods as the original dataset. We then repeat the hindcasting exercise but using the synthetic instead of the actual data. This allows us to compute the MSNE at each time horizon under the null. We repeat this procedure 100 times, which allows us to compute 95% quantiles of the synthetic MSNE. We then compare to the MSNE's of the actual forecasts, averaging over all forecasts with the same values of τ and m.

The results of this analysis are shown in Fig. 1.5, where the MSNE is depicted for different horizons τ and estimation windows m. The empirical forecast errors are substantially lower than what we would expect if the data followed an uncorrelated geometric random walk (blue ribbon). While still far from perfect, the theoretical prediction improves dramatically when allowing for autocorrelation (red ribbon). While the empirical forecast errors still rarely lie within the theoretical bounds, for $\tau < 10$ the agreement is quite good. This suggests that this simple time series model is close

[6]Note that m refers to the number of differenced data points, i.e. the number of patent growth rates, which is one less than the number of years in the estimation window.

[7]As shown in Appendix A.3, Fig. A.5, we can make more forecasts for small m and small τ than we can for large estimation and forecasting window combinations.

to being well-specified for $m < 10$, but does not capture the the data generating process as well for $m > 10$. Note, however, that the number of independent forecasts is much smaller when $m > 10$.

The left panel of Fig. 1.6 investigates biases in the random walk forecasts by plotting the mean normalized forecast error $\epsilon_{i,t+\tau}/\hat{\sigma}_i$ against different forecasting horizon τ and fixed $m = 30$ (see Appendix A.3, Fig. A.6, for other choices of m). Note that we are no longer squaring the errors. We no longer have closed form approximation for the error, but we can still test this with our surrogate data procedure. For relatively short forecasting horizons ($\tau \leq 11$), the empirical forecasts errors tend to be slightly higher than expected under the null, but are compatible with the null hypothesis until very long horizons $\tau > 35$. This can come from substantial non-stationarity in the data, but could also indicate a potential for better models.

We next pool all normalized forecast errors across technologies i and time horizons τ to test how well the errors compare with a Student t-distribution, as the theory would suggest in Eq. (1.7). The right panel of Fig. 1.6 shows a kernel density plot of normalized forecast errors, again for fixed estimation window $m = 30$, collapsing roughly 460 thousand data points into a single distribution. Eq. (1.7) suggests that the normalized forecasts errors, after appropriate rescaling using m, τ and θ_i, should follow a t-distribution with $m-1$ degrees of freedom. The black density curve indicates the theoretical t-distribution with degree $m-1 = 29$, which the errors should converge to if the model is well-specified. Due to the high degree of freedom this closely approximates a normal distribution. For simplicity, we have rescaled the forecast errors with a general autocorrelation parameter $\theta_i = \theta$ which we vary for better observing the parameter's effect. Clearly, when the errors are scaled without taking autocorrelation into account, i.e. the random walk case $\theta = 0$, the normalized forecast errors are not spread out enough to comply with the theoretical prediction (red line). When decreasing the autocorrelation to the average determined from Fig. 1.3, -36%, however, the fit improves substantially (green line). The fit is remarkably good if we use $\theta = -0.45$, which is somewhat more negative than the sample mean (blue line).[8]

[8]It is not unexpected that we need a lower autocorrelation parameter to better approximate the variance of the theoretical distribution, since the empirical forecasts are not fully independent due to the rolling window approach. Alternatively, we could have plotted the the forecast errors from the surrogate data instead of using Eq. (1.7), as done in Farmer & Lafond (2016) and Lafond et al. (2018).

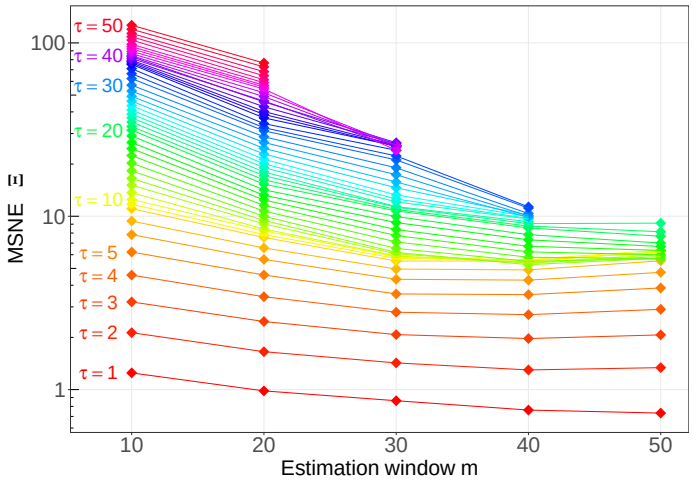

Figure 1.4: **Mean squared normalized forecast errors (MSNE) as a function of estimation window length m.** Color codes indicate different forecast horizons τ. Note the log-scale of the y-axis.

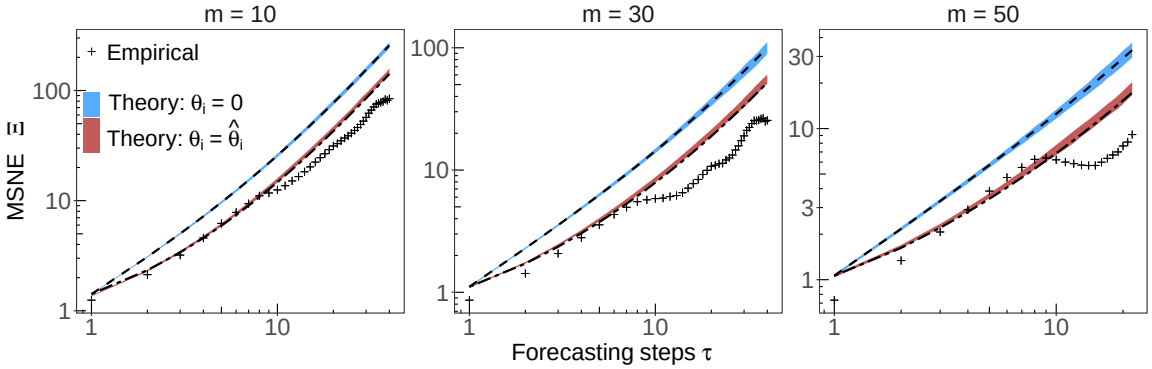

Figure 1.5: **Mean squared normalized forecast errors (MSNE) for the geometric random walk model.** The crosses represent the MSNE obtained from hindcasting the random walk model on the empirical data. Shaded ribbons indicate 5-95% quantiles of MSNE obtained from hindcasting the random walk model on surrogate data. The surrogate time series are generated by simulating from the geometric random walk with $\theta = 0$ (blue) and with $\theta = \hat{\theta}_i$ (red). The dashed line lying within the blue ribbon is the is the theoretical MSNE based on Eq. (1.5) with $\theta = 0$. The dot-dash line lying mostly within the red ribbon is the theoretical MSNE computed using $\theta = -0.36$. Note the log-log scale of the figure.

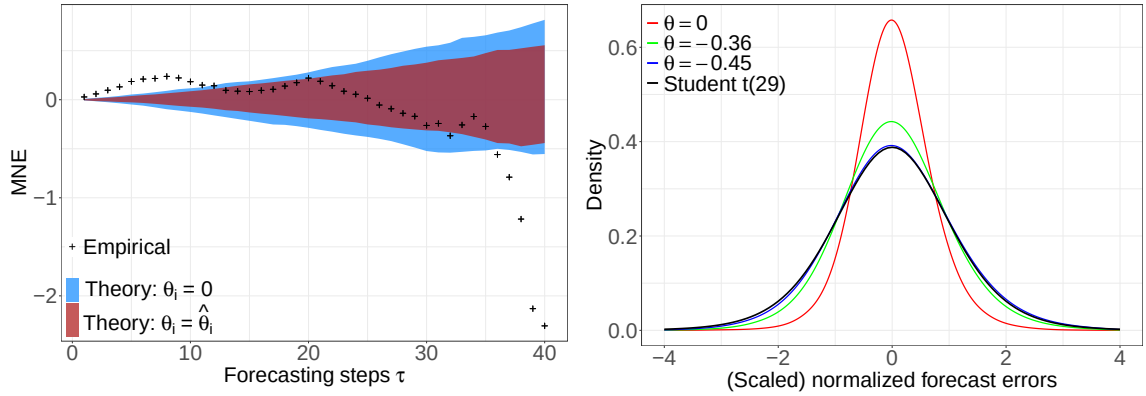

Figure 1.6: Left panel: **Mean normalized forecast errors of the random walk model.** In contrast to Fig. 1.5, the normalized errors are no longer squared, so this can take negative values. Coloring and symbols are the same as in Fig. 1.5. Right panel: **Distribution of the fully normalized forecast errors κ based on Eq.** (1.7)**, pooled across technology i and forecast horizon τ.** The black density curve indicates the theoretical t-distribution with degree $m-1=29$. The red, green and blue curves correspond to rescaled forecast errors following Eq. (1.7) and using different values of θ as indicated. Both panels are based on an estimation window of $m = 30$. Results for alternative choices of m are shown in Appendix A.3, Fig. A.6.

1.2.4 Predictive performance

We now turn to the predictive power of the time series models. Here, we don't compare the empirical model forecasts with their theoretical counterparts. Instead, we compare the forecasting results among different model specifications to determine which model yields the best predictions of the patenting time series. More specifically, we compare the predictive performance between the uncorrelated geometric random walk and the IMA(1,1). Since we have observed strong common trends (Fig. 1.2 and Appendix A.2), we also use pooled parameters for forecasting. Thus, additionally to the IMA(1,1) of Eq. (1.1), we make predictions based on the pooled parameter IMA(1,1) of the form

$$y_{i,t} = y_{i,t-1} + \mu + \theta\epsilon_{t-1} + \epsilon_{i,t}, \tag{1.8}$$

where we compute the pooled parameters simply as arithmetic averages of the individual estimates, $\hat{\mu} = \sum_{i=1}^{N} \hat{\mu}_i/N$ and $\hat{\theta} = \sum_{i=1}^{N} \hat{\theta}_i/N$ (we don't need the standard deviation of the noise $\hat{s}$ for forecasting). Analogously, for the pooled geometric random walk model we have

$$y_{i,t} = y_{i,t-1} + \mu + \epsilon_{i,t}. \tag{1.9}$$

We compare the forecasting performance of the different model specifications in Fig. 1.7. It becomes clear that for small estimation windows ($m = 10$) forecasts improve vastly when pooling the individual model parameters to common system parameters. This is particularly true for long-run forecasts. This result makes sense since individual parameter estimates are highly uncertain when the estimation window is short. In this case it is favorable to take advantage of the strong common trend by incorporating further cross-sectional information instead of projecting individual time series forward based on noisy parameters.

Even for estimation windows of medium length ($m = 30$), the pooled models still perform better in predicting future patenting dynamics. The differences, however, are less pronounced and essentially negligible for forecasts up to ten years. Interestingly, when using sufficiently long estimation windows ($m = 50$), this picture is slightly different. Here, there is essentially no difference in short term forecasting performance. For forecasts in the range of about 10 years ahead the autocorrelated model performs better than the uncorrelated random walk model. For long-term forecasts, however, we observe again that the pooled models predict better than the individual parameter models, although differences are smaller than compared to the short estimation windows.

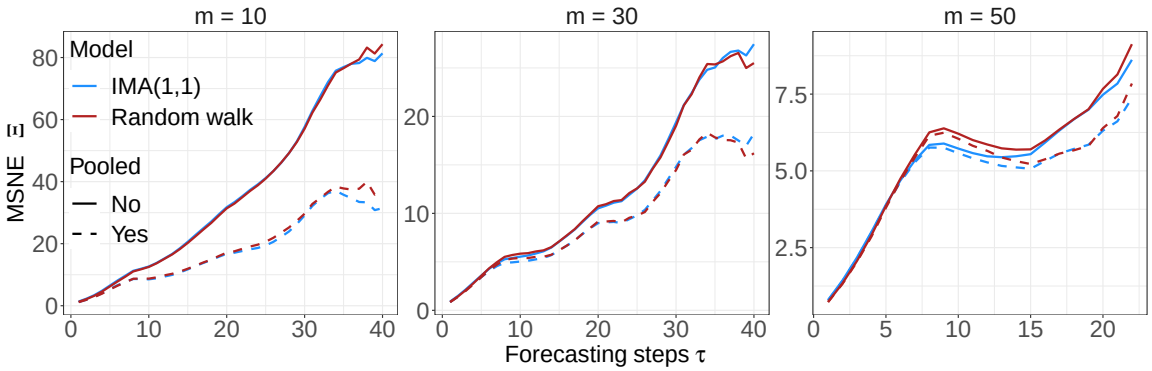

Figure 1.7: **Comparing mean squared normalized forecast errors (MSNE) between the various models.** Blue and red colored lines indicate IMA(1,1) and uncorrelated random walk forecast errors, respectively. Solid lines represent model specifications with technology-specific parameters, whereas dashed lines are cases with pooled parameters. For the (non-squared) mean normalized forecast errors consult Appendix A.3, Fig. A.6.

1.3 Discussion

Patents are an important measure of innovation effort. By analyzing the history of patents in distinct technological categories, we can obtain a quantitative description of innovation dynamics. We can use these insights to make forecasts of future patenting rates, giving us an informed guess on the extent of future inventive activities of different technological fields.

We have studied the entire record of utility patents from the US Patent and Trademark Office and we find that different technological classes exhibit very different growth dynamics. For example, average growth rates are substantially higher for Physics and Electricity patents, than they are for Textiles or Human Necessities. Despite technological idiosyncrasies, we also identify a strong common trend underlying the dynamics of virtually all patent categories.

We have further investigated how well the patenting data can be characterized by standard linear time series models. When comparing the different models based on their predictive performance, we find rather negligible differences among the geometric random walk (ARIMA (0,1,0)) and its autocorrelated extension (ARIMA (0,1,1)), although the latter model provides more empirically valid prediction intervals. When the estimation windows are short, resulting in noisy estimates of model parameters, we find that pooling individual model parameters into common "system parameters" can substantially improve forecasts, in particular for longer forecast horizons.

We have also compared forecasts of the geometric random walk with an alternative, more advanced network model suggested by Acemoglu et al. (2016). Interest-

ingly, the geometric random walk is able to drastically outperform the innovation network model in predicting patenting dynamics. We stress that these results are not disproving network effects in patenting dynamics, as is also discussed in Pichler, Lafond & Farmer (2020). Instead, this work proposes a simple benchmark for other models aimed at capturing patenting dynamics.

In this work we have focused on aggregate measures of forecasting performance, by averaging errors over technologies and years. An interesting avenue of future research would be to model "excess" growth rates of individual technologies, i.e. whether the cross-sectionally demeaned patenting growth rates can be explained and predicted with statistical models. This would shed further light on which technological categories are resistant to or dependent on the overall system dynamics. Future research should also be directed towards a better understanding of innovation dynamics on the more detailed technological level. It would be interesting to see whether general time series models, and further advancements of those, yield accurate results for modeling specific subgroups of technologies. This type of technology forecasting could prove useful for guiding policymakers in designing effective innovation policy schemes.

Chapter 2

Technological interdependencies predict innovation dynamics

This chapter is based on the paper Pichler, Lafond & Farmer (2020).

2.1 Introduction

Technological evolution is often described as a recursive process whereby the recombination of existing components leads to new or improved technological components (Schumpeter 1939, Usher 1954, Kauffman 1993, Fleming 2001, Arthur 2009, McNerney et al. 2011, Tria et al. 2014, Youn et al. 2015, Fink & Reeves 2019). A simple hypothesis, therefore, is that technologies that tend to recombine elements from fast-growing technologies should themselves grow faster. In other words, a technology will tend to progress faster if the technologies it relies on are themselves making fast progress. While these ideas are well established, very little has been done to establish empirically that technological interdependencies help predict future innovation dynamics. Being able to demonstrate this relationship would be very helpful, as it would allow us to support key technologies and overall technological progress by designing and supporting technological ecosystems.

In this chapter, we establish that knowing the technological ecosystem helps predict the dynamics of future innovation. We use a simple model where the innovation rates of a technological class depends on efforts within the class, but also on the stock of knowledge in the class and in supporting classes. We test the model on the record of United States Patent Office (USPTO) patents from 1836 to 2017, which includes around 10M patents, 40M classifications in 630 technological classes, and 90M citations. As predicted by the model, we find that the growth rates of the number of patents in a technological category depends strongly on the growth rates

of its knowledge sources. Given the volatile nature of growth rates, this strong relationship is remarkable. We use this insight for making out-of-sample predictions of patenting activities and find that integrating network effects can improve predictions substantially compared to independent time series models. The results have important policy implications, suggesting that research policy targeted at fostering innovation in a technological class has to take its surrounding technological ecosystem into account.

Several studies have put forward the idea of network-dependent innovation dynamics. For instance, Cowan & Jonard (2003) and König et al. (2011) have proposed models of innovation arising from the recombination of knowledge in R&D partnership networks.

Acemoglu et al. (2016) find that upstream patenting levels are highly correlated with downstream patenting levels. Taalbi (2020) finds broadly similar results using innovation counts and a network constructed to reflect which industry uses innovation from another industry. Patenting levels are persistent, and so are automatically very predictable. In contrast, we focus on changes in patenting levels, which are not persistent and are much more difficult to predict.

Our work on innovation dynamics is also related to evolutionary models such as Farmer et al. (1986), Bagley et al. (1992) and Jain & Krishna (2001). In these models, the exponential growth trajectories of nodes arise due to the existence of autocatalytic sets, i.e. a subset of the network where nodes have at least one positive incoming link from a node of the same subgraph, leading to sustainable self-reinforcing dynamics. Evidence for the presence of autocatalytic sets in technology systems was recently provided by Napolitano et al. (2018) who find that the autocatalytic structure of the patent network has grown in time and that patent classes belonging to the autocatalytic set show higher levels of innovation activity.

Our contribution to the literature can be summarized as follows. After discussing data sources in Chapter 2.2, we show strong empirical evidence of coupled innovation growth in the presence of network linkages and demonstrate that the observed effects are far from what would be expected by chance (Chapter 2.3). Second, we propose a simple theoretical model of network-dependent knowledge production which is able to explain the observed empirical pattern. We then show how the model can be estimated from empirical data, and quantify how network effects in innovation dynamics have become more important over the last 70 years (Chapter 2.4). Finally, we validate the model by predicting future patenting levels in Chapter 2.5. We find

that independent time series models can be substantially improved in case network information is available.

2.2 Data and network construction

We use data on the whole universe of granted United States patents starting in 1836 up to 2017, where the year refers to the publishing date of the patent. The dataset covers 9.83 million patents and over 91 million citations. We present some further details on data sources in Appendix B.1.

Each patent is categorized by the patent office into at least one of several Cooperative Patent Classification (CPC) codes. We regard each 4-digit CPC code as a distinct technological category. To check robustness of our results with respect to alternative classification schemes and aggregation levels, we redo the entire analysis presented here on the level of CPC 3-digits, International Patent Classification (IPC) 3- and 4-digits and report the results in Appendix B.

2.2.1 Temporal technology network

We construct a directed network $\mathbf{C}_t$ where an element $C_{ij,t}$ is the sum of all citations from technology j in year t to technology i. We first show for the static case how simple matrix algebra can be used to construct this technology networks from patent data. Let us define the patent citation matrix $\mathbf{H}$, where H_{pq} is one if patent p cites patent q and zero otherwise. To obtain the number of citations between technologies, the patent citation matrix, $\mathbf{H}$, has to be projected onto the technology space. The patents-technology class relationship can be represented as a binary bipartite network where the link $\tilde{B}_{pi}$ means that patent p belongs to technology class i.

Since we are interested in the knowledge flows between technologies, the focus of this work is the technology citation network which can be obtained as

$$\mathbf{C}^{\text{tech-citation}} = \tilde{\mathbf{B}}^{\top}\mathbf{H}\tilde{\mathbf{B}}. \tag{2.1}$$

This projection yields a integer-valued citation matrix, but inflates the total number of citations in the technology-based network if multiple technology classes are assigned to a patent. To keep total number of citations constant ($\sum_{ij} H_{ij} = \sum_{ij} C_{ij}$), we use the row normalized bipartite network in the projection. Note that the summing across rows $\sum_i \tilde{B}_{pi}$ yields the total number of CPC classes per patent p (the degree of p in the bipartite network) and the normalized entry $B_{pi} \equiv \tilde{B}_{pi} / \sum_i \tilde{B}_{pi}$ is the "share" of

technology i in patent p. Given the row-normalized bipartite network $\mathbf{B}$, we obtain the technology citation matrix considered in the analysis by

$$\mathbf{C} = \mathbf{B}^\top \mathbf{H} \mathbf{B}. \tag{2.2}$$

Fig. 2.1 illustrates the network projection in a schematic toy example.[1]

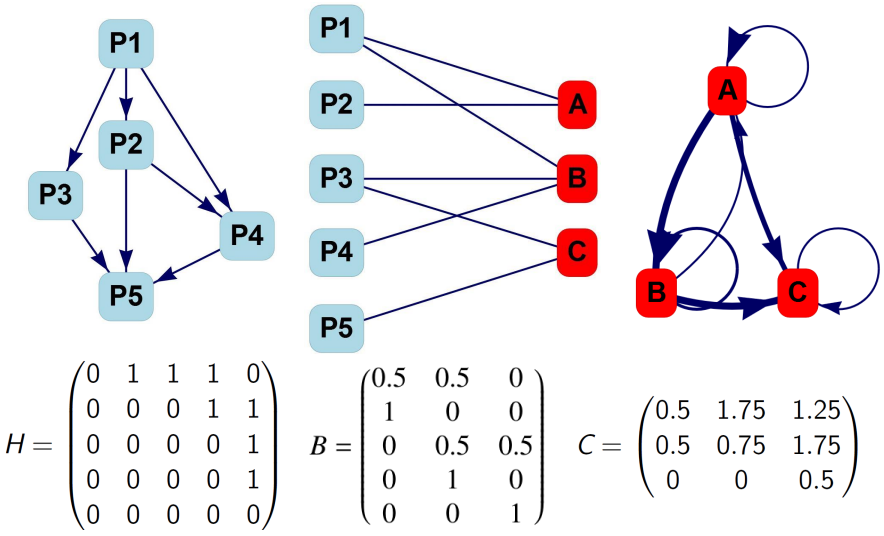

Figure 2.1: A toy example of the one-mode projection, Eq. (2.2). The left network is a patent citation network $\mathbf{H}$ where a directed edge from i to j indicates that patent i cites j. The network is hierarchical as older patents cannot cite more recent patents. Below the network the corresponding adjacency matrix is shown. In the middle is the row-normalized bipartite patent-classification network $\mathbf{B}$ with some patents (blue) assigned to multiple technology classes (red). The projection yields the weighted and directed technology citation network $\mathbf{C}$ with self-loops shown on the right.

To obtain a temporal representation of knowledge flows, we index the patent citation network by time, $\mathbf{H}_t$, where the time index refers to the *citing* patent. The time-indexed bipartite network $\mathbf{B}_t$ includes all the patents which cite and have been cited at time point t, as well as the corresponding technology classes. The time dependent technology citation matrix $\mathbf{C}_t$ is obtained by applying Eq. (2.2) to these matrices.

In our analysis, we use row-normalized matrices $\{\mathbf{W}_t : t = 1947, ..., 2017\}$ where each element $W_{ij,t}$ is defined as

$$W_{ij,t} \equiv \frac{C_{ij,t}}{\sum_{j=1} C_{ij,t}}. \tag{2.3}$$

[1]Note that technological distance can be defined in different ways, for example, with co-classification or co-citation networks. Our methodology is agnostic with respect to the exact definition but can be applied to any specification of matrix $\mathbf{C}$. For example, a co-classification matrix is obtained from substituting $\mathbf{H}$ in Eq. (2.2) by the identity matrix. Similarly, the co-citation matrix can be constructed by $(\tilde{\mathbf{B}}\mathbf{H})^\top \tilde{\mathbf{B}}\mathbf{H}$.

Thus, the row $\{W_{ij,t} : j = 1 \dots N\}$ can be interpreted as technology i's dependence in its patenting activity at time t on all other technologies: each entry is the share of citations from technology i at time t to another technological class j. This is a temporal network with yearly snapshots (Masuda & Lambiotte 2016). We do not construct networks for years before 1947, because earlier citations made by patents are not well documented. But citations *to* earlier patents are well reported. For example, a citation from a patent granted after 1947 to a patent of the 19$^{\text{th}}$ century is included in the network

2.2.2 Significance sampling

As argued in Alstott et al. (2017), links of technology networks are prone to be driven by "impinging factors" and should therefore be bias-corrected. Size effects can drive the number of citations between technologies, because the expected number of citations between two technological classes increases with the number of patents in the classes. Another important source of bias is the age of a patent. A new patent might have few forward citations because other inventions simply did not have the time to cite the patent. To eliminate the size and age biases, Alstott et al. (2017) suggest an algorithm to sample random network realizations conditional on class sizes and patent ages. The algorithm yields networks which preserve these impinging factors, but are random otherwise. The random realizations can then be used to derive a z-score for each link in the network. Links between technological categories which one would expect by chance, will have small z-scores and can be removed.

The randomized controls are generated as follows. First, identify all citations with a specific time lag which have been made in a particular year. An example would be to take all the citations made by patents in the year 2000 which cite patents three years earlier. Then, second, reshuffle all the *cited* patents. In the example, patents issued in 2000 would point randomly to the set of patents issued three years earlier. Larger classes will receive more citations on average. Note that the total number of citations for a given time lag are preserved by this procedure. By repeating the algorithm many times, an expected number of citations between classes can be derived, as well as the standard deviation. The link expectations and standard deviations are then used to compute the z-scores.

Sampling a large number of random networks for every year between 1947 and 2017 is computationally costly. Therefore, we analytically derive z-scores for every potential inter-class link as suggested by Alstott et al. (2017) who have shown that the analytical derivation approximates the numerical randomization process very well.

The randomization procedure can be stated as drawing from a hypergeometric distribution. Recall that a hypergeometric distribution is the probability of observing a certain number of successes in n draws, without replacement, from a sample of size s containing k objects with the success feature.

In our context, the hypergeometric distribution describes the probability of observing a certain number of citations from class i to class j at t with lag l. Let us denote the number of citations between two classes at t and given lag l with $C_{ij,t(t-l)}$. The sample size s is then simply the total number of inter-class citations, $s_{t(t-l)} = \sum_{i,j} C_{ij,t(t-l)}(1-\delta_{ij})$. The number of draws n corresponds to the number of citations made by a given class i. Thus, $n_{i,t-l} = \sum_{j\neq i} C_{ij,t(t-l)}$. The number of successes k is the number of citations received by j at $t-l$, i.e. $k_{j,t-l} = \sum_{i\neq j} C_{ij,t(t-l)}$.

The expected number of patents from i to j at t for all possible lags can then be found by

$$\mathbb{E}[C_{ij,t}] = \sum_{l=0}^{t-1836} \frac{k_{j,t-l}}{s_{t-l}} n_{i,t-l}, \tag{2.4}$$

and, by assuming independent random variables, the standard deviation is given by

$$\sigma_{ij,t} = \left(\sum_{l=0}^{t-1836} n_{i,t-l} \frac{k_{j,t-l}}{s_{t-l}} \left(1 - \frac{k_{j,t-l}}{s_{t-l}}\right) \frac{s_{t-l} - n_{i,t-l}}{s_{t-l} - 1} \right)^{1/2}. \tag{2.5}$$

For each entry in the citation matrix we calculate z-scores based on

$$z_{ij,t} = \frac{C_{ij,t} - \mathbb{E}[C_{ij,t}]}{\sigma_{ij,t}}. \tag{2.6}$$

We eliminate all citations in the technology citation matrix $\mathbf{C}_t$ with z-score smaller 2, roughly corresponding to the 97.5% quantile of the standard normal distribution. As we show in more detail in Appendix B.2, the significance sampling yields substantially sparser networks.

2.3 Preliminary empirical evidence

We now investigate whether technological categories connected to fast-growth technological categories also grow faster. In other words, we study the assortativity of the temporal network with respect to patenting growth rates. For each technology, we compute its neighborhood patenting growth rate as the average nearest neighbor growth rate (ANNG),

$$\tilde{g}_{i,t} = \sum_{j=1}^{N} \frac{C_{ij,t}(1-\delta_{ij})}{\sum_{k=1} C_{ik,t}(1-\delta_{ik})} g_{j,t}, \tag{2.7}$$

where δ_{ij} denotes the Kronecker delta (the matrix diagonal is excluded in the summation so that the ANNG measures neighbor growth rates only). We then test if there is a positive correlation between growth rates of patenting activity in technological classes and their ANNG.

Fig. 2.2a) plots the 10-year growth rates against the 10-year ANNG for the technology network in 2000, revealing a highly significant positive relationship in growth rates from 2000 to 2010 between technologies and their neighborhood. Since the technology network is based on the year 2000, the figure suggests that if a technology draws a significant share of knowledge from another technology in 2000, we would expect them to exhibit similar growth rates over the next ten years. This pronounced positive relationship is striking given that it is based on growth rates, which tend to be much noisier than patenting levels. While one could expect the network to be assortative with respect to degrees and patenting levels, assortativity with respect to growth rates is far less trivial.

Although interesting, we should test if this relationship holds over time and if we really can be sure that the observed relationship is significantly different from what a null model would produce. To benchmark our results, we calculate the correlations between growth rates and ANNG rates for a randomized version of the given technology network. For each year t and a fixed row i in the matrix $\mathbf{W}_t$, we resample all off-diagonal entries without replacements. In the randomized control network the nodes still have the same weighted outgoing links, but now randomly pointing to other nodes (excluding self-loops). We repeat the randomization process 1,000 times and report the average correlation between growth rates and ANNG as the corresponding null model. Fig. 2.2b) plots the Pearson correlation over time between growth rates and ANNG, once computed in the given technology network and once in the randomized control. We report the results for three different growth rates, 1-year (upper panel), 5-year (center panel) and 10-year growth rates (bottom panel). The positive correlation structure between knowledge sources growth and own growth rates over time is far from what one would expect from a network where nodes distribute their out-links randomly over the whole set of potential knowledge sources. Intuitively, one would expect that network effects materialize over the long run instead of showing immediate effects. The figure confirms this hypothesis, but even for 1-year growth rates we find relatively strong and significant positive correlations. Another interesting aspect is that for all three growth rate types, the positive relationship tends to get

stronger over time.[2] Further evidence on the impact of a technology's neighborhood on its patenting dynamics is discussed in Appendix B.

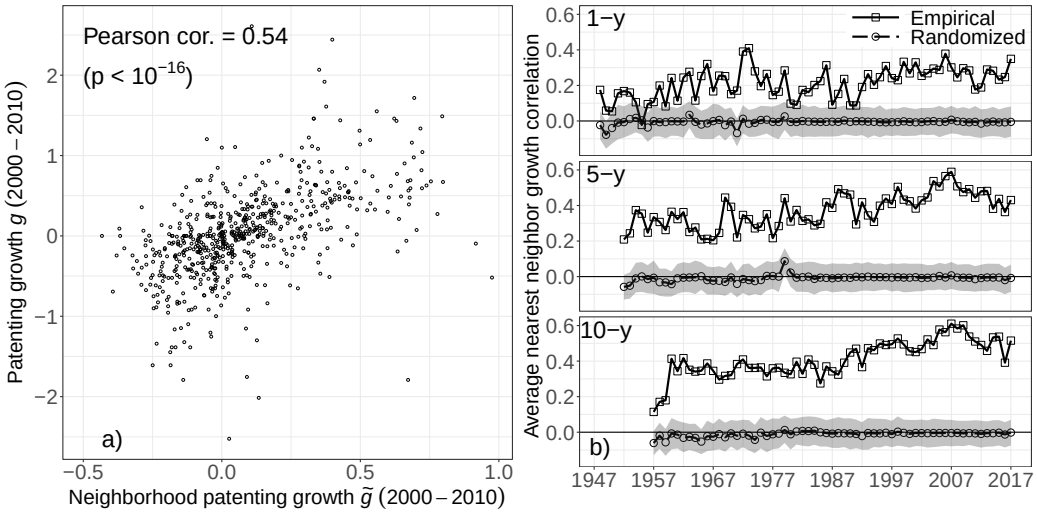

Figure 2.2: a) Technologies' 10-year growth rates $\boldsymbol{g}$ for the horizon 2000 to 2010 plotted against average nearest neighbor growth rates $\tilde{\boldsymbol{g}}$ (ANNG). The neighbors' growth rates are calculated based on the technology network in 2000. The plot thus suggests that technologies experienced a similar growth rate over the next ten year as the technologies they were linked to in 2000. b) Network assortativity over time with respect to 1-year (upper panel), 5-year (center panel) and 10-year (bottom panel) growth rates. There is pronounced correlation between growth rates of nodes and their neighbors in the technology network, while the randomized assortativity measure hovers around zero. The shaded area around the randomized network predictions depict the 0.05-0.95 quantile range of the 1,000 sample realizations. The correlations tend to get stronger in time in the empirical network.

2.4 A model of network-dependent knowledge creation

To explain the empirical observations, we now introduce a simple model of network-dependent knowledge creation and discuss its implications for the long-term evolution of technologies. Let us consider N distinct technological classes. The creation of new knowledge in a technological class requires active research effort, and depends positively on the existing stock of knowledge, from the same class or from specific other classes.

[2]It should be noted that technology citation networks are by construction correlated with co-classification networks, and thus, their marginal effects are difficult to disentangle. To test whether there is signal in the citations themselves, we computed the same ANNG measures for USPC main classes. When considering only USPC main classes, there is no co-classification, allowing us to separate these effects. Although slightly noisier, we also find strongly positive ANNG values for the less granular USPC main classes.

More precisely, we assume that the creation of new knowledge follows the dynamical system

$$\dot{K}_i(t) = \theta_i R_i(t)^\alpha \prod_{j=1}^{N} K_j(t)^{\beta W_{ij}}, \tag{2.8}$$

where θ_i is a technology-specific productivity parameter, $R_i(t)$ the research effort in category i at time t and $K_i(t)$ is the stock of knowledge in i at t. The technological ecosystem is represented by the weighted adjacency matrix $\mathbf{W}$ which is normalized to be row-stochastic and for mathematical convenience assumed to be fixed in time. A technological class is then a node in the network whose innovation rates depend on its location in the technological ecosystem. If a technology i draws knowledge from the knowledge stock of node j in the innovation process, the two nodes are connected through a directed edge from i to j. The directed link has the weight W_{ij} denoting the share of technology j's knowledge in the creation of new knowledge in technology i. θ_i captures the fact that intrinsic characteristics of technologies affect how easy innovation rates can be influenced. $\alpha \geq 0$ denotes the elasticity of knowledge output with respect to research efforts. The elasticity of knowledge output in i ($\dot{K}_i$) with respect to the knowledge stock in j (K_j) is βW_{ij}, thus β denotes the sum over all classes j of these elasticities (since $\sum_j W_{ij} = 1$).

Eq. (2.8) also relates to the knowledge production functions used in classical endogenous growth models (Romer (1990), Aghion & Howitt (1990), Grossman & Helpman (1991) and Jones (1995)), but incorporates network effects. It simplifies to the standard Cobb-Douglas knowledge production function in case each technology uses only its own knowledge stock ($\mathbf{W} = \mathbb{I}$). Similar equations have been estimated empirically within the "R&D spillovers" literature which estimates the impact of R&D in one entity on outcomes in another entity, such as countries, regions, firms, and sectors (Ertur & Koch 2007, Hall et al. 2010, Ho et al. 2018).

Research effort is considered to be an exogenous policy variable which we assume to grow at a constant rate $R_i(t) = R_{i,0}e^{\lambda_i t}$. Dividing Eq. (2.8) by $K_i(t)$, taking logs and the derivative with respect to time, we obtain after rearranging the nonlinear autonomous system

$$\dot{g}_i(t) = \alpha\lambda_i g_i(t) + (\beta W_{ii} - 1)g_i(t)^2 + \beta g_i(t) \sum_{j=1}^{N} W_{ij} g_j(t)(1 - \delta_{ij}), \tag{2.9}$$

where $g_i(t) \equiv \dot{K}_i(t)/K_i(t)$ is the growth rate of the knowledge stock. Eq. (2.9) has been extensively studied without network effects in traditional endogenous growth models and in this case its dynamics are well-understood (Romer (2012), ch.3). When

solving the model without network effects, the steady state growth rate ("balanced growth path") is $g_i^* = \alpha\lambda_i/(1-\beta)$ which is globally stable under the standard assumption of $\beta < 1$. Here, the long run-growth rate of knowledge in technology i can only be increased by increasing the long-run growth rates of research efforts in the particular technology.

When network effects are included, a technology's long-term growth path depends also on the growth rates of other technologies. By setting $\dot{g}_i = 0$ for all i, we find the steady state of the form

$$g_i^* = \frac{\alpha\lambda_i}{1-\beta W_{ii}} + \frac{\beta}{1-\beta W_{ii}} \sum_{j=1}^{N} W_{ij} g_j^* (1-\delta_{ij}). \tag{2.10}$$

The steady state growth path for technology i can be understood as a sum of two components. The first part is the idiosyncratic term, which equals the endogenous growth result without network effects. The second component suggests that the long-run growth of a technology depends positively on its neighbors' growth rates.

To see the difference between the network-dependent model and a simple endogenous growth model version without network effects, let us assume that we could increase the constant growth rate of research effort in a technology i from λ_i to λ_i'. Without network effects, this would simply increase the growth rate of i by $\alpha(\lambda_i' - \lambda_i)/(1-\beta)$ with no impact on other technologies. Yet if network effects are included, this initial increase of i's growth rate will also impact neighboring technologies which draw upon the knowledge stock of i, which in turn, will again affect their downstream neighbors and so on. If node i points to any of the affected technologies, it will again experience a change in its growth rate and trigger the process again. We see that the convergence (if any) to the steady state after a shock in the research effort variable is more involved if the innovation rate of a technology depends on other technologies. The phase portraits depicted in Fig. 2.3 visualize how a change in research efforts impact innovation dynamics in a simple system of only two technologies, one with network effects (right panel) and one without (left panel).

2.4.1 Calibration

While the dynamical system can have a non-trivial transient, for simplicity we calibrate the model based on the steady state results, using maximum likelihood. To do so, we reformulate Eq. (2.10) as the econometric model

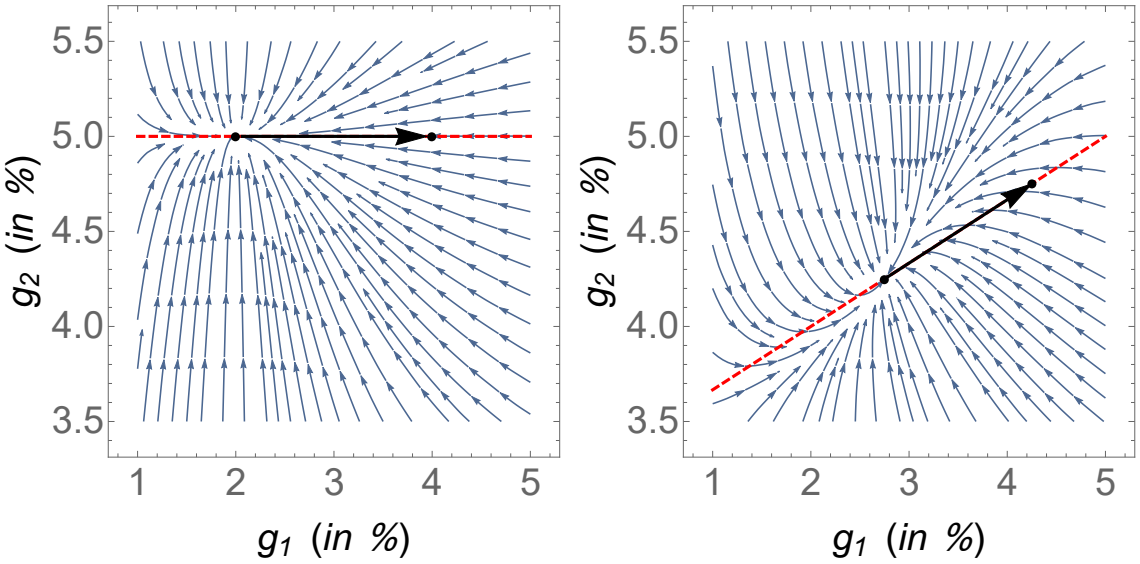

Figure 2.3: Phase portraits of growth trajectories for a simple example of only two technologies. The horizontal and vertical axes denote the growth rates of technology 1 and technology 2, respectively. In the left panel there are no network effects, $\mathbf{W} = \mathbb{I}$. In the right panel a network is present with elements $W_{ij} = 0.5$ for all i, j. The phase diagrams are shown for initial research growth rates $\lambda_1 = 0.02$ and $\lambda_2 = 0.05$. The remaining parameters are set as follows: $\alpha = \beta = 0.5$. With this parametrization all trajectories converge to the steady state $(g_1^*, g_2^*) = (0.02, 0.05)$ in the no-network case and to the steady state $(g_1^*, g_2^*) = (0.0275, 0.0425)$ if network effects are included. In the absence of network effects, the steady state growth rates are simply equal the research effort growth rates. If network effects are included, technology 1 is growing faster in steady state since it draws knowledge from technology 2 which experiences higher research effort growth. In contrast, for technology 2 the steady state growth rate is lower due to the network effects. The black arrows indicate how the steady state of the system moves if research effort growth λ_1 is doubled to $\lambda_1' = 0.04$, but leaving $\lambda_2 = 0.05$ unchanged. In the left panel, this increases only the steady state growth rate for technology 2, but leaves technology 2 unaffected. In the right panel where network effects are included both technologies benefit from the increase in research effort in technology 1. The red dashed line shows the trajectory of the steady state as a function of λ_1.

$$g_{i,t} = \frac{a_i}{1 - \beta W_{ii,t}} + \frac{\beta}{1 - \beta W_{ii,t}} \sum_{j=1}^{N} W_{ij,t} g_{j,t} (1 - \delta_{ij}) + \epsilon_{i,t}, \tag{2.11}$$

where $\epsilon_{i,t} \sim N(0, \sigma^2)$ and a_i is capturing the composite variable $\alpha\lambda_i$.

We estimate the parameters using maximum likelihood since the OLS estimator is biased and inconsistent due to the reasons outlined in Anselin (2013). The specification is related to spatial autoregressive (SAR) models for panel data (Elhorst 2014, Wang & Yu 2015), but there are also significant differences. In contrast to most spatial econometric applications, our model uses a time-varying network term and includes diagonal elements which complicate the estimation. We restate the econometric model in matrix form which will prove useful in the derivation of the estimators.

The model and its corresponding data generating process can be written as

$$\boldsymbol{g}_t = \mathbf{V}_t \boldsymbol{a} + \beta \mathbf{V}_t \tilde{\boldsymbol{g}}_t + \boldsymbol{\epsilon}_t, \quad t \in \{1, 2, ..., T\}, \tag{2.12}$$

$$\boldsymbol{g}_t = \left[\mathbb{I} - \beta \mathbf{W}_t\right]^{-1} \boldsymbol{a} + \left[\mathbb{I} - \beta \mathbf{W}_t\right]^{-1} \boldsymbol{\epsilon}_t, \tag{2.13}$$

$$\boldsymbol{\epsilon}_t \sim N(0, \sigma^2 \mathbb{I}),$$

where the matrix $\mathbf{V}_t = \mathbf{V}_t(\beta) \in \mathbb{R}^{N \times N}$ is diagonal with elements $V_{ii,t} \equiv (1 - \beta W_{ii,t})^{-1}$. $\boldsymbol{g}_t$, $\boldsymbol{a}$, $\tilde{\boldsymbol{g}}_t$ and $\boldsymbol{\epsilon}_t$ are column vectors of length N. The i^{th} element of $\tilde{\boldsymbol{g}}_t$ is defined as $\tilde{g}_{i,t} := \sum_{j=1}^{N} W_{ij,t} g_{j,t} (1 - \delta_{ij})$.

The likelihood function is given by

$$\begin{aligned} &\text{LogL}\left((\sigma^2, \boldsymbol{a}, \beta) | \{\boldsymbol{g}_t\}\right) = \\ &-\frac{NT}{2} \ln(2\pi\sigma^2) + \sum_{t=1}^{T} \ln \left|\mathbb{I} - \beta \mathbf{V}_t \mathbf{W}_t\right| - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \boldsymbol{\epsilon}_t^\top \boldsymbol{\epsilon}_t. \end{aligned} \tag{2.14}$$

To find the network specific parameter β, we first optimize the log-likelihood function with respect to σ^2 and a_i. Following the usual procedure in spatial autoregressive models, we then numerically optimize the concentrated log-likelihood function with respect to β (Ord 1975, Elhorst 2014, Wang & Yu 2015). Maximizing the log-likelihood function with respect to σ^2 and a_i yields the following expressions:

$$\begin{aligned} \hat{\sigma}^2 &= \frac{\sum_t \boldsymbol{\epsilon}_t^\top \boldsymbol{\epsilon}_t}{NT}, \\ \hat{\boldsymbol{a}} &= \left(\sum_t \mathbf{V}_t^2\right)^{-1} \sum_t \mathbf{V}_t \boldsymbol{g}_t - \beta \left(\sum_t \mathbf{V}_t^2\right)^{-1} \sum_t \mathbf{V}_t^2 \tilde{\boldsymbol{g}}_t. \end{aligned}$$

By plugging the found expressions into Eq. (2.14), we find parameter β through the following optimization problem

$$\underset{\beta}{\text{argmax}} \quad constant + \sum_{t=1}^{T} \ln \left|\mathbb{I} - \beta \mathbf{V}_t \mathbf{W}_t\right| - \frac{NT}{2} \ln \sum_{t=1}^{T} \tilde{\boldsymbol{\epsilon}}_t^\top \tilde{\boldsymbol{\epsilon}}_t, \tag{2.15}$$

with $\tilde{\boldsymbol{\epsilon}}_t \equiv \boldsymbol{g}_t - \mathbf{V}_t \left(\sum_t \mathbf{V}_t^2\right)^{-1} \sum_t \mathbf{V}_t \boldsymbol{g}_t - \beta \mathbf{V}_t \left(\sum_t \mathbf{V}_t^2\right)^{-1} \sum_t \mathbf{V}_t^2 \tilde{\boldsymbol{g}}_t$. The (asymptotic) variance-covariance matrix is obtained as the Fisher information matrix, i.e. by inverting the numerically computed Hessian of the negative log-likelihood function.

We allow the technology network to be time-varying, since we observe changing citation networks over time. In Appendix B.5 we present results of extensive robustness checks, covering alternative model specifications such as using time-fixed networks and spatial autoregressive models. There we also control for further possible explanatory variables, such as technological class sizes.

	Dependent variable: patenting growth		
	(1-y)	(5-y)	(10-y)
average $\hat{a}_i$	0.01	0.02	0.04
$\hat{\beta}$	0.86^{***}	0.94^{***}	0.95^{***}
	(0.0059)	(0.0057)	(0.0070)
$\hat{\sigma}^2$	0.10	0.19	0.29
Log-likelihood	-12,956	-5,333	-3,526
Observations	42,403	8,482	4,223
Note:			$^{***}\text{p}< 10^{-16}$

Table 2.1: Results from estimating the econometric network model Eq. (2.11). The results are shown for 1-, 5- and 10-year growth rates.

The model is derived in terms of knowledge stocks $K_{i,t} = \sum_{\tau=0}^{t} P_{i,\tau}$, but knowledge stocks grow very smoothly, leaving little variance to exploit for estimating the parameters of interest. In the steady state, however, by definition growth rates are constant over time within each category. In that case, the (deterministic) growth rates of stocks and flows are the same. Thus, for empirical convenience, we let $g_{i,t} \equiv \ln(P_{i,t}/P_{i,t-\tau})$ be the τ-year patenting growth rate of the technology class i at a time t.

We fit the model to the whole time series up to 2017 and report the estimated parameters in Table 2.1. Since the results of Chapter 2.3 suggest varying magnitudes of network effects for different time lags, we estimate the parameters for 1-year, 5-year and 10-year growth rates (columns one to three, respectively). Unsurprisingly, we find research growth parameters $a_i = \alpha\lambda_i$ which are positive on average and higher for larger time lags. The highly significant network parameter β is large, ranging from 0.86 to 0.94, depending on the growth rate lag. This value is large, because if $\beta > 1$, we would expect exploding dynamics where an increasing knowledge stock leads to ever larger growth rates. β is smaller, but close to one, exemplifying the importance of the existing technology network for future innovation dynamics.

To explore the network impact on innovation growth more systematically, we rewrite the derived result in Eq. (2.10) into matrix notation,

$$\boldsymbol{g}^* = \alpha \mathbf{L} \boldsymbol{\lambda}, \tag{2.16}$$

where $\mathbf{L} \equiv [\mathbb{I} - \beta \mathbf{W}]^{-1}$. The matrix representation is useful as it shows how the long-run trajectory of innovation in a given technological class depends on the research

efforts in the entire technological ecosystem. In particular, if research effort is subject to policy, we can use Eq. (2.16) to study two interesting policy experiments.

First, Eq. (2.16) allows us to identify the key supporting technologies for each technology. The naïve way to foster innovation in a technology is to increase research efforts in the particular technological category. But Eq. (2.16) shows that knowledge growth depends also on research in other technologies and that there could even be cases where knowledge spillovers from research in other technologies will be more beneficial than simply devoting additional research efforts in the focal technology's class.[3] Knowledge spillovers from technology j to technology i are made explicit when looking at the Jacobian matrix

$$\frac{\partial g_i^*}{\partial \lambda_j} = \alpha L_{ij}. \tag{2.17}$$

The matrix element L_{ij} therefore informs us on how much we would expect the long-term innovation growth rate of technology i to change as a consequence of a marginal change in technology j. A large row sum of off-diagonal elements $\sum_{j=1} L_{ij}(1-\delta_{ij})$ indicates a large dependence of technology i's growth on external research activities. Ordering technologies based on their size in a given column yields a ranking from the most to least important supporting technologies for the focal technology.

As a second policy experiment, we can ask how a sustained change in R&D in a particular sector affects the long-run growth rate of the entire technological ecosystem. This can be relevant for a policy-maker who wants to devote her research investments in an efficient manner. For example, if a choice has to be made in what technology to invest for fostering overall innovation activities, the decision can be supported with Eq. (2.16). If an economy's total knowledge is the sum of sectoral knowledge stocks $K_{tot}(t) = \sum_{i=1}^{N} K_i(t)$, the growth rate of the total knowledge stock can be expressed as $g_{tot}(t) = \sum_{i=1}^{N} g_i(t) \frac{K_i(t)}{K_{tot}(t)}$. The impact on total knowledge growth as a consequence of a sustained change in research effort growth in a single class i is then given by

$$\frac{\partial g_{tot}^*}{\partial \lambda_j} = \alpha \sum_{i=1}^{N} L_{ij} \frac{K_i}{K} + \alpha \lambda_j \sum_{i=1}^{N} \frac{\partial (K_i/K)}{\partial \lambda_j} L_{ij}. \tag{2.18}$$

The impact on overall innovation growth is again a sum of two components. The first part in the sum takes the form of a output multiplier, analogous to the output

[3]When estimating the model based on 1-year growth rates, we find five technologies with larger off-diagonal entries than diagonal elements. Although these technologies are very different, ranging from *multipurpose hand tools* (B25F) to *emergency protective circuit arrangements* (H02H), they all rely heavily on the general purpose technology class A61B, *diagnosis, surgery and identification*, which exhibits the largest downstream spillovers, $\sum_j L_{ij}(1-\delta_{ij})$, of all technologies.

multiplier in input-output economics. In contrast to traditional input-output economics (Leontief 1936, Miller & Blair 2009), however, the output multiplier has to be weighted with the relative category sizes since the model is in growth rates instead of levels. Increasing research efforts in a category with a high weighted output multiplier entails large positive effects on overall innovation rates. The second part in the sum accounts for the fact that the weights themselves change due to changes in research effort.

2.5 Predicting innovation rates

The empirical evidence for network effects on innovation dynamics is strong. In this chapter we take advantage of this finding to test the predictive power of the estimated network model. We use the model to make out-of-sample forecasts of patenting levels and evaluate their performance by comparing them with predictions based on time series models which do not incorporate network effects explicitly.

We test the network model in two separate prediction exercises. In the first prediction exercise we ask whether knowledge on innovation dynamics around a focal technology can help predicting its growth rates. Or to put it differently, conditional on growth rates of i's neighbors, what is i's growth rate? We call these the *conditional* forecasts. In a second prediction exercise we make forecasts where we do not assume knowledge on the neighbors' growth rates, and use only information available up to t. We call these the *unconditional* forecasts.

To evaluate the forecast performance of the network model, we compare its predictions with the forecasts obtained from standard ARIMA(p,1,q) time series models

$$g_{i,t} = \nu_i + \sum_{s=0}^{p} \phi_{i,s} g_{i,t-s} + \sum_{s=0}^{q} \psi_{i,s} u_{i,t-s} + u_{i,t}, \tag{2.19}$$

where $\phi_{i,0} = \psi_{i,0} = 0$. Note that an ARIMA simply reduces to a geometric random walk if $p = q = 0$ and that predictions from the ARIMA model are the same for both prediction exercises, since here forecasts for a specific class i are completely independent of other classes.

The conditional predictions of the network model can be written as

$$\hat{g}^{cond.}_{i,t+\tau} \equiv \frac{\hat{a}_i}{1 - \hat{\beta} W_{ii,t}} + \frac{\hat{\beta}}{1 - \hat{\beta} W_{ii,t}} \sum_{j=1}^{N} W_{ij,t} g_{j,t+\tau} (1 - \delta_{ij}). \tag{2.20}$$

For the unconditional forecasts we use the specification

$$\hat{\boldsymbol{g}}_{t+\tau}^{uncond.} \equiv \left(\sum_{k=0}^{k'} \hat{\beta}^k \mathbf{W}_t^k \right) \hat{\boldsymbol{a}} \approx \mathbf{L}(\hat{\beta}) \hat{\boldsymbol{a}}, \tag{2.21}$$

where the approximation is exact in case $k' \to \infty$. We choose this specification since we observe better predictions for $k' \leq \infty$.

To see why the network model yields better prediction with a $k' \leq \infty$, note that the data generating process in Eq. (2.13) for given parameters implies $\mathbb{E}[\boldsymbol{g}_t] = (\mathbb{I} - \beta \mathbf{W}_t)^{-1} \boldsymbol{a}$ and $\mathbb{V}[\boldsymbol{g}_t] = \sigma^2 \left[(\mathbb{I} - \beta \mathbf{W}_t^\top)(\mathbb{I} - \beta \mathbf{W}_t) \right]^{-1}$, entailing an exploding variance for β approaching one. Since our estimations suggest that β is actually close to one, forecast errors are expected to be large. To alleviate this issue, we take advantage of the power series expansion, $(\mathbb{I} - \beta \mathbf{W}_t)^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{W}_t^k$, by defining the predictor as $\hat{\boldsymbol{g}}_t^{uncond.} = (\sum_{k=0}^{k'} \hat{\beta}^k \mathbf{W}_t^k) \hat{\boldsymbol{a}}$. Increasing k' reduces the bias, versus reducing k' lowers the variance.

To find the best choice of k' without using future information into account, we split the data set into three parts: a training set (1948-1987), a validation set (1988-2002) and a test set (2003-2017). We first calibrate the model to the training set to predict patenting in every class in the validation set for varying k'. We then choose the model which yield the lowest median absolute percentage error in the validation set forecasts. We use median absolute percentage errors over mean absolute percentage errors, since the well-known downward bias of the mean absolute percentage error (Armstrong 1985) tends to favor predictors which are systematically underestimating. This selection procedure yields $k' = 3$. After choosing the best model specification, we re-estimate the model to the whole data set up to 2002. Using the parametrization of the "optimal" models, we then predict patenting activities for every single year and technology from 2003 to 2017. Since we found different magnitudes of network dependence for different growth rate lags, we estimated the network model for every growth rate lag (1-year, 2-years, ..., 15-years) separately.

This approach assumes that the optimal choice in the validation set also corresponds to the best choice in the test set which does not always have to be the case. The dependence between variables can change over time, an aspect which is not taken into account here. Alternative approaches could be considered in future research.

We follow the same approach for the ARIMA predictions to determine the "optimal" number of MA and AR terms. Following the proposed principles, we first estimate all possible ARIMA models with lags $\{(p,q) |\ p,q \in \{0,...,5\}\}$ for a given

technology in the training set. Second, we use all estimated configurations to predict patenting levels in the validation set. The model specification which reduces the median absolute percentage error in the validation set is then chosen to predict patenting activities in the test set. This procedure is repeated for every single time series.

We estimate all models based on the data between 1947 and 2002 and use the result to predict yearly patenting from 2003 to 2017. To assess the predictive performance of the network models, we calculate the predictability gain for each year and technology as

$$PG_{i,t} = \frac{|P_{i,t} - \hat{P}_{i,t}^{ARIMA}|}{P_{i,t}} - \frac{|P_{i,t} - \hat{P}_{i,t}^{network}|}{P_{i,t}}, \tag{2.22}$$

where $\hat{P}_{i,t}^{ARIMA}$ denotes the predicted number of patents from the ARIMA model and $\hat{P}_{i,t}^{network}$ the predictions from the network models. Eq. (2.22) is simply the difference between absolute percentage errors from the ARIMA and the network model predictions.

The average predictability gains for each year are shown in Fig. 2.4, demonstrating substantial predictive power of the network model. In the conditional forecasts, the predictability gain is relatively small for one year growth rates ($\approx 3\%$), but reaches a maximum of roughly 62% in 2009. Note that if a technology's evolution is not influenced by its surrounding technological ecosystem, we would not expect any systematic predictability gain at all. Obviously, it is harder to beat the ARIMA models in the unconditional forecasting scenario where no information on future innovation dynamics of neighboring technologies is available. Nevertheless, the network model performs significantly better every single year, with an average predictability gain of around 20%. In Appendix B.6 we investigate further aspects of the predictability gain distribution and present results for alternative forecasting benchmarks.

When taking time averages of Eq. (2.22) for each time series, we find for 85% of technologies positive mean predictability gains in the conditional forecasts. In the unconditional forecasts, we get positive predictability gains in 63%. The result looks similar when taking time medians instead of averages where we find positive predictability gains for 84% of all conditional forecasts and for 63% of unconditional forecasts. For each technology, we also conduct a one-sided t-test to check if the predictability gains are significantly greater than zero. We find that predictability gains are significantly larger than zero (on the 5% level: note that we have only 15 observations per series) in 68% of all cases for the conditional forecasts and in 40%

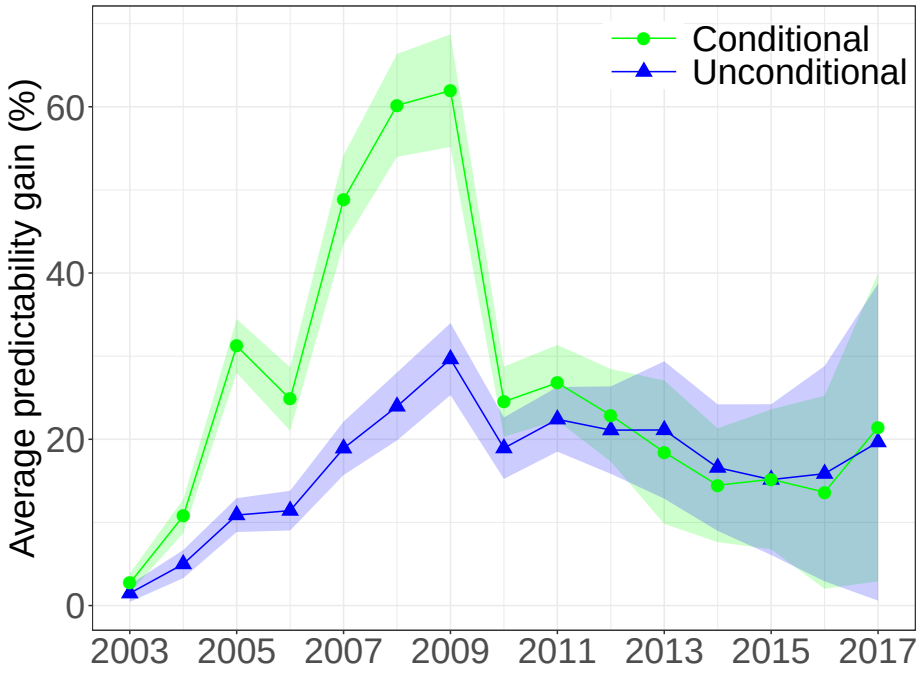

Figure 2.4: Average predictability gain in %, $\sum_i PG_{it}/N$, over ARIMA time series models for the conditional (Eq. (2.20)) and unconditional forecasts (Eq. (2.21)). Results are based on forecasts made in 2002 for each year in the future. Note that conditional forecasts use future information on neighborhood growth rates. Shaded areas indicate twice the standard errors.

for unconditional forecasts. On the other hand, only 6% of all time series are found to have significantly negative predictability gains in the conditional forecasting setup and 16% in the unconditional one.

The results of the prediction exercises add further support to the analysis of the previous chapters. Innovation rates are network-dependent, and having knowledge on the technology network helps improving forecasts of patenting dynamics.

2.6 Discussion

We have provided evidence that technological classes co-evolve if they are linked through the patent citation network. This empirical observation can be explained with a simple model of network-dependent knowledge creation. Motivated by the recursive nature of the innovation process, technologies take advantage of distinct knowledge sources in the technological ecosystem. We also validated the model by making out-of-sample prediction of growth rates, conditional and unconditional of neighboring growth rates. We have shown that the network model improves forecasts of patenting growth rates substantially when compared to standard time series mod-

els. Thus, the network of technological interdependencies is informative for inferring future patenting dynamics.

It is also important to point out the limitations and caveats of our analysis. By keeping the number of nodes fixed in the technological network, we ignored the emergence of breakthrough technologies. In our approach, technological categories are taken as given and well-defined by the latest technology classification system. But technological classifications themselves are subject to evolutionary forces emerging from the changing technological base (Lafond & Kim 2019). New technology classes are created, old ones merge or are abandoned and innovations can be reclassified – mechanisms which were not made explicit in this work.

In our empirical framework, we allowed the technology network to vary in time, but we did not model this explicitly. An interesting avenue for future research is to gain a better understanding of the mechanisms driving these rewiring processes. Furthermore, it would be important to understand how patenting activity and technology network metrics such as centrality translate into cost reductions of technologies (Farmer & Lafond 2016, Triulzi et al. 2020). Finally, another interesting direction of research is to further investigate the interaction of network-dependent innovation dynamics with the real economy by coupling innovation network with input-output networks (Hötte 2021). This could illuminate new aspects of how structural change happens in the economy and improve economic forecasts.

We conclude by discussing the policy implications of our analysis. Since innovation dynamics in technological classes reveal strong network dependence, the allocation of research resources has to consider the ecosystem in which the focal category is embedded in. Innovation in a technology is not an isolated process, but results from complex mechanisms involving research efforts, institutional settings, and technological interdependencies. Facing the enormous challenge of transitioning the current economic system to a low-carbon economy, we will need substantial improvements in key low-carbon technologies such as photovoltaics and wind energy. Better understanding their technological interdependencies emerging from the technological ecosystem and making it fruitful for research policy will be a crucial step in this direction.

Chapter 3

In and out of lockdown: Propagation of supply and demand shocks in a dynamic input-output model

This chapter is based on Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2020) and Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021).

3.1 Introduction

The social distancing measures imposed to combat the Covid-19 pandemic created severe disruptions to economic output, causing shocks that were highly industry specific. Some industries were shut down almost entirely by lack of demand, labor shortages restricted others, and many were largely unaffected. Meanwhile, feedback effects amplified the initial shocks. The lack of demand for final goods such as restaurants or transportation propagated upstream, reducing demand for the intermediate goods that supply these industries. Supply constraints due to a lack of labor under social distancing propagated downstream, creating input scarcity that sometimes limited production even in cases where the availability of labor and demand would not have been an issue. The resulting supply and demand constraints interacted to create bottlenecks in production, which in turn led to unemployment, eventually decreasing consumption and causing additional amplification of shocks that further decreased final demand.

In this chapter we develop a model that respects the three main features that made the Covid-19 episode exceptional: (1) The shocks were highly heterogeneous

across industries, making it necessary to model the economy at the sectoral level, taking sectoral inter-dependencies into account; (2) the shocks affected both supply and demand simultaneously, making it necessary to consider both upstream and downstream propagation; (3) the shocks were so strong and were imposed and relaxed so quickly that the economy never had time to converge to a new steady state, making dynamic models better suited than static models.

In a first version of this work, results were released online on 21 May 2020, not long after social distancing measures first began to take effect in March (Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2020). Based on data from 2019 and predictions of the shocks by del Rio-Chanona et al. (2020), we predicted a 21.5% contraction of GDP in the UK economy in the second quarter of 2020 with respect to the last quarter of 2019. This forecast was remarkably close to the actual contraction of 22.1% estimated by the UK Office of National Statistics. (The median forecast by several institutions and financial firms was 16.6% and the forecast by the Bank of England was 30%).[1]

In a substantially revised version (Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2021) we took advantage of the benefits of hindsight to perform a "post-mortem", allowing us to better understand what worked well, what did not work well, and why. We systematically investigated the sensitivity of the results under different specifications, including variations in the shock scenario, production function, and key parameters of the model. We showed how getting good aggregate results depends on making the right trade-off between the severity of the shocks and the rigidity of the production function, as well as other key factors such as the right level of inventories. This provides valuable lessons about modeling the economic effects of disasters such as the Covid-19 pandemic.

In this chapter we synthesize our insights from our earlier versions but with a special focus on the model and theoretical results. Our model is inspired by previous work on the economic response to natural disasters (Hallegatte 2008, Henriet et al. 2012, Inoue & Todo 2019). As in these models, industry demand and production decisions are based on simple rules of thumb, rather than resulting from optimization in a dynamic general equilibrium setup. We think that the Covid-19 shock was

[1] Sources: ONS estimate: `https://www.ons.gov.uk/economy/grossdomesticproductgdp/bulletins/gdpfirstquarterlyestimateuk/apriltojune2020`; median forecast of institutions and financial firms (in May): `https://www.gov.uk/government/statistics/forecasts-for-the-uk-economy-may-2020`; forecast by the Bank of England (in May): `https://www.bankofengland.co.uk/-/media/boe/files/monetary-policy-report/2020/may/monetary-policy-report-may-2020`.

so sudden and unexpected that agent expectations had little time to converge to an equilibrium over the short time period that we consider (Evans & Honkapohja 2012). Our work here is one of several studies using sectoral models with input-output linkages that appeared in the first few months of the pandemic (Barrot et al. 2020, Mandel & Veetil 2020*a*, Fadinger & Schymik 2020, Bonadio et al. 2020, Baqaee & Farhi 2020, Guan et al. 2020, Richiardi et al. 2020). Several further papers in this tradition have been put forward in the meantime (Reissl et al. 2021, Delli Gatti & Grugni 2021).[2] Our work belongs to this effort, but differs in a number of important ways.

The most important conceptual difference is our treatment of the production function. The most common production functions can be ordered by the degree to which they allow substitutions between inputs. At one extreme, the Leontief production function assumes a fixed recipe for production, allowing no substitutions and restricting production based on the limiting input (Inoue & Todo 2020). Under the Leontief production function, if a single input is reduced, overall production will be reduced proportionately, even if that input is ordinarily relatively small. This can lead to unrealistic behaviours. For example, the steel industry has restaurants as an input, presumably because steel companies have a workplace canteen and sometimes entertain their clients and employees. A literal application of the Leontief production function predicts that a sharp drop in the output of the restaurant industry will dramatically reduce steel output. This is unrealistic, particularly in the short run.

The alternatives used in the literature are the linear production function implying an infinite elasticity of substitution (Richiardi et al. 2020), the Cobb-Douglas production function (Fadinger & Schymik 2020), which has an elasticity of substitution of 1, and the CES production function, where typically calibration for short term analysis uses an elasticity of substitution less than 1 (Barrot et al. 2020, Mandel & Veetil 2020*a*, Bonadio et al. 2020). Some papers (Baqaee & Farhi 2020) consider a nested CES production function, which can accommodate a wide range of technologies. In principle, this can be used to allow substitution between some inputs and forbid it between others in an industry-specific manner. However, it is hard to calibrate all the elasticities, so that in practice many models end up using only limited nesting structure or assuming uniform substitutability. Consider again our example of the steel

[2]Our work is also related to some Agent-Based Models (ABMs) of the Covid-19 pandemic (Basurto et al. 2020, Delli Gatti & Reissl 2020, Sharma et al. 2021). While our model does not consider individual households and firms, it is simulated forward in time, rather than "solved" for some set of prices and quantities that clear the market. In this sense, our dynamic input-output model is closer to ABMs than to general equilibrium models.

industry: Under the calibrations of the CES production function that are typically used, firms can substitute energy or even restaurants for iron, while still producing the same output. For situations like this, where the production process requires a fixed technological recipe, this is obviously unrealistic.

To solve this problem we introduce a new production function that distinguishes between critical and non-critical inputs at the level of the 55 industries in the World Input-Output tables. This production function allows firms to keep producing as long as they have the inputs that are absolutely necessary, which we call *critical inputs*. The steel industry cannot produce steel without the critical inputs iron and energy, but it can operate for a considerable period of time without non-critical inputs such as restaurants or management consultants. We apply the Leontief function only to the critical inputs, ignoring the others. Thus we make the assumption that during the pandemic the steel industry requires iron and energy in the usual fixed proportions, but the output of the restaurant or management consultancy industries is irrelevant. Of course restaurants and logistics consultants are useful to the steel industry in normal times – otherwise they wouldn't use them. But during the short time-scale of the pandemic, we believe that neglecting them provides a better approximation of economic behavior than either a Leontief or a CES production function with uniform elasticity of substitution. In the appendix, we show that our production function is very close to a limiting case of an appropriately constructed nested CES, which we could have used in principle, but is less well-suited to our calibration procedure.

To determine which inputs are critical and which are not, we use a survey performed by IHS Markit at our request. This survey asked "Can production continue in industry X if input Y is not available for two months?". The list of possible industries X and Y was drawn from the 55 industries in the World Input-Output Database. This question was presented to 30 different industry analysts who were experts of industry X. Each of them was asked to rate the importance of each of its inputs Y. They assigned a score of 1 if they believed input Y is critical, 0 if it is not critical, and 0.5 if it is in-between, with the possibility of a rating of NA if they could not make a judgement. We then apply the Leontief function to the list of critical inputs, ignoring non-critical inputs. We experimented with several possible treatments for industries with ratings of 0.5 and found that we get somewhat better empirical results by treating them as half-critical (though at present we do not have sufficient evidence to resolve this question unambiguously).

Besides the bespoke production function discussed above, we also introduce a Covid-19-specific treatment of consumption. Most models do not incorporate the

demand shocks that are caused by changes in consumer preferences in order to minimize risk of infection. The vast majority of the literature has focused on the ability to work from home, and some studies incorporate lists of essential vs. inessential industries, but almost no papers have also explicitly added shocks to consumer preferences. (Baqaee & Farhi (2020) is an exception, but the treatment is only theoretical). Here we use the estimates from del Rio-Chanona et al. (2020), which are taken from a prospective study by the Congressional Budget Office (2006). These estimates are crude, but we are not aware of estimates that are any better. The currently available data on actual consumption is qualitatively consistent with the shocks predicted by the CBO, with massive shocks to the hospitality industry, travel and recreation, milder (but large) shocks elsewhere (Andersen et al. 2020, Carvalho et al. 2020, Chen et al. 2021, Surico et al. 2020). The largest mismatch between the CBO estimates and consumption data is in the healthcare sector, whose consumption decreased during the pandemic, in contrast to the estimated increase by the CBO. There was also an increase in some specific retail categories (groceries) which the CBO estimates did not consider. However, overall, as del Rio-Chanona et al. (2020) argue, the estimates remain qualitatively accurate. Besides the initial shock, we also attempt to introduce realistic dynamics for recovery and for savings. The shocks to on-site consumption industries are more long lasting, and savings from the lack of consumption of specific goods and services during lockdown are only partially reallocated to other expenses.

A key finding is that there is a trade-off between the severity of the shocks and the rigidity of the production function. At one extreme, severe shocks with a Leontief production function lead to an almost complete collapse of the economy, and at the other, small shocks with a linear production function fail to reproduce the massive recession observed in the data. There is a sweet spot in the middle, with empirically good results based on various mixes of production function rigidity and shock severity.

We also investigate several theoretical properties of the model. Our analysis makes clear that bottlenecks in supply chains can strongly suppress aggregate economic output. The extent to which this is true depends on the production function. These effects are extremely strong with the Leontief production function, are much weaker with a linear production function (which allows unrealistically strong substitutions) and have an intermediate effect with our modified Leontief function. Network effects can strongly inhibit recovery, and can cause counter-intuitive results, such as situations in which reopening a few industries can actually depress economic output. We further ask whether popular metrics of centrality, such as upstreamness and output multipliers, are useful to understand the diffusion of supply and demand shocks in

our model. We find that static measures are only partially able to explain modeling results. Nevertheless, an industry's upstreamness is a strong indicator of its potential to amplify shocks - this is true for supply shocks, as expected, and for demand shocks, which is somewhat more surprising.

3.2 A dynamic input-output model

Our model combines elements of the input-output models developed by Battiston et al. (2007), Hallegatte (2008), Henriet et al. (2012) and Inoue & Todo (2019), together with new features that make the model more realistic in the context of a pandemic-induced lockdown.

The main data input to our analysis is the UK input-output network obtained from the latest year (2014) of the World Input-Output Database (WIOD) (Timmer et al. 2015), allowing us to distinguish 55 sectors.

3.2.1 Timeline

A time step t in our economy corresponds to one day. There are N industries, one representative firm for each industry, and one representative household. The economy initially rests in a steady-state until it experiences exogenous lockdown shocks. These shocks can affect the supply side (labor compensation, productive capacity) and the demand side (preferences, aggregate spending/saving) of the economy. Every day:

1. Firms hire or fire workers depending on whether their workforce was insufficient or redundant to carry out production in the previous day.

2. The representative household decides its consumption demand and industries place orders for intermediate goods.

3. Industries produce as much as they can to satisfy demand, given that they could be limited by lack of critical inputs or lack of workers.

4. If industries do not produce enough, they distribute their production to final consumers and to other industries on a pro rata basis, that is, proportionally to demand.

5. Industries update their inventory levels, and labor compensation is distributed to workers.

3.2.2 Model description

It will become important to distinguish between demand, that is, orders placed by customers to suppliers, and actual realized transactions, which might be lower.

3.2.2.1 Demand

Total demand. The total demand faced by industry i at time t, $d_{i,t}$, is the sum of the demand from all its customers,

$$d_{i,t} = \sum_{j=1}^{N} O_{ij,t} + c^d_{i,t} + f^d_{i,t}, \tag{3.1}$$

where $O_{ij,t}$ (for *orders*) denotes the intermediate demand from industry j to industry i, $c^d_{i,t}$ represents (final) demand from households and $f^d_{i,t}$ denotes all other final demand (e.g. government or non-domestic customers).

Intermediate demand. Intermediate demand follows a dynamics similar to the one studied in Henriet et al. (2012), Hallegatte (2014), and Inoue & Todo (2019). Specifically, the demand from industry i to industry j is

$$O_{ji,t} = A_{ji} d_{i,t-1} + \frac{1}{\tau}[n_i Z_{ji,0} - S_{ji,t}]. \tag{3.2}$$

Intermediate demand thus is the sum of two components. First, to satisfy incoming demand (from $t-1$), industry i demands an amount $A_{ji}d_{i,t-1}$ from j. Therefore, industries order intermediate inputs in fixed proportions of total demand, with the proportions encoded in the technical coefficient matrix $\mathbf{A}$, i.e. $A_{ji} = Z_{ji,0}/x_{i,0}$ where $Z_{ji,0}$ is realized intermediate consumption and $x_{i,0}$ is total output of industry i. Before the shocks, both of these variables are considered to be in the pre-pandemic steady state. While we will consider several scenarios where industries do not strictly rely on fixed recipes in production, demand always depends on the technical coefficient matrix.

The second term in Eq. (3.2) describes intermediate demand induced by desired reduction of inventory gaps. Due to the dynamic nature of the model, demanded inputs cannot be used immediately for production. Instead industries use an inventory of inputs in production. $S_{ji,t}$ denotes the stock of input j held in i's inventory. Each industry i aims to keep a target inventory $n_i Z_{ji,0}$ of every required input j to ensure production for n_i further days.[3] The parameter τ indicates how quickly

[3]Considering an input-specific target inventory would require generalizing n_i to a matrix with elements n_{ji}, which is easy in our computational framework but difficult to calibrate empirically.

an industry adjusts its demand due to an inventory gap. Small τ corresponds to responsive industries that aim to close inventory gaps quickly. In contrast, if τ is large, intermediate demand adjusts slowly in response to inventory gaps.

Consumption demand. We let consumption demand for good i be

$$c_{i,t}^d = \theta_{i,t}\tilde{c}_t^d, \tag{3.3}$$

where $\theta_{i,t}$ is a preference coefficient, giving the share of goods from industry i out of total consumption demand $\tilde{c}_t^d$. The coefficients $\theta_{i,t}$ evolve exogenously, following assumptions on how consumer preferences change due to exogenous shocks (in our case, differential risk of infection across industries, see Chapter 3.3).

Total consumption demand evolves following an adapted and simplified version of the consumption function in Muellbauer (2020). In particular, $\tilde{c}_t^d$ evolves according to

$$\tilde{c}_t^d = \left(1 - \tilde{\epsilon}_t^D\right) \exp\left\{\rho \log \tilde{c}_{t-1}^d + \frac{1-\rho}{2} \log\left(m\tilde{l}_t\right) + \frac{1-\rho}{2} \log\left(m\tilde{l}_t^p\right)\right\} \tag{3.4}$$

(The exact specification of Eq. (3.4) has been contributed mainly by Marco Pangallo.) In the equation above, the factor $\left(1 - \tilde{\epsilon}_t^D\right)$ accounts for direct aggregate shocks and will be explained in Chapter 3.3. The second factor accounts for the endogenous consumption response to the state of the labor market and future income prospects. In particular, $\tilde{l}_t$ is current labor income, $\tilde{l}_t^p$ is an estimation of permanent income (see Chapter 3.3), and m is the share of labor income that is used to consume final domestic goods, i.e. that is neither saved nor used for consumption of imported goods. In the pre-pandemic steady state with no aggregate exogenous shocks, $\tilde{\epsilon}_t^D = 0$ and by definition permanent income corresponds to current income, i.e. $\tilde{l}_t^p = \tilde{l}_t$. In this case, total consumption demand corresponds to $m\tilde{l}_t$.[4]

Other components of final demand. In addition, an industry i also faces demand $f_{i,t}^d$ from sources that we do not model as endogenous variables in our framework, such as government or industries in foreign countries. We discuss the composition and calibration of $f_{i,t}^d$ in detail in Chapter 3.3.

[4]To see this, note that in the steady state $\tilde{c}_t^d = \tilde{c}_{t-1}^d$. Taking logs on both sides, moving the consumption terms on the left hand side and dividing by $1-\rho$ throughout yields $\log \tilde{c}_t^d = \log\left(m\tilde{l}_t\right)$.

3.2.2.2 Supply

Every industry aims to satisfy incoming demand by producing the required amount of output. Production is subject to the following two economic constraints:

Productive capacity. First, an industry has finite production capacity $x_{i,t}^{\text{cap}}$, which depends on the amount of available labor input. Initially every industry employs $l_{i,0}$ of labor and produces at full capacity $x_{i,0}^{\text{cap}} = x_{i,0}$. We assume that productive capacity depends linearly on labor inputs,

$$x_{i,t}^{\text{cap}} = \frac{l_{i,t}}{l_{i,0}} x_{i,0}^{\text{cap}}. \tag{3.5}$$

Input bottlenecks. Second, the production of an industry might be constrained due to an insufficient supply of critical inputs. This can be caused by production network disruptions. Intermediate input-based production capacities depend on the availability of inputs in an industry's inventory and its production technology, i.e.

$$x_{i,t}^{\text{inp}} = \text{function}_i(S_{ji,t}, A_{ji}). \tag{3.6}$$

We consider five different specifications for how input shortages impact production, ranging from a Leontief form, where inputs need to be used in fixed proportions, to a Linear form, where inputs can be substituted arbitrarily. As intermediate cases we consider specifications with industry-specific dependencies of inputs. For this purpose, IHS Markit analysts rated at our request whether a given input is *critical*, *important* or *non-critical* for the production of a given industry (see Appendix C.3 for details). We then make different assumptions on how the criticality ratings of inputs affect the production of an industry. We now introduce the five specifications in order of stringency with respect to inputs.

(1) Leontief: As first case we consider the Leontief production function, in which every positive entry in the technical coefficient matrix A is a binding input to an industry. This is the most rigid case we are considering, leading to the functional form

$$x_{i,t}^{\text{inp}} = \min_{\{j:\, A_{ji}>0\}} \left\{ \frac{S_{ji,t}}{A_{ji}} \right\}. \tag{3.7}$$

In this case, an industry would halt production immediately if inventories of any input are run down, even for small and potentially negligible inputs.

(2) IHS1: As most stringent case based on the IHS Markit ratings, we assume that production is constrained by critical and important inputs, which need to use in fixed

proportions. In contrast to the Leontief case, however, production is not constrained by the lack of non-critical inputs. Thus, an industry's production capacity with respect to inputs is

$$x_{i,t}^{\text{inp}} = \min_{j \in \{\mathcal{V}_i \cup \mathcal{U}_i\}} \left\{ \frac{S_{ji,t}}{A_{ji}} \right\}, \tag{3.8}$$

where $\mathcal{V}_i$ is the set of *critical* inputs and $\mathcal{U}_i$ is the set of *important* inputs to industry i.

(3) IHS2: As second case using the input ratings, we leave the assumptions regarding *critical* and *non-critical* inputs unchanged but assume that the lack of an *important* input reduces an industry's production by a half. We implement this production scenario as

$$x_{i,t}^{\text{inp}} = \min_{\{j \in \mathcal{V}_i,\, k \in \mathcal{U}_i\}} \left\{ \frac{S_{ji,t}}{A_{ji}}, \frac{1}{2} \left(\frac{S_{ki,t}}{A_{ki}} + x_{i,0}^{\text{cap}} \right) \right\}. \tag{3.9}$$

This means that if an *important* input goes down by 50% compared to initial levels, production of the industry would decrease by 25%. When the stock of this input is fully depleted, production drops to 50% of initial levels.

(4) IHS3: Next we treat all *important* inputs as *non-critical*, such that only *critical* suppliers can create input bottlenecks. This reduces the input bottleneck equation, Eq. (3.6), to

$$x_{i,t}^{\text{inp}} = \min_{j \in \mathcal{V}_i} \left\{ \frac{S_{ji,t}}{A_{ji}} \right\}. \tag{3.10}$$

(5) Linear: Finally, we also implement a linear production function for which all inputs are perfectly substitutable. Here, production in an industry can continue even when inputs cannot be provided, as long as there is sufficient supply of alternative inputs. In this case we have

$$x_{i,t}^{\text{inp}} = \frac{\sum_j S_{ji,t}}{\sum_j A_{ji}}. \tag{3.11}$$

Note that while production is linear with respect to intermediate inputs, the lack of labor supply cannot be compensated by other inputs.

We assume for any production function that imports never cause bottlenecks. Thus, imports are treated as non-critical inputs or, equivalently, there is no shortages in foreign intermediate goods.

Input bottlenecks are most likely to arise under the Leontief assumption, and least likely under the linear production function. The IHS production functions assume intermediate levels of input specificity. In Appendix C.4 we show that the IHS production functions are almost equivalent to (suitably parameterised) CES production functions and that results are even identical using these CES specifications instead of the IHS production functions.

Output level choice and input usage. Since an industry aims to satisfy incoming demand within its production constraints, realized production at time step t is

$$x_{i,t} = \min\{x^{\text{cap}}_{i,t}, x^{\text{inp}}_{i,t}, d_{i,t}\}. \tag{3.12}$$

Thus, the output level of an industry is constrained by the smallest of three values: labor-constrained production capacity $x^{\text{cap}}_{i,t}$, intermediate input-constrained production capacity $x^{\text{inp}}_{i,t}$, or total demand $d_{i,t}$.

The output level $x_{i,t}$ determines the quantity used of each input according to the production recipe. Industry i uses an amount $A_{ji}x_{i,t}$ of input j, unless j is not critical and the amount of j in i's inventory is less than $A_{ji}x_{i,t}$. In this case, the quantity consumed of input j by industry i is equal to the remaining inventory stock of j-inputs $S_{ji,t} < A_{ji}x_{i,t}$.

Rationing. Without any adverse shocks, industries can always meet total demand, i.e. $x_{i,t} = d_{i,t}$. However, in the presence of production capacity and/or input bottlenecks, industries' output may be smaller than total demand (i.e., $x_{i,t} < d_{i,t}$) in which case industries ration their output across customers. We assume simple proportional rationing, although alternative rationing mechanisms could be considered (Pichler & Farmer 2021). The final delivery from industry j to industry i is a share of orders received

$$Z_{ji,t} = O_{ji,t}\frac{x_{j,t}}{d_{j,t}}. \tag{3.13}$$

Households receive a share of their demand

$$c_{i,t} = c^{d}_{i,t}\frac{x_{i,t}}{d_{i,t}}, \tag{3.14}$$

and the realized final consumption of agents with exogenous final demand is

$$f_{i,t} = f^{d}_{i,t}\frac{x_{i,t}}{d_{i,t}}. \tag{3.15}$$

Inventory updating. The inventory of i for every input j is updated according to

$$S_{ji,t+1} = \min\left\{S_{ji,t} + Z_{ji,t} - A_{ji}x_{i,t}, 0\right\}. \tag{3.16}$$

In a Leontief production function, where every input is critical, the minimum operator would not be needed since production could never continue once inventories are run down. It is necessary here, since industries can produce even after inventories of non-critical input j are depleted and inventories cannot turn negative.

Hiring and separations. Firms adjust their labor force depending on which production constraints in Eq. (3.12) are binding. If the capacity constraint $x_{i,t}^{\text{cap}}$ is binding, industry i decides to hire as many workers as necessary to make the capacity constraint no longer binding. Conversely, if either input constraints $x_{i,t}^{\text{inp}}$ or demand constraints $d_{i,t}$ are binding, industry i lays off workers until capacity constraints become binding. More formally, at time t labor demand by industry i is given by $l_{i,t}^{d} = l_{i,t-1} + \Delta l_{i,t}$, with

$$\Delta l_{i,t} = \frac{l_{i,0}}{x_{i,0}} \left[\min\{x_{i,t}^{\text{inp}}, d_{i,t}\} - x_{i,t}^{\text{cap}} \right]. \tag{3.17}$$

The term $l_{i,0}/x_{i,0}$ reflects the assumption that the labor share in production is constant over the considered period. We assume that it takes time for firms to adjust their labor inputs. Specifically, we assume that industries can increase their labor force only by a fraction γ_{H} in direction of their target. Similarly, industries can decrease their labor force only by a fraction γ_{F} in the direction of their target. In the absence of strong labor market regulations, we usually have $\gamma_{\text{F}} > \gamma_{\text{H}}$, indicating that it is easier for firms to lay off employed workers than to hire new workers. Industry-specific employment evolves then according to

$$l_{i,t} = \begin{cases} l_{i,t-1} + \gamma_{\text{H}} \Delta l_{i,t} & \text{if } \Delta l_{i,t} \geq 0, \\ l_{i,t-1} + \gamma_{\text{F}} \Delta l_{i,t} & \text{if } \Delta l_{i,t} < 0. \end{cases} \tag{3.18}$$

The parameters γ_{H} and γ_{F} can be interpreted as policy variables. For example, the implementation of a furloughing scheme makes re-hiring of employees easier, corresponding to an increase in γ_{H}. (Eq. (3.18) has been contributed primarily by Marco Pangallo.)

3.3 Pandemic shock

Simulations of the model described in Chapter 3.2.2 start in the pre-pandemic steady state. While there is evidence that consumption started to decline prior to the lockdown (Surico et al. 2020), for simplicity we apply the pandemic shock all at once, at the date of the start of the lockdown (23 March 2020 in the UK).

We initialize the model with pandemic supply and demand shocks derived by del Rio-Chanona et al. (2020) that we have adapted to accommodate for the UK

context.[5] During the lockdown, workers who cannot work on-site and are unable to work from home become unproductive, resulting in lowered productive capacities of industries. At the same time demand-side shocks hit as consumers adjust their consumption preferences to avoid getting infected, and reduce overall consumption out of precautionary motives due to the depressed state of the economy.

3.3.1 Supply shocks

We have studied several supply shock scenarios and compared how they perform in terms of aggregate and sectoral forecast accuracy. In an extensive validation exercise discussed in Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021) we have chosen the best performing scenario as our "baseline" supply shock scenario. Since the empirical validation of various supply shock types has mostly been done by Marco Pangallo, we refer to Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021) for details. Below we focus only on the chosen baseline scenario.

At every time step during the lockdown an industry i experiences an (exogenous) first-order labor supply shock $\epsilon^S_{i,t} \in [0, 1]$ that quantifies reductions in labor availability. Letting $l_{i,0}$ be the initial labor supply before the lockdown, the maximum amount of labor available to industry i at time t is given as

$$l^{\max}_{i,t} = (1 - \epsilon^S_{i,t}) l_{i,0}. \tag{3.19}$$

If $\epsilon^S_{i,t} > 0$, the productive capacity of industry i is smaller than in the initial state of the economy. We assume that the reduction of total output is proportional to the loss of labor. In that case the productive capacity of industry i at time t is

$$x^{\text{cap}}_{i,t} = \frac{l_{i,t}}{l_{i,0}} x^{\text{cap}}_{i,0} \leq (1 - \epsilon^S_{i,t}) x_{i,0}. \tag{3.20}$$

Recall from Chapter 3.2.2.2 that firms can hire and fire to adjust their productive capacity to demand and supply constraints. Thus, productive capacity can be lower than the initial supply shock, in case industry i has some idle workers that are not prevented to go to work by lockdown measures. In any case, during lockdown firms can

[5] As described below, here we use estimates of the shocks that we develop a priori, based on empirically motivated assumptions on labor supply constraints and changes in preferences. We could have, in principle, used a traditional macro approach (see Brinca et al. (2020) in the case of Covid-19) where one can infer the shocks based on data. We did not pursue this here for several reasons. Our model does not feature price changes (which are crucial to separately identify supply and demand shocks), and we would have had to develop a method to infer the shocks based on model's fit to the data, a problem that is unlikely to have a unique solution. Most importantly, inferring shocks this way is only useful in hindsight, whereas our goal was to make predictions.

never hire more than $l_{i,t}^{\max}$ workers. When the lockdown is lifted for a specific industry i, first-order supply shocks are removed, i.e., we set $\epsilon_{i,t}^{S} = 0$, for $t \geq t_{\text{end\_lockdown}}$.

To estimate the supply shocks we follow del Rio-Chanona et al. (2020) by calculating for each industry the number of workers who can work remotely, and considering the government regulations that dictate whether an industry is essential, i.e. whether an industry can operate during a lockdown even if working from home is not possible. For instance, if an industry is non-essential, and none of its employees can work from home, it faces a labor supply reduction of 100% during lockdown, i.e. $\epsilon_{i,t}^{S} = 1 \ \forall t \in [t_{\text{start\_lockdown}}, t_{\text{end\_lockdown}}$. Instead, if an industry is classified as fully essential, it faces no labor supply shock and $\epsilon_{i,t}^{S} = 0 \ \forall t$.

We define a Remote Labor Index RLI_i and an Essential Score ESS_i at the WIOD industry level. The Remote Labor Index of industry i is the probability that a worker from industry i can work from home. The Essential Score is the probability that a worker from industry i has an essential job. The supply shock is then given by

$$\epsilon_{i}^{S} = (1 - \text{RLI}_i)(1 - \text{ESS}_i). \tag{3.21}$$

In Appendix C.1 we show in detail how we have derived the Remote Labor Index and the Essential Score and provide further industry-level shock statistics.

3.3.2 Demand shocks

The Covid-19 pandemic caused strong shocks to all components of demand. We consider shocks to private consumption demand, which we further distinguish into shocks due to fear of infection and due to fear of unemployment, and shocks to other components of final demand, such as investment, government consumption and exports. (Marco Pangallo is the prime contributer to the exact definitions of private consumption shocks due to infections and unemployment shown below.) We outline the basic assumptions on demand shocks below and show in in Appendix C.1 the detailed cross-sectional and temporal shock profiles. We further demonstrate in Appendix C.5 that alternative plausible demand shock assumptions only mildly influence model results.

Demand shocks due to fear of infection. During a pandemic, consumption/ saving decisions and consumer preferences over the consumption basket are changing, leading to first-order demand shocks (Congressional Budget Office 2006, del Rio-Chanona et al. 2020). For example, consumers are likely to demand less services from the hospitality industry, even when the hospitality industry is open. Transport is very likely to face substantial demand reductions, despite being classified as an essential

industry in many countries. A key question is whether reductions in demand for "risky" goods and services is compensated by an increase in demand for other goods and services, or if lower demand for risky goods translates into higher savings.

We consider a demand shock vector ϵ_t^D, whose components $\epsilon_{i,t}^D$ are the relative changes in demand for goods of industry i at time t. Recall from Eq. (3.3), $c_{i,t}^d = \theta_{i,t}\tilde{c}_t^d$, that consumption demand is the product of the total consumption scalar $\tilde{c}_t^d$ and the preference vector θ_t, whose components $\theta_{i,t}$ represent the share of total demand for good i. We initialize the preference vector by considering the initial consumption shares, that is $\theta_{i,0} = c_{i,0}/\sum_j c_{j,0}$. By definition, the initial preference vector θ_0 sums to one, and we keep this normalization at all following time steps. To do so, we consider an auxiliary preference vector $\bar{\theta}_t$, whose components $\bar{\theta}_{i,t}$ are obtained by applying the shock vector $\epsilon_{i,t}^D$. That is, we define $\bar{\theta}_{i,t} = \theta_{i,0}(1 - \epsilon_{i,t}^D)$ and define $\theta_{i,t}$ as

$$\theta_{i,t} = \frac{\bar{\theta}_{i,t}}{\sum_j \bar{\theta}_{j,t}} = \frac{(1 - \epsilon_{i,t}^D)\theta_{i,0}}{\sum_j (1 - \epsilon_{j,t}^D)\theta_{j,0}}. \tag{3.22}$$

The difference $1 - \sum_i \bar{\theta}_{i,t}$ is the aggregate reduction in consumption demand due to the demand shock, which would lead to an equivalent increase in the saving rate. However, households may not want to save all the money that they are not spending. For example, they most likely want to spend on food the money that they are saving on restaurants. Therefore, we define the aggregate demand shock $\tilde{\epsilon}_t^D$ in Eq. (3.4) as

$$\tilde{\epsilon}_t^D = \Delta s \left(1 - \sum_{i=1}^{N} \bar{\theta}_{i,t}\right), \tag{3.23}$$

where Δs is the change in the savings rate. When $\Delta s = 1$, households save all the money that they are not planning to spend on industries affected by demand shocks; when $\Delta s = 0$, they spend all that money on goods and services from industries that are affected less.

To parameterize $\epsilon_{i,t}^D$, we adapt consumption shock estimates by the Congressional Budget Office (2006) and del Rio-Chanona et al. (2020). Roughly speaking, these shocks are massive for restaurants and transport, mild for manufacturing and null for utilities. We make two modifications to these estimates. First, we remove the positive shock to the health care sector, as in the UK the cancellation of non-urgent treatment for other diseases than Covid-19 far exceeded the additional demand for health due to Covid-19.[6] Second, we apportion to manufacturing sectors the reduced

[6] `https://www.ons.gov.uk/economy/grossdomesticproductgdp/bulletins/gdpfirstquarterlyestimateuk/apriltojune2020`

demand due to the closure of non-essential retail. For example, retail shops selling garments and shoes were mandated to shut down, and so we apply a consumption demand shock to the manufacturing sector producing these goods.[7]

We keep the intensity of demand shocks constant during lockdown. We then reduce demand shocks when lockdown is lifted according to the situation of the Covid-19 pandemic in the UK. In particular, we assume that consumers look at the daily number of Covid-19 deaths to assess whether the pandemic is coming to an end, and that they identify the end of the pandemic as the day in which the death rate drops below 1% of the death rate at the peak. Given official data[8], this happens on August 11$^{\text{th}}$. Thus, we reduce $\epsilon^D_{i,t}$ from the time lockdown is lifted (13 May 2020) by linearly interpolating between the value of $\epsilon^D_{i,t}$ during lockdown and $\epsilon^D_{i,t} = 0$ on 11 August 2020. The choice of modeling behavioral change in response to a pandemic by the death rate has a long history in epidemiology (Funk et al. 2010).

Demand shocks due to fear of unemployment. A second shock to consumption demand occurs through reductions in current income and expectations for permanent income.

Reductions in current income are due to firing/furloughing, due to both direct shocks and subsequent upstream or downstream propagation, resulting in lower labor compensation, i.e. $\tilde{l}_t < \tilde{l}_0$, for $t \geq t_{\text{start\_lockdown}}$. To support the economy, the government pays out social benefits to workers to compensate income losses. In this case, the total income $\tilde{l}_t$ that enters Eq. (3.4) is replaced by an effective income $\tilde{l}^\star_t = b\tilde{l}_0 + (1-b)\tilde{l}_t$, where b is the fraction of pre-pandemic labor income that workers who are fired or furloughed are able to retain.

A second channel for shocks to consumption demand due to labor market effects occurs through expectations for permanent income. These expectations depend on whether households expect a V-shaped vs. L-shaped recovery, that is, whether they expect that the economy will quickly bounce back to normal or there will be a prolonged recession. Let expectations for permanent income $\tilde{l}^p_t$ be specified by

$$\tilde{l}^p_t = \xi_t \tilde{l}_0 \tag{3.24}$$

[7]To be fully consistent with the definition of demand shock, we should model non-essential retail closures as supply shocks, and propagate the shocks to manufacturing through reduced intermediate good demand. However, there are two practical problems that prevent us to do so: (i) the sectoral aggregation in the WIOD is too coarse, comprising only one aggregate retail sector; (ii) input-output tables only report margins of trade, i.e. they do not model explicitly the flow of goods from manufacturing to retail trade and then from retail trade to final consumption. Given these limitations, we conventionally interpret non-essential retail closures as demand shocks.

[8]`https://coronavirus.data.gov.uk/`

In this equation, the parameter ξ_t captures the fraction of pre-pandemic labor income $\tilde{l}_0$ that households expect to retain in the long run. We first give a formula for ξ_t and then explain the various cases.

$$\xi_t = \begin{cases} 1, & t < t_{\text{start\_lockdown}}, \\ \xi^L = 1 - \frac{1}{2}\frac{\tilde{l}_0 - \tilde{l}_{t_{\text{start\_lockdown}}}}{\tilde{l}_0}, & t \in [t_{\text{start\_lockdown}}, t_{\text{end\_lockdown}}], \\ 1 - \rho + \rho\xi_{t-1} + \nu_{t-1}, & t > t_{\text{end\_lockdown}}. \end{cases} \tag{3.25}$$

Before lockdown, we let $\xi_t \equiv 1$, i.e. permanent income expectations are equal to current income. During lockdown, following Muellbauer (2020) we assume that ξ_t is equal to one minus half the relative reduction in labor income that households experience due to the direct labor supply shock, and denote that value by ξ^L. (For example, given a relative reduction in labor income of 16%, $\xi^L = 0.92$.)[9] After lockdown, we assume that 50% of households believe in a V-shaped recovery, while 50% believe in an L-shaped recovery. We model these expectations by letting ξ_t evolve according to an autoregressive process of order one, where the shock term ν_t is a permanent shock that reflects beliefs in an L-shaped recovery. With 50% of households believing in such a recovery pattern, it is $\nu_t \equiv -(1-\rho)(1-\xi^L)/2$.[10]

Other final demand shock scenarios Note that WIOD distinguishes five types of final demand: (I) *Final consumption expenditure by households*, (II) *Final consumption expenditure by non-profit organisations serving households*, (III) *Final consumption expenditure by government* (IV) *Gross fixed capital formation* and (V) *Changes in inventories and valuables*. Additionally, all final demand variables are available for every country, meaning that it is possible to calculate imports and exports for all categories of final demand. The endogenous consumption variable $c_{i,t}$ corresponds to (I), but only for domestic consumption. All other final demand categories, including all types of exports, are absorbed into the variable $f_{i,t}$.

We apply different shocks to $f_{i,t}$. We do not apply any exogenous shocks to categories (III) *Final consumption expenditure by government* and (V) *Changes in inventories and valuables*, while we apply the same demand shocks to category (II) as

[9] During lockdown, labor income may be further reduced due to firing. For simplicity, we choose not to model the effect of these further firings on permanent income.

[10] The specification in Eq. (3.25) reflects the following assumptions: (i) time to adjustment is the same as for consumption demand, Eq. (3.4); (ii) absent permanent shocks, $\nu_t = 0$ after some t, ξ_t returns to one, i.e. permanent income matches current income; (iii) with 50% households believing in an L-shaped recovery, ξ_t reaches a steady state given by $1 - (1 - \xi^L)/2$: with $\xi^L = 0.92$ as in the example above, ξ_t reaches a steady state at 0.96, so that permanent income remains stuck four percentage points below pre-lockdown income.

we do for category (I). To determine shocks to investment (IV) and exports we start by noticing that, before the Covid-19 pandemic, the volatility of these variables has generally been three times the volatility of consumption.[11] The overall consumption demand shock is around 5% so, as a baseline, we assume shocks to investment and exports to be 15%. In Appendix C.5 we show that the model results are fairly robust with respect to alternative choices.

3.4 Economic impact of Covid-19 on the UK economy

As already mentioned, in an early version of this work (Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2020), we released results in May 2020 predicting a 21.5% contraction of GDP in the UK economy in the second quarter of 2020 with respect to the last quarter of 2019.[12] This is in comparison to the contraction of 22.1% that was actually observed. Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021) further developed the economic model and conducted a "post-mortem" analysis to better understand why the initial predictions were so accurate. Here, we show the model results of the UK economy during the pandemic based on the best-performing model parametrization identified in Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021). The specific set of parameters used in these simulations are shown in Table 3.1.

Some of the parameters can be obtained directly from the data. We calibrate the inventory target parameters n_i using ONS data for the usual stock of inventories that different industries typically have (see details in Appendix C.2). These parameters are highly heterogeneous across industries; typically manufacturing and trade have much higher inventory targets than services. Another parameter which we can directly calibrate from data is the propensity to consume m (see Eq. (3.4)). Directly reading the share of labor income that is used to buy final domestic goods from the input-output tables, we find $m = 0.82$. Finally, we calibrate benefits b based on official UK policy, $b = 0.8$.

[11] This is computed by calculating the standard deviation of consumption, investment and export growth over all quarters from 1970Q1 to 2019Q4. These are 1.03%, 2.87% and 3.24% respectively. Source: `https://www.ons.gov.uk/file?uri=\%2feconomy\%2fgrossdomesticproductgdp\%2fdatasets\%2frealtimedatabaseforukgdpcomponentsfortheexpenditureapproachtothemeasureofgdp\%2fquarter2aprtojune2020firstestimate/gdpexpenditurecomponentsrealtimedatabase.xls`

[12] Due to an error in how we dealt with Real Estate shocks, our prediction was slightly worse than it would have been without the error, see Appendix C.1

Name	Symbol	Value
Inventory targets	n_i	Appendix C.2
Propensity to consume	m	0.82
Government benefits	b	0.80
Production function		IHS2
Inventory adjustment	τ	10
Upward labor adjustment	γ_H	1/30
Downward labor adjustment	γ_F	1/15
Consumption adjustment	ρ	0.99
Change in savings rate	Δs	0.50

Table 3.1: Assumptions and parameters of the model that: (top) are varied across scenarios; (middle) are fixed due to little effect on results; (bottom) are fixed due to direct data calibration.

The other parameters cannot easily been taken from the data. Model results are most sensitive with respect to the exact choice of the production function. We therefore show model runs for different production function assumptions below. As we show in Appendix C.5, the other parameters have less effects on modeling results and so we fix them to reasonable values. These are:

- The parameter τ, capturing responsiveness to inventory gaps. We fix $\tau = 10$ days, which indicates that firms aim at filling most of their inventory gaps within two weeks. This lies in the range of values used by related studies (e.g. $\tau = 6$ in Inoue & Todo (2019), $\tau = 30$ in Hallegatte (2014)).

- The hiring and firing parameters γ_{H} and γ_{F}. We choose $\gamma_{\mathrm{H}} = 1/30$ and $\gamma_{\mathrm{F}} = 2\gamma_{\mathrm{H}}$. Given our daily time scale, this is a rather rapid adjustment of the labor force, with firing happening faster than hiring.

- The parameter ρ, indicating sluggish adjustment to new consumption levels. We select the value assumed by Muellbauer (2020), adjusted for our daily timescale.[13]

[13]Assuming that a time step corresponds to a quarter, Muellbauer (2020) takes $\rho = 0.6$, implying that more than 70% of adjustment to new consumption levels occurs within two and a half quarters. We modify ρ to account for our daily timescale: By letting $\bar{\rho} = 0.6$, we take $\rho = 1 - (1 - \bar{\rho})/90$ to obtain the same time adjustment as in Muellbauer (2020). Indeed, in an autoregressive process like the one in Eq. (3.4), about 70% of adjustment to new levels occurs in a time ι related inversely to the persistency parameter ρ. Letting Q denote the quarterly timescale considered by Muellbauer (2020), time to adjustment ι^Q is given by $\iota^Q = 1/(1 - \bar{\rho})$. Since we want to keep approximately the same time to adjustment considering a daily time scale, we fix $\iota^D = 90\iota^Q$. We then obtain the parameter ρ in the daily timescale such that it yields ι^D as time to adjustment, namely $1/(1-\rho) = \iota^D = 90\iota^Q = 90/(1 - \bar{\rho})$. Rearranging gives the formula that relates ρ and $\bar{\rho}$.

- The savings parameter Δs. We take $\Delta s = 0.5$, meaning that households save half the money they are not spending in goods and services due to fear of infection, and direct half of that money to spending for other "safer" goods and services.

Fig. 3.1 shows model results for production (gross output) under the given model parametrization. When the lockdown starts, there is a sudden drop in economic activity, shown by a sharp decrease in production. Some other industries further decrease production over time as they run out of critical inputs. Throughout the simulation, service sectors tend to perform better than manufacturing, trade, transport and accommodation sectors. The main reason for that is most service sectors face both lower supply and demand shocks, as a high share of workers can effectively work from home, and business and professional services depend less on consumption demand.

We take 13 May 2020 as the date in which lockdown measures are lifted. By the end of June, the economy is still far from recovering. In part, this is due to the fact that the aggregate level of consumption does not return to pre-lockdown levels, due to a reduction in expectations of permanent income associated with beliefs in an L-shaped recovery (Chapter 3.2.2.1), and due to the fact that we do not remove shocks to investment and exports (see Chapter 3.3).

We now turn to evaluating how well the selected scenario describes the economic effects of Covid-19 on the UK economy. In terms of gross output, one can see in Table 3.2 that the model slightly underestimates the recession in April and slightly overestimates it in May, while it correctly estimates a strong recovery in June.[14] Additionally, aggregate value added in the second quarter of 2020 is very close to the data.

Our model, however, considers other macroeconomic variables than gross output and value added, so we also compare these other variables to data (Table 3.2). From national accounts, we collect data on private and government consumption, investment, change in inventories, exports and imports (expenditure approach to GDP); wages and salaries and profits (income approach to GDP). Looking at the expenditure approach to GDP, some variables have a worse reduction in the model, such as investment and exports, while other variables have a worse reduction in the data,

[14]We aggregate empirical output from sectoral indexes using our steady-state output shares as weights.

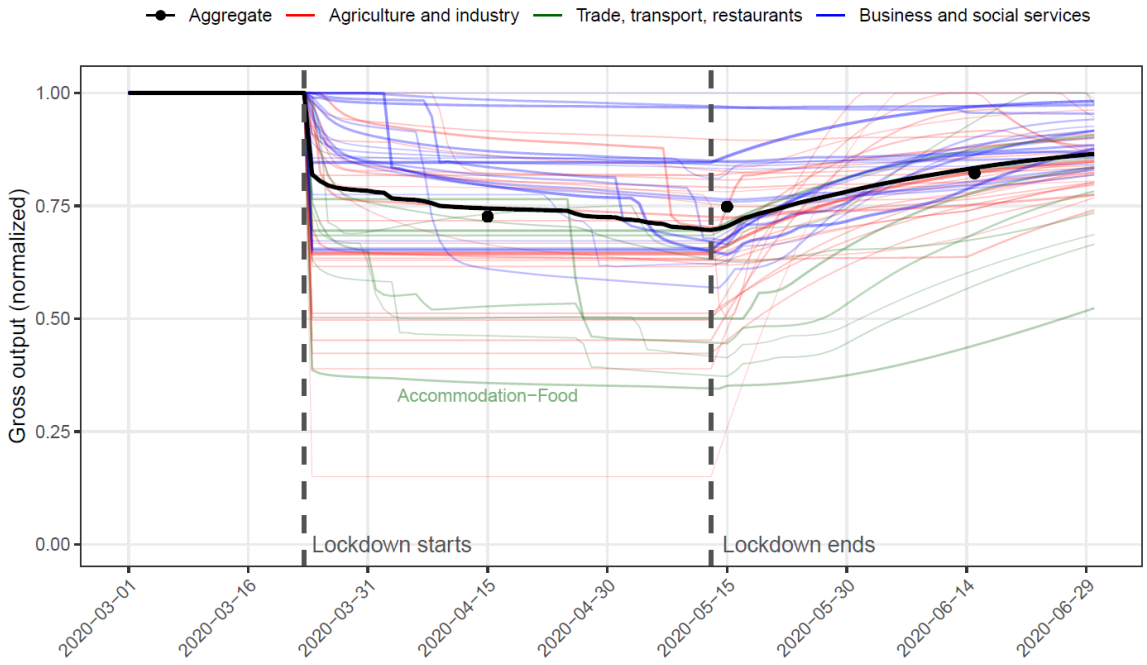

Figure 3.1: **Economic production for the chosen scenario as a function of time.** We plot production (gross output) as a function of time for each of the 55 industries. Aggregate production is a thick black line and each sector is colored. Agricultural and industrial sectors are colored red; trade, transport, and restaurants are colored green; service sectors are colored blue. All sectoral productions are normalized to their pre-lockdown levels, and each line size is proportional to the steady-state gross output of the corresponding sector. For comparison, we also plot empirical gross output as normalized with respect to March 2020.

such as private and government consumption, inventories and imports.[15] However, the model predicts the relative reductions fairly well, as we find a stronger collapse in private consumption and investment than in government consumption or inventories, as in the data. Finally, considering the income approach to GDP, we overestimate the reduction in wages and salaries and underestimate the reduction in profits.[16] Nonetheless, we correctly predict that the absolute reduction in wages and salaries is much smaller than in profits (due to government subsidies).

Turning to the performance of our model at the disaggregate level of industries, Fig. 3.2 shows gross output as predicted by the model and in the data. Here, gross output is averaged over the values it takes in March, April and June, both in the

[15]If one considers Exports-Imports as a component, the model predicts a current account deficit, while in reality there was a current account surplus. We should however note that we do not model international trade, we treat exports as exogenous and imports like locally produced goods and services.

[16]Note that these categories are not jointly exhaustive, as the ONS also considers mixed income and taxes less subsidies, which are difficult to compare to variables in our model.

Variable (compared to Q4-2019)	Data	Model
Gross output April	-27.4%	-25.3%
Gross output May	-25.2%	-26.9%
Gross output June	-17.8%	-16.8%
Value added Q2	-21.5%	-22.1%
Private consumption Q2	-25.3%	-21.3%
Investment Q2	-26.3%	-29.7%
Government consumption Q2	-17.5%	-14.2%
Inventories Q2	-2.2%	-0.5%
Exports Q2	-23.3%	-27.8%
Imports Q2	-30.6%	-23.9%
Wages and Salaries Q2	-1.1%	-4.3%
Profits Q2	-26.7%	-22.3%

Table 3.2: Comparison between data and predictions of the selected scenario for the main aggregate variables. Percentages refer to changes from the last quarter of 2019, which we take to represent the pre-pandemic economic situation, to the second quarter of 2020.

model and in the data, and compared to the value it had in Q4 2019. To interpret this figure, note that for all points on the left of the identity line, model predictions are lower than in the data, i.e. the model is pessimistic. Conversely, on the right of the identity line model predictions are higher than in the data, i.e. the model is optimistic. Note that model predictions and empirical data are correlated, although not perfectly: the Pearson correlation coefficient weighted by industry size is 0.75. The majority of sectors decreased production up to 60% of initial levels, both in the model and in the data, but a few sectors were forced to decrease production much more.

While the predictions are generally very good, there are also some dramatic failures. We conjecture that this is due to idiosyncratic features of these industries that we could not take into account without overly complicating our model. For example, one sector for which the predictions of our model are completely off is C29 - Manufacturing of vehicles. Almost all car manufacturing plants were completely closed in the UK in April and May, and so production was essentially zero (7% of the pre-pandemic level in April and 14% in May). While they reopened in June, production in Q2 is slightly above 20% of the pre-pandemic level. However, our model predicts a level of production around 63% of the pre-pandemic level. It is difficult to account for the complete shutdown of car manufacturing plants in our model, as our selected scenario for supply shocks does not require manufacturing plants to completely close during lockdown. We think that this discrepancy between model predictions and data can be explained by two factors. First, car manufacturing is highly integrated

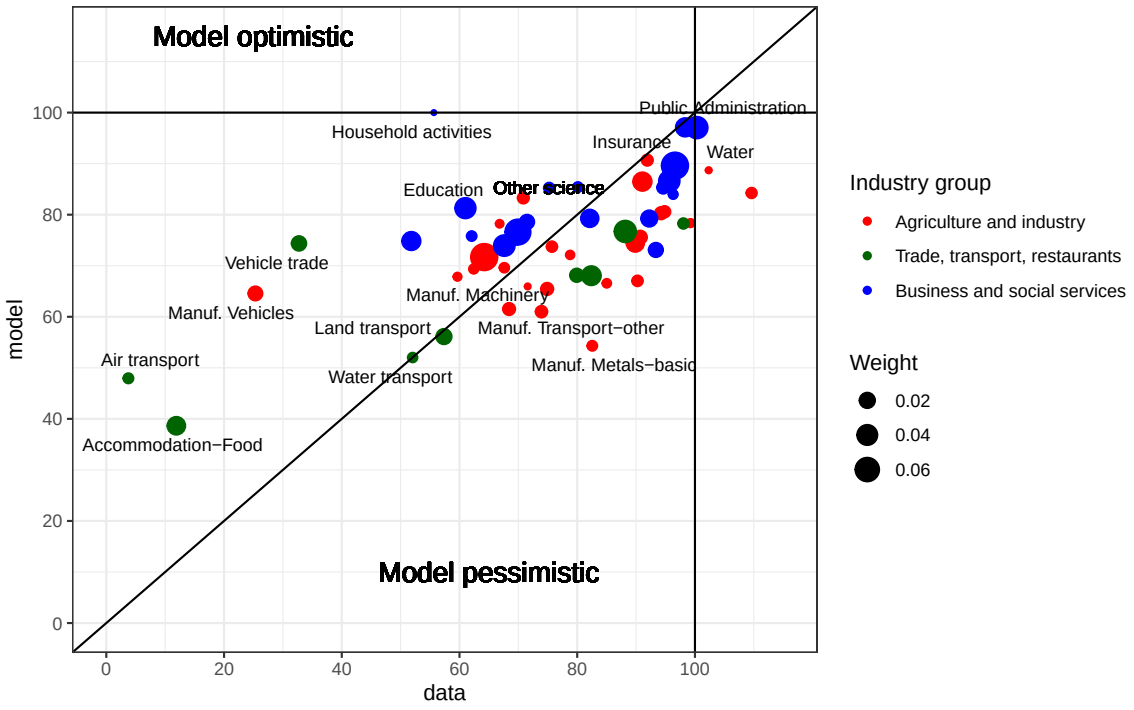

Figure 3.2: **Comparison between model predictions and empirical data.** We plot predicted production (gross output) for each of 53 industries against the actual values from the ONS data, averaged across April, May and June. Values are relative to pre-lockdown levels and dot size is proportional to the steady-state gross output of the corresponding sector. Agricultural and industrial sectors are colored red; trade, transport, and restaurants are colored green; service sectors are colored blue. Dots above the identity line means that the predicted recession is less severe than in the data, while the reverse is true for dots below the identity line.

internationally, and, in a period where most developed countries were implementing lockdown measures, international supply chains were highly disrupted. For simplicity, however, in our model we did not model input bottlenecks due to lack of imported goods. Second, it is possible that firms producing non-essential goods voluntarily decided to stop production to protect the health of their workers, even if they were not forced to do so. Another example for which the predictions of our model are off is H51 - Air transport. Production in the data is 3% of pre-lockdown levels, while our model predicts around 50%. In the model, most activity of the air transport industry during lockdown is due to business travel, which is a non-critical input to many industries that we do not exogenously shut down (recall that industries aim at using non-critical inputs if they are available).

In some other cases, however, our model gave accurate predictions, even when the answer was far from obvious. Compare, for instance, industries M74_M75 (Other Science) and O84 (Public Administration). Both received a very weak supply shock (3% for Other Science and 1.1% for Public Administration), and no private consumption

demand shocks. Yet, Public Administration had almost no reduction in production, while Other Science reduced its production to about 75% of its pre-pandemic level. The ONS report in May quotes reduced intermediate demand as the reason why Other Science reduced its activity. Conversely, Public Administration's output is almost exclusively sold to the government, which did not reduce its consumption. Because of its ability to take into account supply chain effects and the resulting reductions in intermediate demand, our model is able to endogenously capture the difference between these two sectors, even though their shocks are small in both cases.

3.5 Theoretical results

3.5.1 Production functions and input criticality

A key innovation of our model is that it uses a production function that distinguishes between critical and non-critical inputs, based on the IHS Markit analyst ratings (Chapter 3.2.2.2). The most rigid case is the Leontief production function where every input, regardless of its share in the input mix, creates input bottlenecks when missing. Since the IO network at the level of 55 industries is extremely dense ($\approx$ 98% of possible links are present), almost every industry could cause substantial downstream disruptions of economic production under Leontief production. Linear production is the least rigid case, assuming no critical inputs at all. Here, downstream shock propagation only occurs in case of insufficient total input levels. The various IHS production functions lie in-between. Here, only a subset of inputs are critical and required in fixed proportions for short-term production. The IHS1 specification is closest to the Leontief case, considering around 25% of inputs as critical. In contrast, IHS3 with only 15% of inputs being critical is closest to linear production and IHS2 is between the two cases (15% critical, 10% half-critical). To better understand how different assumptions on input criticality influence outcomes, we run simulations under varying production function specifications.

In the left panel of Fig. 3.3 we show simulation results for gross output under alternative production function assumptions. Except for the production function, all other parameters are set as outlined in Chapter 3.4. As expected we find the largest drop in production for the Leontief production function, where every input can potentially become binding (black line). For the Leontief economy our model predicts a drop in gross output of up to 60% compared to initial levels. Again in line with intuition, we obtain the mildest economic impacts for the linear production function

(red). There is still a substantial drop due to first-order shocks, but essentially no further adjustments resulting from network effects.

The production function specifications using the results of the IHS Markit survey lie between these two cases, but much closer to the linear production function. In contrast to the linear production function, however, we observe second-order effects materializing some time after the initial shock hits the economy. While here the overall differences between the various IHS production functions seem to be rather small, it is important to note that these results are not generic. In particular, they largely depend on industry-specific inventory levels. In the right panel of Fig. 3.3 we redo the simulation, but after downscaling inventory levels to half their actual size.[17] In this case the differences among the three IHS production function specifications are much more pronounced.

These results indicate that production function assumptions are central for modeling higher-order effects of economic shocks. Relatedly, inventory levels play a key role too as they can act as buffer against adverse shocks. A better understanding of shock propagation dynamics in production networks thus requires detailed information on industries' inventories and their ability in dealing with shortages of certain types of inputs.

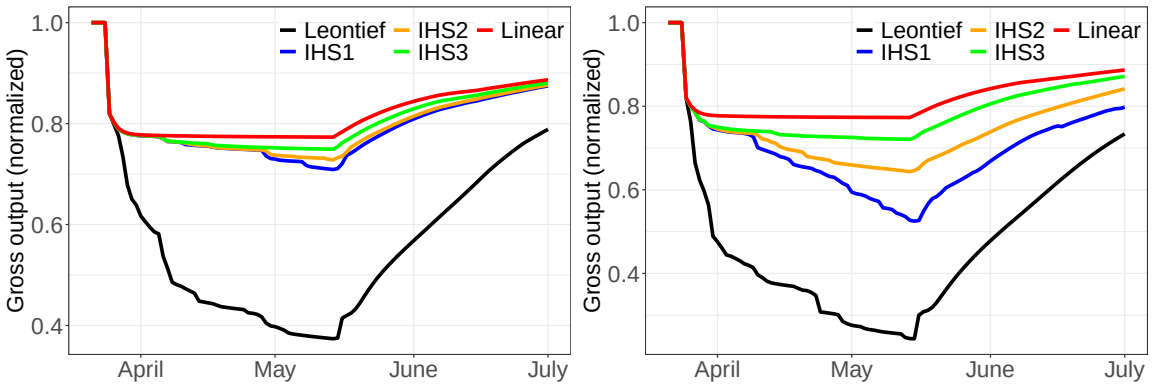

Figure 3.3: **Comparison of economic output under different production functions.** The left panel shows results when the model is parametrized as outlined in Chapter 3.4 and only the production function is varied. The right panel is the same, except that we downscale inventories to a half of actual levels. Shocks are applied on 23 March 2020 and partly relaxed from 13 May 2020 onward as discussed in Chapter 3.4. Output values are normalized to the initial no-lockdown steady state.

[17] We achieve this by taking $n_i^{\text{reduced}} = \max(1, \frac{1}{2} n_i)$ and then substituting n_i^{reduced} for n_i. The maximum operator is needed to ensure that the model starts in steady state, as $n_i^{\text{reduced}} < 1$ would immediately lead to a decline in economic output.

3.5.2 Direct shocks and higher-order impacts

Our model is initialized with the first-order supply and demand shocks discussed in Chapter 3.3 and simulates how the economic system translates these shocks into overall economic impacts. We next investigate the relative contribution of direct shocks and network effects to total economic impacts. Recall that exogenous supply shocks lead to an immediate reduction in gross output. We let the *direct* output reduction be

$$\mathrm{OR}_i^{\mathrm{direct}} \equiv \frac{x_{i,0}^{\mathrm{shocked}} - x_{i,0}}{x_{i,0}} = \frac{(1-\epsilon_{i,0}^S)x_{i,0} - x_{i,0}}{x_{i,0}} = -\epsilon_{i,0}^S, \tag{3.26}$$

which is equivalent to the first-order supply shocks multiplied by minus one. Similarly, exogenous demand shocks instantaneously decrease final consumption. More formally, we define the *direct* final consumption reduction as

$$\mathrm{CR}_i^{\mathrm{direct}} \equiv \frac{c_{i,0}^{\mathrm{shocked}} + f_{i,0}^{\mathrm{shocked}} - c_{i,0} - f_{i,0}}{c_{i,0} + f_{i,0}}, \tag{3.27}$$

where $c_{i,0}^{\mathrm{shocked}} = (1-\epsilon_{i,0}^D)c_{i,0}$ and $f_{i,0}^{\mathrm{shocked}}$ denotes non-household consumption after shocks have been applied as indicated in Chapter 3.3.2.

The upper panel of Fig. 3.4 visualizes for each sector of the UK economy first-order reductions in gross output $\mathrm{OR}_i^{\mathrm{direct}}$ and final consumption $\mathrm{CR}_i^{\mathrm{direct}}$. While some industries such as several manufacturing industries face larger immediate reductions in gross output than final consumption, transport industries (H49-51) experience much larger negative shocks in final consumption. Reductions in both final consumption and gross output are enormous for Accommodation and Food Service Activities (I).

We next quantify higher-order supply- and demand-side impacts. These indirect effects are time-dependent since overall economic performance changes in time as can be seen from the simulations above. Note that higher-order impacts in gross output do not necessarily need to be caused by supply-side shocks but could also result from a lack of demand. Conversely, final consumption reductions can stem from lowered production levels. We let the *total* reduction in output at any time denote

$$\mathrm{OR}_{i,t}^{\mathrm{total}} = x_{i,t}/x_{i,0} - 1.$$

The indirect output reduction is then computed as the residual of the total output reduction $\mathrm{OR}_{i,t}^{\mathrm{total}}$ and direct output reduction $\mathrm{OR}_i^{\mathrm{direct}}$. Thus, for the *indirect* output reduction we have

$$\mathrm{OR}_{i,t}^{\mathrm{indirect}} = \mathrm{OR}_{i,t}^{\mathrm{total}} - \mathrm{OR}_i^{\mathrm{direct}}. \tag{3.28}$$

Similarly, we compute *indirect* final consumption reductions as

$$\mathrm{CR}_{i,t}^{\mathrm{indirect}} = \mathrm{CR}_{i,t}^{\mathrm{total}} - \mathrm{CR}_{i}^{\mathrm{direct}}, \tag{3.29}$$

where $\mathrm{CR}_{i,t}^{\mathrm{total}} = (c_{i,t} + f_{i,t})/(c_{i,0} + f_{i,0}) - 1$ is the *total* reduction in final consumption.

The center panel of Fig. 3.4 shows a scatter plot where the x-axis denotes direct output reductions $\mathrm{OR}_{i}^{\mathrm{direct}}$ and the y-axis indirect output reductions $\mathrm{OR}_{i,t}^{\mathrm{indirect}}$ after one month into lockdown. Points scatter in an inverted L-shape indicating that industries that experience large direct output shocks do not reduce production much more in the course of the lockdown. In contrast, many industries that experience little or no direct shocks to their productive capacities downsize economic production substantially in the following month.

The bottom panel of Fig. 3.4 is the same but for final consumption instead of output. Here, we do not observe a similar inverted L-shape pattern. Instead, several industries lie closer to the identity line, indicating similar first-order and second-order reductions in final consumption. While all higher-order supply-side effects are non-positive, some industries face positive higher-order consumption effects, although they tend to be very small. Positive values on the y-axis indicate that higher-order impacts on final consumption are slightly mitigating initial shocks. The vast majority of industries ($\approx 95\%$), however, still faces negative total consumption effects.

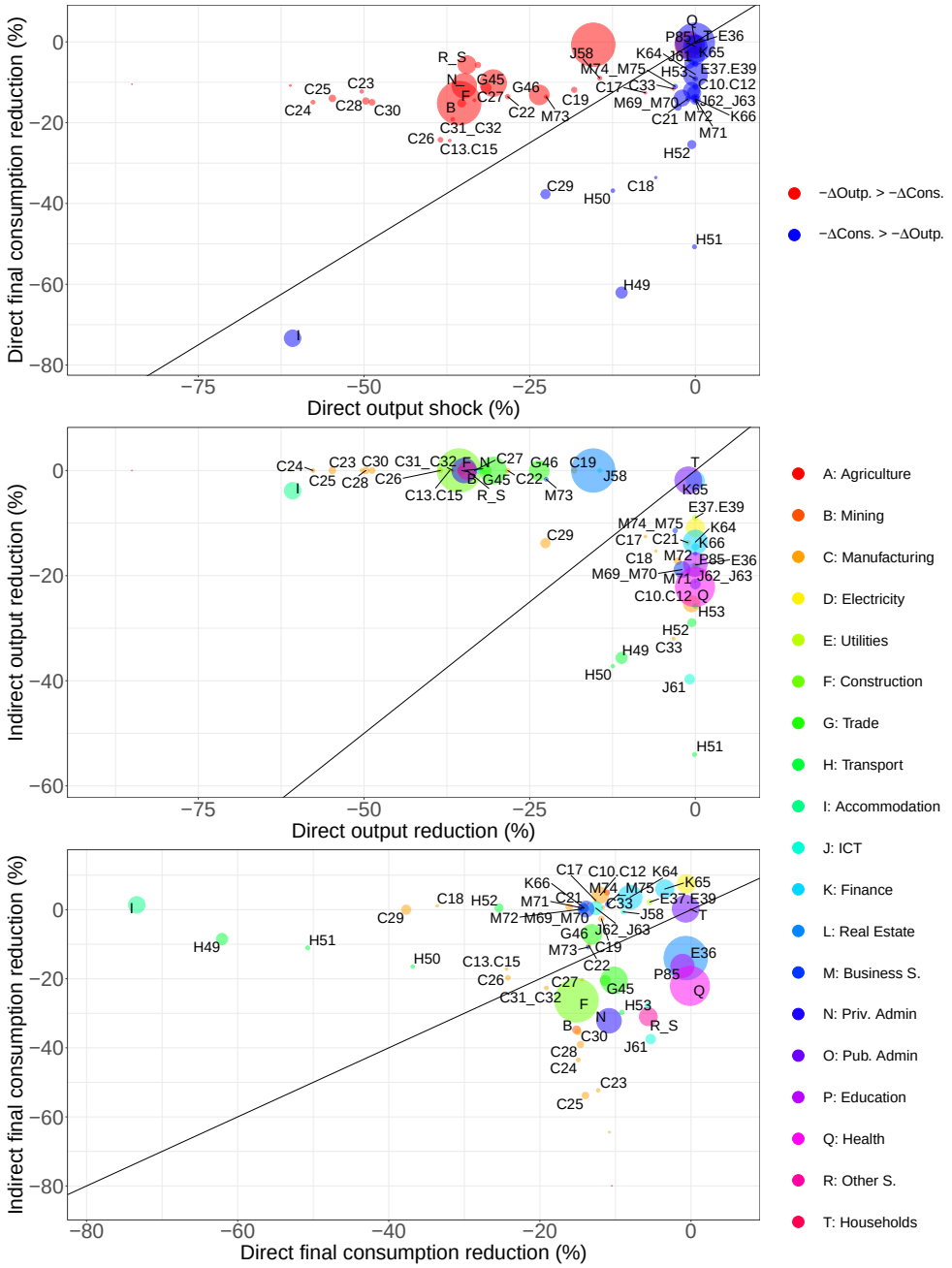

Figure 3.4: *Upper panel:* **Sectoral first-order shocks in the UK economy** on the supply and demand sides. The axes give %-reductions in gross output (x-axis) and in final consumption (y-axis). A blue disk indicates that the monetary shock in absolute terms is larger on the demand side than on the supply side (blue disks can thus lie below the identity line and vice versa). Disk size corresponds to initial gross output of industries. Details on industry labels can be found in Appendix C.3, Table C.1. *Center panel:* **Comparison of direct and indirect output reductions.** Direct reductions in sectoral output (x-axis) plotted against indirect impacts on sectoral production (y-axis). Industries that face a large initial supply shock tend to experience smaller higher-order impacts, while higher-order effects can be large for industries that experienced little or no initial supply shock. Disk size corresponds to initial gross output of industries. *Bottom panel:* **Comparison of direct and indirect final consumption reductions.** Direct final demand reductions (x-axis) plotted against indirect impacts on final demand (y-axis). Disk size corresponds to initial level of final demand satisfied per industry.

3.5.3 Economic impacts of supply, demand and combined shocks

We now demonstrate how the first-order shocks shown in Fig. 3.4 are amplified through the production network and how they affect overall economic impact. To evaluate the relative contributions of supply and demand shocks to overall outcomes, we run the following simulations: First, we run the baseline model setup where both initial supply and demand shocks are present as discussed above. Second, we run the model considering only supply shocks, i.e. we set all consumer demand shocks to zero, $\epsilon_i^D = 0$, as well as remove all shocks to other final demand categories, i.e. $f_{i,t}^d = f_{i,0}^d$. As a third simulation scenario we run the model with all initial supply shocks switched off, $\epsilon_i^S = 0$, and include only initial demand shocks.

Fig. 3.5 shows normalized gross output values for all three simulation scenarios, both when lockdown continues (solid lines) and when the lockdown is lifted for all industries after six weeks (dashed lines). It becomes clear that demand shocks lead to much smaller economic impacts than supply shocks (blue vs. red solid lines). On the other hand, we find that the economy recovers much quicker after the lockdown if there are no demand shocks, as indicated by the large positive slope of the blue dashed line. When there are only demand shocks, recovery is slow. In this case unwinding the economy from lockdown brings limited positive effects due to persistence in exports and investment shocks, sluggish consumption adaptation and a portion of consumers believing in a L-shaped recovery.

Strikingly, during lockdown we observe for an extended period larger negative economic impacts if the model economy faces only supply shocks instead of being exposed to supply and demand shocks simultaneously (blue vs. black solid line). Why is it that turning off demand shocks leads to larger adverse overall impacts? The reason is that in case of large supply constraints and no reduction in final demand, there is higher competition for relatively few goods. If producers cannot satisfy aggregate demand, they need to ration their output to customers (recall that we use proportional rationing). In case of large aggregate demand every customer receives only a relatively small share, which could be even less for some industries compared to the scenario where demand shocks are turned on. If these goods are critical inputs, production in concerned industries will come down once inventories of these inputs are run down. Thus, removing demand shocks can increase the risk of input bottlenecks in production. Put simply, decreasing final demand for certain types of goods ensures their continued supply to other intermediate industries, which might rely critically on the provision of these goods.

In the particular case considered here, it turns out that several large industries have to reduce production as a consequence of fierce competition for critical inputs, as can be seen in left panel of Fig. 3.6. Without demand shocks industries, such as Publishing (J58), Education (P85), and Insurance (K65), produce around 25% less when compared to the baseline case six weeks after lockdown. Although several other industries such as Manufacturing Chemicals (C20) and Manufacturing Vehicles (C29) produce substantially more without demand shocks, this does not offset the overall adverse effect on the economy as a whole.

To better understand how this happens we zoom into two illustrative industries. In the right panel of Fig. 3.6 we show how production constraints in Telecommunications (J62) lead to larger input bottlenecks in IT (J62-3) in the absence of initial demand shocks. The red and blue solid lines show normalized output values of the IT sector for the baseline and the no demand shock scenario, respectively. At the beginning of the lockdown production of the IT sector is less adversely affected in absence of demand shocks. As indicated with the dashed lines, however, Telecommunication input inventories are faster depleted in this case. This is because the Telecommunication sector faces larger aggregate demand (blue vs. red crosses), and due to rationing, delivers only a smaller amount of output to the IT sector. Telecommunication is a critical input to the IT sector and thus, IT output experiences a substantial decline in production once its inventories are run down.

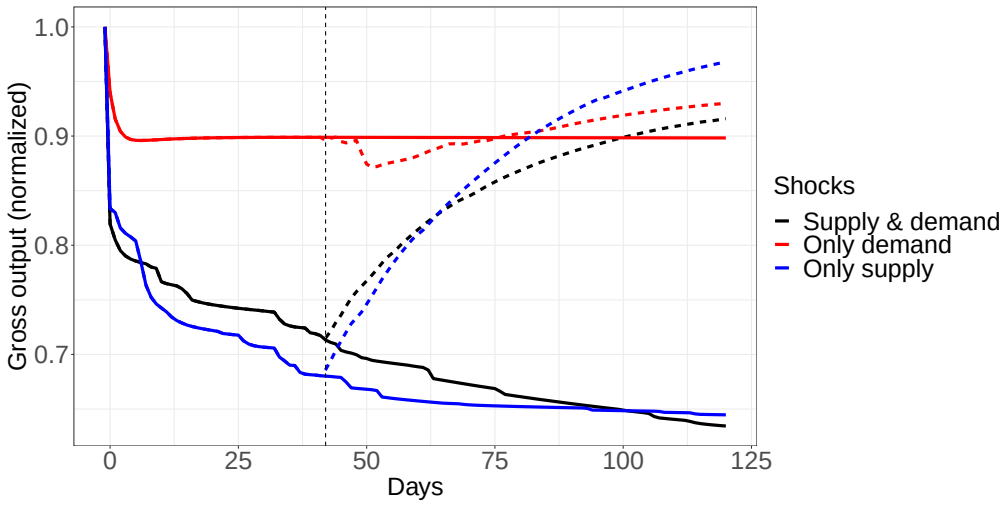

Figure 3.5: **Economic impacts of supply, demand and combined effects.** The black line denotes the model default setup where both supply and demand shocks are applied. The red/blue line shows the case where only demand/supply shocks are switched on. The figure shows results for an indefinite lockdown (shocks are not removed, solid lines to the right of the vertical dashed line) and the case where the lockdown is relaxed after six weeks at $t = 42$ (dashed lines showing an economic recovery).

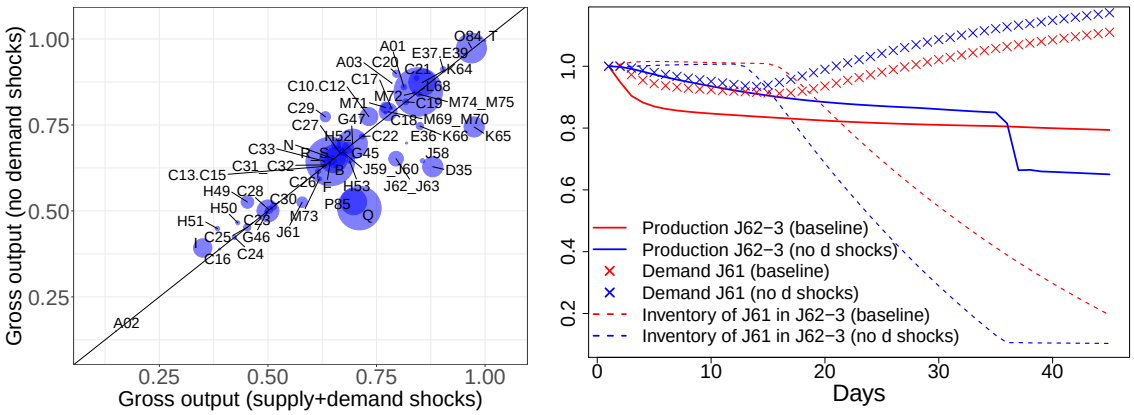

Figure 3.6: **How can production decrease when demand shocks are removed?** *Left panel:* Comparing gross output (normalized with pre-pandemic levels) at the sectoral level six weeks after lockdown when supply and demand shocks are turned on (x-axis) and with demand shocks turned off (y-axis). A few large sectors achieve substantially less production if there are no initial demand shocks in the economy. *Right panel:* A sectoral example of downstream shock propagation of the "Baseline" and "Only supply shocks" scenarios in Fig. 3.5. Sector IT (ISIC J62-3) produces less after around 35 time steps if there are no demand shocks in the economy (blue solid line) compared to both supply and demand shocks being present (red solid), since it runs out of critical input J61, Telecommunications, inventories (blue vs. red dashed line). If there are no initial demand shocks, J61 faces higher aggregate demand (blue vs. red crosses). Due to higher demand for J61 goods and constraint production of J61 goods, IT receives less J61 if there are no demand shocks in the economy.

The case of Telecommunication production and the coupled output of the IT sector exemplifies the complexity of shock propagation through production networks. We identify similar mechanisms for other industries, but the exact dynamics depend on several specificities such as the availability of critical inputs and shock heterogeneity. This analysis highlights several interesting features that frequently are not considered in most macroeconomic analysis, even if they incorporate industry-specific effects.

First, the specification of production functions and input criticality plays an important role. Most economic analyses use some form of CES production functions with non-zero substitution coefficients. Under this approach, while it is in principle possible to construct elaborated CES nests where different degrees of substitutability are allowed between different inputs of a given sector (Baqaee & Farhi 2020), in practice it is so hard to calibrate the parameters that only one (Barrot & Sauvagnat 2016) or a few (Baqaee & Farhi 2020, Bonadio et al. 2020) parameters are specified. The survey considered here (Appendix C.3) instead introduces a distinction between critical and non-critical inputs, for each separate industry, allowing us to keep the Leontief assumption of a strong lack of substitutability for critical inputs, which is arguably a key feature of short-run dynamics after large shocks, while at the same

time not allowing some non-critical inputs to prevent production. This is a step toward more realism, as in exceptional circumstances like a pandemic, we believe that it is likely that firms can still operate even if several inputs that they usually use are not available. Of course, this is admittedly imperfect and could be improved, and we have made the strong assumption that the lack of use a non-critical input simply does not decrease production and translate into higher profits. Assuming a drop in productivity in this case would change the quantitative results, but would not, however, fundamentally change the dynamics.

Second, the size of inventories held by industries is crucial. Similar to equity buffers in financial distress models, inventories act as buffers against production shocks originating upstream and propagating downstream. Detailed information on input-specific inventories on industry and firm levels, as well as on behaviour and inventory management rules, could vastly improve our understanding of shock propagation in production networks.

Third, dynamics really matter over the short time horizons relevant for the pandemic lockdown and its immediate aftermath. It can take days to weeks for shocks to cascade through several layers of a large production network and our simulations suggest that it can take months until the dynamics reach a steady state. The propagation of shocks is path-dependent. This is due to the fact that different industries can have very different customers, resulting in heterogeneous contagion dynamics that depend on "who gets hit first". General equilibrium models and most input-output models implicitly assume zero adjustment time and compare pre- and post-shock equilibrium states of the economy to quantify overall impacts. Our analysis suggests that the shock propagation dynamics play an important role in the short time horizons of pandemic lockdowns. In other words, equilibrium comparative static is warranted only when adjustment is faster than the arrival of new shock (Ando et al. 1963). This is not the case during the pandemic, where the lifting of the lockdown happens before the system has had a chance to reach the lockdown steady-state.

3.5.4 Re-opening a network economy

How to effectively unwind social distancing measures without letting the pandemic situation get out of control was one of the key questions during the first pandemic wave. In an early version of this work we have investigated the trade-off between epidemic spreading and economic output due to re-opening certain industries (Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2020). Here, we focus on the economic consequences of re-opening single industries. We do this by studying the following

simulations: As before, we represent the economy in lockdown by initialising the model as usual with first-order shocks. We then consider two scenarios. First, the re-opening scenario where lockdown is relieved after six weeks for a given set of industries, i.e. for those industries we set $\epsilon^S_{i,t} = 0$ and demand adjusts as discussed in Chapter 3.3. Second, the lockdown scenario, where the lockdown continues and no shocks are removed. We then compare the two scenarios to quantify the boost in economic activity of re-opening a given industry compared to the lockdown.

Fig. 3.7 summarises our findings. Each panel shows total production normalized by pre-shock output on the y-axes for both scenarios (re-open sectors in red, continued lockdown in black). The x-axes shows the number of days where day zero is when the lockdown is lifted in the re-opening scenario. Thus, the economy was already six weeks in lockdown before day zero which is not shown since production is identical for both scenarios during that period. Panel columns represent simulation results for different production function specifications. Panel rows indicate the industries which are re-opened if the lockdown is relaxed.

The economic boost of re-opening varies largely between different sectors and also depends strongly on the production mode assumed. Let us first consider re-opening only the upstream primary sectors Agriculture and Mining, while keeping the lockdown otherwise unchanged (top row of panels). Under the assumption of Leontief production (left column) this leads to a dramatic increase in economic output. Note that primary sectors only account for 2% of UK's total economic output. Opening primary sectors has much smaller effects when using the baseline IHS2 production function (center column), where inputs are only partially critical, and the linear production function (right column), where inputs are not critical at all.

When re-opening the much larger manufacturing sectors (15% of total output), we obtain a completely different impact on economic output (second row). Strikingly, we observe almost no positive effects on economic output under a Leontief production function. The reason is similar to why smaller aggregate shocks can lead to larger overall impacts as discussed in Fig. 3.5. Manufacturing is a large sector relying on many inputs which are critical inputs for other sectors too. Production constraints in other industries might render it impossible to provide larger amounts of those inputs. If manufacturing sectors are re-opened, competition for those scarce inputs increases, resulting in less intermediate consumption for non-manufacturing industries which might face input bottlenecks as a consequence. We do not observe this for the alternative production function setups which relax the strong Leontief assumption.

Here, economic output lies several percentage points above the continued lockdown scenario.

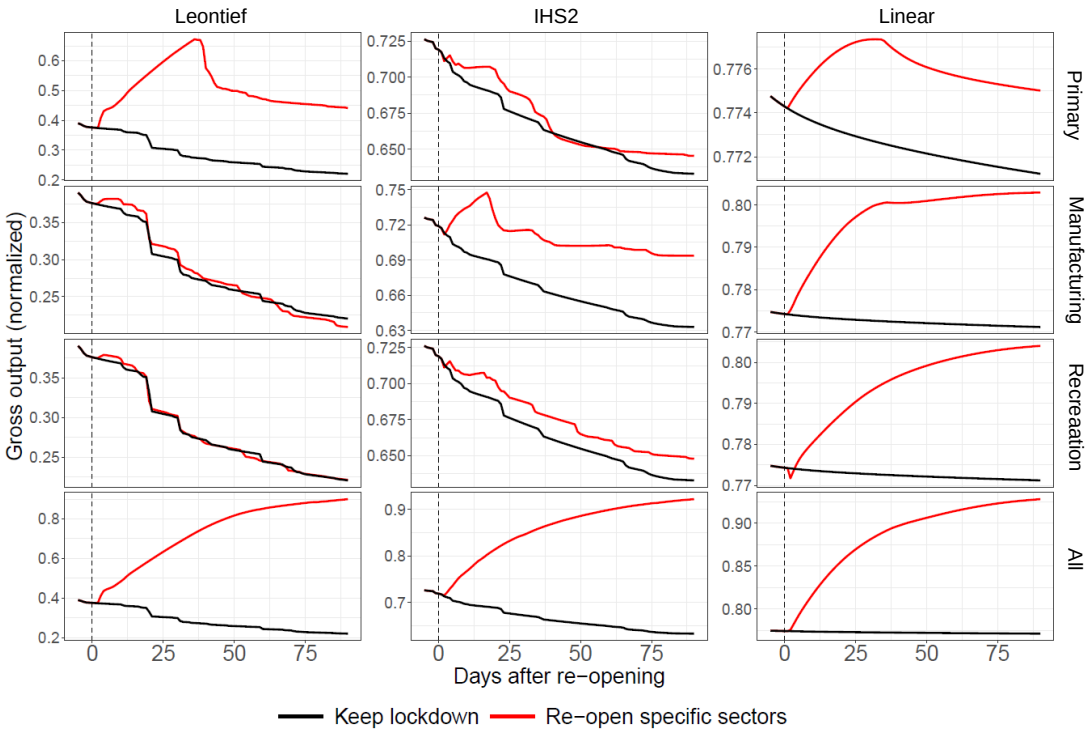

Figure 3.7: **Economic performance of indefinite lockdown and re-opening single sectors for different production functions.** Black lines indicate total output if the lockdown continues and red lines if only one given sector is re-opened at $t = 0$. Columns represent different production functions, rows denote the sectors which are re-opened. The left column show the effect of re-opening for different sectors in a Leontief economy. The center and right columns show the same for the basline IHS2 and linear production functions, respectively. Row 1 shows economic effects of re-opening only primary sectors (ISIC A and B). Row 2 shows the same for opening Manufacturing (ISIC C), row 3 for "Recreation" (ISIC I, R, S) and row 4 for opening all sectors simultaneously. Note that the y-axes scales vary between panels.

We find again very different results when reopening Other Services and Food and Accommodation, here for brevity called Recreation (third row). These sectors are large (6% of total output) and heavily affected by the lockdown since they include theaters, hotels, restaurants and other social activities. It is interesting that opening these industries has no impact on overall economic production when assuming a Leontief production function. This is because these are highly downstream industries and their economic output is of little significance for the intermediate consumption of other industries. Thus, opening recreational sectors has mostly demand-side effects, and given the capacity constraints of upstream sectors this extra demand cannot be satisfied, resulting in no change in overall production. We find positive effects of

opening recreational sectors under the alternative production function assumptions where total production increases slightly above lockdown levels.

The bottom row of panels compares the restart of the economy when the lockdown is lifted for all industries simultaneously. The largest economic boost is given in the Leontief economy which re-starts at very low levels after lockdown. The recovery paths are similar for our baseline and the linear production function, although there is differences in the actual magnitudes. Note that recovery is not instantaneous, but takes a considerable amount of time. A month after re-opening, the economy still operates well below initial production levels.

3.5.5 Up- and downstream propagation of economic shocks

In the standard Cobb-Douglas Equilibrium IO model, productivity shocks propagate downstream and demand shocks propagate upstream (Carvalho & Tahbaz-Salehi 2019). Thus, the elasticity of aggregate output to a shock to one sector depends on the type of shock, and on the position of an industry in the input-output network. What can we say about these questions in our model? Are there properties of an industry that could be computed ex-ante to know how systemic it is? Is this different for supply and demand shocks?

To answer these questions, we run the model with a single shock – either supply or demand – to a single industry. All the other industries do not experience any shocks but have initial productive capacities and face initial levels of final demand. We then let the economy evolve under this specific setting for one month.[18] We repeat this procedure for every industry, every production function specification and different shock magnitudes. We then investigate if the decline in total output can be explained by simple measures such as shock magnitude, output multipliers or upstreamness levels which we formally define below. We first explain the supply and demand shock based scenarios in somewhat more detail.

Supply shock scenarios. When considering supply shocks only, we completely switch off any adverse demand effects, i.e. $\epsilon^D_{i,t} = 0$ (which implies $\theta_{i,t} = \theta_{i,0}$ and $\tilde{\epsilon}^D_t = 0$), $\xi_t = 1$ and $f^d_{i,t} = f^d_{i,0}$ for all i and t. We also set all supply shocks equal to zero $\epsilon^S_{i,t} = 0$, except for a single industry j which experiences a supply shock from the

[18] We also did the analysis with model simulations up to two months after the initial shock is applied. Since results are similar for the two cases, we only report results for the one month simulations.

set $\epsilon^S_{j,t} \in \{0.1, 0.2, ..., 1\}$. We then loop over every possible j. We do this for each of the different production function assumptions.

Demand shock scenarios. In our demand shock scenarios there are no supply shocks ($\epsilon^S_{i,t} = 0 \; \forall \; i, t$) and similarly there is no demand shocks for all but one industry j ($\epsilon^D_{i,t} = 0$, $f^d_{i,t} = f^d_{i,0} \; \forall i \neq j, \forall t$). For the single industry j we again let shocks vary between 10 and 100%; $\epsilon^D_{j,t} \in \{0.1, 0.2, ..., 1\}$. For simplicity we assume uniform shocks across all final demand categories of the given industry, resulting in $f^d_{j,t} = (1-\epsilon^D_{j,t}) f^d_{j,0}$. To keep things as simple as possible, we further assume that there is no fear of unemployment $\xi_t = 1$ and that final consumers do not switch to alternative products at all ($\Delta s = 1$). Under these assumptions the values for $\theta_{i,t}$ and $\tilde{\epsilon}^D_t$ are then computed as outlined in Chapter 3.3.2.

Fig.3.8 shows the simulation results broken down into the various shock magnitudes (vertical axis) and production function categories (horizontal axis). The coloring and the value of a tile represent the average aggregate output (as a fraction of initial output), where the average is taken over all N runs. Results obtained from the demand shock scenarios, Fig. 3.8(b), do not differ across alternative production function specifications and thus we only show one representative case. The reason for this is that demand shocks to an industry are passed on proportionally to all of its suppliers as shown in Eq. (3.2), regardless of the underlying production mechanisms. This is in stark contrast to the supply shock scenarios, Fig. 3.8(a), where economic impacts depend strongly on the choice of production function. While the economic impact of supply constraints is limited in the linear production model and depends fairly smoothly on the shock magnitude, this is not the case in the Leontief model. Here the economy experiences a major truncation if single industries are shut down. The results lie in-between these two extremes when production functions differentiate between critical and non-critical inputs (IHS 1-3).

Except for the linear production function simulations, we observe for several severe supply shock scenarios fairly wide distributions of economic impacts. Thus, not only the shock size matters but also *which* industries are affected by these shocks. For example, applying an 80% supply shock shock to industry Repair-Installation (C33) under the IHS2 assumption, collapses the economy by about 50%, although the industry accounts for less than half a percent of the overall economy. On the other hand, applying the same shock to the comparatively large industry Other Services (R_S, 3% of the economy) leads to a mere 6% reduction of aggregate output.

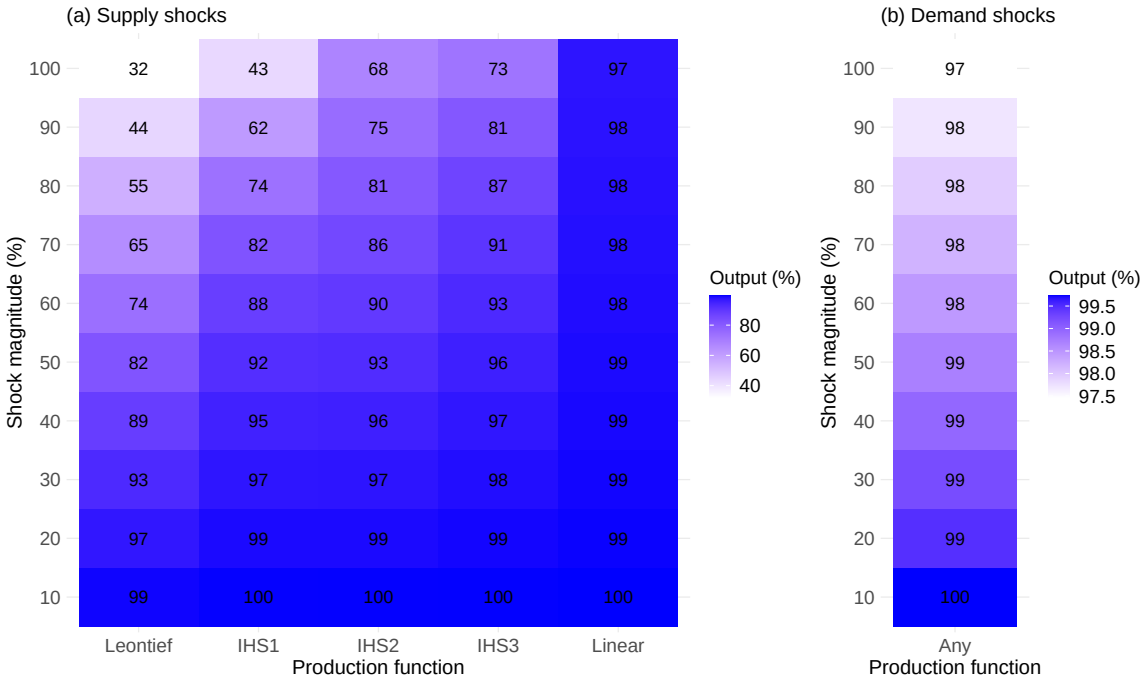

Figure 3.8: **Aggregate gross output as percentage of pre-shock levels after shocking single industries.** A column depicts different production functions and rows distinguish the supply (a) and demand (b) shock magnitude which an industry is exposed to. Results in (b) are only shown for one production function, since they are identical across alternative specifications. The values in the tiles and their coloring denote aggregate output levels as percentage of pre-shock levels one month after the shock hits a given industry. This values are computed as averages from N runs, always shocking only a single industry.

For a policymaker it is important to know what properties of an industry drive the amplification of shocks, as this could inform both the design of lockdown measures as well as reopening policies. To explore this more systematically, we regress output levels against potential explanatory factors such as upstreamness, output multipliers and industry sizes. An industry's upstreamness in a production network is its average distance to the final consumer (Antràs et al. 2012) and also known as Total Forward Linkages (Miller & Temurshoev 2017). High upstreamness implies that the output of this industry requires several subsequent production steps before it is purchased by final consumers. Thus, relaxing shocks on industries with high upstreamness could potentially stimulate further economic activity. Since upstreamness boils down to the row sums of the Gosh inverse (Miller & Temurshoev 2017), we obtain the N-dimensional upstreamness vector as $\boldsymbol{u} = (\mathbb{I} - \mathbf{B})^{-1}\mathbf{1}$, where a matrix element $B_{ij} = Z_{ij,0}/x_{i,0}$ represents "allocation coefficients". Upstreamness ranges from 1.004 (Household activities) to 2.742 (Warehousing) in our sample of UK industries.

Output multipliers, or alternatively Total Backward Linkages, are a core metric in many economic studies. In input-output analysis multipliers quantify the impact of a

change in final demand in a given sector on the entire economy. Multipliers are related to various network centrality concepts and have been shown to be strongly predictive of long-term growth (McNerney et al. 2018). Since shocking an industry with a high multiplier should lead to larger decreases in intermediate demand, it is plausible that high-multiplier industries tend to amplify shocks more. The output multiplier is computed as the column sum of the Leontief inverse, i.e. $\boldsymbol{m} = (\mathbb{I} - \mathbf{A}^\top)^{-1}\mathbf{1}$. Multipliers range from 1 (Household activities) to 2.379 (Forestry) in our sample. Upstreamness and multipliers are different, but are fairly highly correlated, with a Pearson correlation of 0.45.

We then regress log aggregate output one month after the initial shock hits the economy, $\log(\sum_k x_{k,30})$, against the shocked industry's log upstreamness and multiplier levels. Naturally, we would expect a larger decline if supply shocks hit an overall large industry and similarly for demand shocks and the size of final consumption. Thus we also control for total industry size measured in log gross output, $\log(x_{i,0})$, and industry-specific total demand, $\log(c^d_{i,0} + f^d_{i,0})$, in our regressions.[19] We then run the regression for every given shock magnitude and production function separately.

Table 3.3 summarizes the regression results for the supply shock scenarios. For supply shocks we find that upstreamness is a very good predictor of adverse economic impacts if the economy builds upon Leontief production mechanisms. If industries use linear production technologies, on the other hand, it is rather the size of the industry than its upstreamness level that explains reductions in aggregate output. When using the intermediate assumption of an IHS2 production function, both upstreamness levels and industry size significantly affect aggregate impacts, although the overall model fit (R^2) drops substantially.

Note that supply shocks are to a large extent a policy variable as they are directly coupled to non-pharmaceutical interventions such as industry-specific shutdowns. Our results indicate that upstreamness levels of industries are an important aspect for designing lockdown scenarios. In case of limited production flexibilities, upstreamness might be even a better indicator of industry closure related aggregate impacts than the actual size of the industry.

Regression results for the demand shock experiments are shown in Table 3.4. The first four columns show the results from univariate regressions where we include only one of the potential covariates: upstreamness, multipliers, final consumption and

[19] We do not include gross output and final demand values together as regressors to avoid multicolinearity. Industry gross output and final consumption are highly correlated; $\text{cor}\{\log(x_{i,0}), \log(c_{i,0} + f_{i,0})\} = 0.89$ (p-value $< 10^{-16}$).

gross output. Somewhat surprisingly, output multipliers, a key metric for quantifying aggregate impacts resulting from demand side perturbations in simpler input-output models, do not exhibit any predictive power in our case. Upstreamness, on the other hand, is positively associated with aggregate output values, indicating that demand shocks to upstream industries have less adverse impacts on the economy than downstream industries. Better model fits, however, are obtained when regressing aggregate output against indicators of industry size gross output or final consumption.

Interestingly, combining final consumption values with the network-based metrics in a multivariate regression (column five) does not improve explanatory power at all. On the other hand, upstreamness and output size explain complementary parts of aggregate impacts (column six), resulting in a better model fit than any regression using final consumption values. Thus, industries' upstreamness and output sizes do not only play a role for the propagation of supply shocks but are also key determinants for demand shock amplification.

Overall, we find that static measures are indicative of our model results but only explain them partially, even in this simplified case where we studied either demand or supply to only one industry at time. In reality, supply and demand shocks are mixed and multiple industries affected at the same time to varying degrees. In that case the propagation of pandemic shocks will be more complex. Nevertheless, our results indicate that upstream industries play an important role in the amplification of exogenous shocks.

3.6 Discussion

In this chapter we have investigated how locking down and re-opening the UK economy as a policy response to the Covid-19 pandemic affects economic performance. We introduced a novel economic model specifically designed to address the unique features of the pandemic. The model is industry-specific, incorporating the production network and inventory dynamics. We use survey results by industry experts to model how critical different inputs are in the production of a specific industry.

We have shown that the model can be used for empirical analysis as well as for theoretically exploring shock propagation mechanisms in production networks. We find that industries are affected by direct demand and supply shocks in highly heterogeneous ways. While many manufacturing industries face large supply shocks, transport industries experience mostly demand-side shocks. Other industries including hotels

	Dependent variable: $\log(\sum_i x_{i,30})$					
Shock size:	40%	40%	40%	80%	80%	80%
Production:	Leontief	IHS2	Linear	Leontief	IHS2	Linear
$\log(u_{i,0})$	-0.233^{***}	-0.114^{***}	0.008	-0.440^{***}	-0.383^{***}	0.016
	(0.018)	(0.019)	(0.003)	(0.023)	(0.071)	(0.007)
$\log(m_{i,0})$	0.067	0.065	-0.013	0.037	0.306	-0.025
	(0.038)	(0.041)	(0.007)	(0.049)	(0.150)	(0.014)
$\log(x_{i,0})$	-0.006	-0.019^{***}	-0.008^{***}	-0.011	-0.079^{***}	-0.016^{***}
	(0.004)	(0.004)	(0.001)	(0.005)	(0.016)	(0.001)
Constant	15.512^{***}	15.673^{***}	15.560^{***}	15.224^{***}	16.175^{***}	15.641^{***}
	(0.050)	(0.053)	(0.009)	(0.065)	(0.197)	(0.018)
Observations	55	55	55	55	55	55
Adjusted R^2	0.776	0.456	0.722	0.890	0.448	0.719
Note:						*p<0.01; **p<0.001; ***p<0.0001

Table 3.3: **Regression results for supply shock experiments.**

	Dependent variable: $\log(\sum_i x_{i,30})$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(u_{i,0})$	0.040^{**}				0.012	0.036^{***}
	(0.010)				(0.010)	(0.008)
$\log(m_{i,0})$		0.021			-0.024	-0.030
		(0.024)			(0.019)	(0.018)
$\log(c_{i,0} + f_{i,0})$			-0.012^{***}		-0.011^{***}	
			(0.002)		(0.002)	
$\log(x_{i,0})$				-0.014^{***}		-0.013^{***}
				(0.002)		(0.002)
Constant	15.443^{***}	15.453^{***}	15.588^{***}	15.618^{***}	15.586^{***}	15.599^{***}
	(0.006)	(0.013)	(0.016)	(0.023)	(0.025)	(0.023)
Observations	55	55	55	55	55	55
Adjusted R^2	0.217	-0.004	0.522	0.452	0.522	0.580
Note:						*p<0.01; **p<0.001; ***p<0.0001

Table 3.4: **Regression results for demand shock experiments.** As representative case we show regression results for the scenario of 60% shocks to final consumption, because results are similar for alternative shock sizes. Note that economic impacts are identical across alternative production functions.

and restaurants are substantially exposed to both shock types simultaneously. We find similar industrial heterogeneity for higher-order impacts.

We analysed how shocks propagate through the production network, resulting in non-trivial economic impacts. We have shown that input criticality plays an important role in the downstream amplification of shocks. We also found that inventory levels can act as buffers against production shocks and are crucial for understanding economic impacts. It has become evident that time scales matter as shock propagation is not immediate but takes time.

We find that first-order shocks can be translated into overall impacts in highly nonlinear ways. We even find cases where smaller aggregate shocks can lead to larger economic impacts as a result of unbalanced supply and demand dynamics. This 'coordination failure' suggests that it could be dangerous to re-open single sectors of the economy by themselves without understanding how they are embedded in the production network. Our results suggest that the economic boost from opening an industry depends on the up-/downstream location of that industry as well as how severe the economy suffers from input bottlenecks. In case the economy faces serious productive constraints, re-opening a single sector can even have adverse effects on economic output. These results underline that network measures or simple industry characteristics might be insufficient for quantifying economic impacts resulting from national lockdowns and subsequent re-opening.

Our remarkably accurate real time GDP predictions for the UK economy presented in Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2020) and the detailed "post-mortem" analysis of our results in Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021) indicate that dynamic models of the type that we developed here can do a good job of forecasting disruptions in the economy. We want to emphasize that, while there is wide variation in the results under different scenarios, this variation could be dramatically reduced by collecting better data. A clear example is the choice of inventory levels. In our original model we had no data for inventory levels of UK industries, so we used data for the US. Replacing this by UK data made a substantial improvement in our results. Similarly, the economic data we had available was from 2019, and the IO data was from 2014 – better real time measurements about the response of industries to the pandemic would likely have improved our predictions. A more extensive study of the importance of different inputs to production could have reduced ambiguity about the choice of production function.

At the highest level, our model illustrates the value of its key features. These are: modeling at the sectoral level, allowing both supply and demand shocks to operate

simultaneously, using a realistic production function that properly captures nonlinearities without exaggerating them, and using a dynamic model that incorporates inventory effects. With better data and better measurement of parameters, our results demonstrate that a model of this type can give useful insight into the economics of a disaster such as the Covid-19 pandemic.

Chapter 4

Simultaneous supply and demand constraints in input-output networks: The case of Covid-19 in Germany, Italy, and Spain

This chapter is based on Pichler & Farmer (2021).

4.1 Introduction

An immanent feature of many natural and anthropogenic disasters is that they affect both the supply and the demand side of the economy. In this chapter we study the Covid-19 pandemic as an exemplary case of simultaneous supply and demand shocks. Supply shocks from the pandemic arise from different sources. While deaths and sickness of employees can limit productive capacity, these effects are minor compared to nation-wide lockdown measures imposed by governments to curb the spreading of the virus. During lockdown workers employed in non-essential industries who cannot work remotely are unable to perform their jobs (del Rio-Chanona et al. 2020, Dingel & Neiman 2020, Koren & Pető 2020).

The pandemic also affects demand in heterogeneous ways (Congressional Budget Office 2006, del Rio-Chanona et al. 2020, Chetty et al. 2020, Carvalho et al. 2020, Chen et al. 2021). While we expect demand shocks to be comparatively small for some industries (e.g. manufacturing), other industries are strongly affected by demand shocks. An illustrative example is the transportation industry which is considered as essential in many countries and thus would not experience adverse supply side effects. But the transportation industry faces large demand shocks, since consumers reduce demand for air travel and public transport to avoid infectious exposure.

Since economic agents are embedded in production networks, we expect that the overall economic impact is larger than the initial shocks to supply and demand suggest. Demand shocks will reduce the sales of firms and, by backward propagation, also diminish the sales of their suppliers. Supply shocks, on the other hand, will spread downstream and upstream. Downstream effects materialize if the limited productive capacity of suppliers creates input bottlenecks for customers. Due to lower productive capacities, firms will also require less inputs for production, thus adversely affecting upstream suppliers of these inputs.

Input-output (IO) models are frequently used to model higher-order economic impacts arising from such exogenous shocks. Most of these models account for either supply or demand shocks but do not incorporate them concurrently (Galbusera & Giannopoulos 2018). In this paper we revisit existing IO modeling techniques and introduce a novel dynamic approach, to account for both types of shocks. Our analysis is based on industry-level data but could easily be extended to firm-level analysis that has become the focus of recent studies (Diem et al. 2021, Borsos & Mero 2020, Mandel & Veetil 2020*b*, Inoue & Todo 2019). In particular, our results relate directly to the effects of data aggregation on impact estimates.

Several macroeconomic models that account for the production network structure have been suggested to study the economic impacts of the Covid-19 pandemic. Inoue & Todo (2020) use an agent-based model calibrated to more than a million Japanese firms to model nation-wide economic effects of a lockdown in Tokyo. Using the World Input-Output Database (WIOD), Mandel & Veetil (2020*a*) study the effects of national lockdowns on global GDP in a non-equilibrium framework. Fadinger & Schymik (2020), Baqaee & Farhi (2020), Bonadio et al. (2020) and Barrot et al. (2020) use general equilibrium models to quantify economic impacts of social distancing. All these studies focus on the supply-side shocks of the pandemic and without further considering changes in final consumption. Exceptions are Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2020) and Guan et al. (2020) who incorporate both supply and demand shocks in their production network models.

In contrast to these studies, we consider simpler modeling techniques derived from traditional IO analysis which we extend to account for simultaneous supply and demand shocks. We follow this approach to better isolate underlying mechanisms of shock propagation in production networks that can be hard to disentangle in more sophisticated macroeconomic models. When applying pandemic shocks to data from Germany, Italy and Spain, we find that existing IO modeling techniques yield infeasible solutions of economic impacts. To understand the nature of this problem we

introduce a simple linear programming method that allows us the determine feasible market allocations, representing a lower bound of minimal shock propagation. To find more realistic impact estimates, we further introduce a new bottom-up approach based on rationing outputs. This setup allows us to dynamically model the propagation of shocks in production networks and uncover several theoretical implications, such as the effect of alternative behavioral assumptions on economic impacts.

We intentionally keep our modeling approach simple. As mentioned above, even more complicated models currently applied to assess the economic effects of the pandemic do not incorporate supply and demand shocks simultaneously. The focus of this paper is on the theoretical implications of simultaneous supply and demand constraints for IO-based impact assessment methods. For more realistic empirical impact assessment of natural disasters, we stress that further aspects that we do not consider here could play an important role, e.g. inventory dynamics (Bak et al. 1993, Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2021), price and substitution effects, adaptive behavior and governmental intervention (see Oosterhaven (2017) for a recent discussion).

Our paper contributes to the field of IO models in disaster impact assessment, for which a recent review can be found in Galbusera & Giannopoulos (2018). For recent demand-side and supply-side focused contributions see Klimek et al. (2019) and Avelino & Dall'erba (2019), respectively. Similar to our derivation of a lower bound of shock propagation, Koks & Thissen (2016) and Oosterhaven & Bouwmeester (2016) rely on optimization techniques to account for supply constraints in demand-driven models. Constrained optimization techniques are frequently used in extended IO models (for example Duchin (2005) and Duchin & Levine (2012)). The modeling of alternative rationing algorithms is related to the studies of Steenge & Bočkarjova (2007), Li et al. (2013) and Koks et al. (2015) which allow for imbalanced IO tables immediately after the disaster strikes and evaluate different recovery paths.

Our findings show that shock amplification is much larger in the bottom-up approaches than what best case scenarios would suggest. Micro-level coordination failures can severely exacerbate adverse economic shocks. Rationing assumptions play a key role in impact estimates and alternative behavioral rules can lead to very different aggregate outcomes. These effects strongly interact with the overall shock magnitude as well as the level of connectedness in the production network. While we find that economic impacts increase with higher levels of network density in general, the extent to which this is true strongly depends on the underlying rationing mechanisms. Our

results also imply that estimates of economic impact can be highly sensitive with respect to data quality and aggregation.

4.2 Pandemic shocks to supply and demand

Our analysis is based on the most recent year (2014) of the World-Input-Output-Database (WIOD) (Timmer et al. 2015).[1] We use estimates of supply and demand shocks from the literature to determine maximum final consumption and production values for 54 industries in three major European economies, i.e. Germany, Italy and Spain.[2] We focus on these economies because existing research provides us with lists of essential industries and thus allows us to make reasonable estimates of supply shocks. For this reason, we only allow for domestic supply shocks in our analysis, despite acknowledging the importance of international supply chain disruptions.

We follow the approach of del Rio-Chanona et al. (2020) to compute supply shocks for every industry during a lockdown. A supply shock is caused by the removal of labor in non-essential industries due to social distancing measures. In contrast, an industry which is defined as essential will not be affected. Even if employed in non-essential industries, workers who can accomplish their work from home will do so and are assumed to keep pre-lockdown productivity levels. The Remote Labor Index RLI_i indicates the share of an industry's labor force that can work from home. The supply shock to industry i is then computed as $\epsilon_i^S = (1 - \mathrm{RLI}_i)(1 - \mathrm{e}_i)$, where e_i is the share of an industry defined as essential. Thus, the supply shock of an industry during lockdown is the share of labor in non-essential industries that cannot work from home. We assume that the supply shock to an industry determines the maximum production of an industry, i.e.

$$x_i^{\max} = (1 - \epsilon_i^S) x_{i,0}, \tag{4.1}$$

where $x_{i,0}$ denotes the production in industry i before lockdown. Following this approach, every industry faces binding supply side constraints, limiting its production to $x_i \in [0, x_i^{\max}]$. The Remote Labor Index is taken from del Rio-Chanona et al. (2020) and industry-specific essential ratings for Germany, Italy and Spain are obtained from Fana et al. (2020).

[1] All code and data to reproduce this chapter are made accessible online: `https://www.doi.org/10.5281/zenodo.4326815`.

[2] We removed industries T (Household activities) and U (Extraterritorial activities) from our analysis since they are not connected to other industries in the data and thus don't play any role in the propagation of shocks in the production network.

To determine first-order demand shocks to every industry, ϵ_i^D, we use estimates of a prospective Congressional Budget Office (CBO) study aiming at quantifying the demand-side impact of a pandemic (Congressional Budget Office 2006). Demand and supply shock data have been mapped to WIOD industry categories by Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021) and are presented in detail in Appendix D.1. As for supply shocks, we assume that demand shocks determine maximum final consumption values according to

$$f_i^{\max} = (1 - \epsilon_i^D) f_{i,0}, \tag{4.2}$$

where $f_{i,0}$ represents pre-lockdown final consumption. Since consumption cannot be negative, we have $f_i \in [0, f_i^{\max}]$.[3] Note that WIOD distinguishes different final consumption categories.[4] We apply the CBO estimates to final consumption of households and non-profit organizations and assume a 10% shock to investments and exports as in Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021). Due to our methodological focus we do not further distinguish between different final consumption categories but for simplicity only consider total final consumption values for every industry. It is important to point out that this is a major simplification and we expect the various types of final consumption to be highly relevant in empirical assessments. Table 4.1 shows country aggregates of essentialness scores, Remote Labor Index, supply and demand shocks.

	x	f	ϵ^S	ϵ^D	RLI	e
Germany	7,066.74	4,447.11	0.31	0.09	0.42	0.49
Italy	4,075.40	2,343.12	0.27	0.11	0.41	0.55
Spain	2,567.91	1,552.88	0.33	0.13	0.40	0.44

Table 4.1: **Country aggregates of gross output, final consumption and supply and demand shock inputs.** Columns $x = \sum_i x_{i,0}$ and $f = \sum_i f_{i,0}$ are country gross output and final consumption in billion USD based on 2014 values. $\epsilon^S = 1 - \sum_i x_i^{\max}/x$ and $\epsilon^D = 1 - \sum_i f_i^{\max}/f$ are total supply and demand shocks, respectively. $\text{RLI} = \sum_i \text{RLI}_i x_{i,0}/x$ and $e = \sum_i e_i x_{i,0}/x$ denote the country-wide Remote Labor Index and essentialness score on which supply shocks are based. In all three countries total direct supply shocks are larger than demand shock.

[3] In principle, f_i could be negative if there is extremely large inventory depletion. This is not the case in our data. In national accounts inventory adjustment merely represents a variable to rebalance row and column sums of IO tables. We therefore do not consider the possibility of negative final demand.

[4] Given the multi-regional nature of WIOD, imports from and exports to other industries are accounted for in the intermediate consumption matrix. Since we treat international trade as exogenous, we aggregate all intermediate exports to a single final consumption category, as is usually the case with national IO tables. Similarly, we treat intermediate imports as an exogenous input that does not inhibit production, or equivalently, that is always available if demanded.

4.3 IO framework

We first introduce basic national accounting identities that build the basis for most IO models. Let us consider an economy consisting of N industries, each producing a unique, industry-specific good. Inter-industry purchases and sales are encoded in the intermediate consumption matrix $\mathbf{Z}$ where an element Z_{ij} denotes the total monetary value of goods produced by industry i that are consumed by industry j. For the N-dimensional vectors of gross output, total final consumption and value added, we write $\boldsymbol{x}$, $\boldsymbol{f}$ and $\boldsymbol{v}$, respectively. In this economy the following identities hold

$$\boldsymbol{x} = \mathbf{Z}\mathbf{1} + \boldsymbol{f} = \mathbf{Z}^\top \mathbf{1} + \boldsymbol{v}, \tag{4.3}$$

where $\mathbf{1}$ is a N-dimensional column vector of ones.

A core assumption in a Leontief-inspired modeling framework is that industries produce based on fixed production recipes, allowing us to rewrite the first identity of Eq. (4.3) as

$$\boldsymbol{x} = \mathbf{A}\boldsymbol{x} + \boldsymbol{f} = \mathbf{L}\boldsymbol{f}, \tag{4.4}$$

where $\mathbf{A}$ is the technical coefficient matrix with elements $A_{ij} \equiv Z_{ij}/x_j$ (the production recipe) and $\mathbf{L}$ the Leontief inverse $\mathbf{L} \equiv (\mathbb{I} - \mathbf{A})^{-1}$. A conventional assumption is that gross output is determined endogenously, whereas final consumption is taken to be as exogenous. It then follows that value added is obtained as a residual variable.

Eq. (4.4) is at the core of many IO models that are frequently used for disaster impact analysis. This specification defines a demand-driven model. However, as we have already stressed, many economic shocks actually act on the supply side of the economy. For example, natural catastrophes such as earthquakes or floods destroy physical capital, putting upper limits on an industry's production in the disaster aftermath. Eq. (4.4), however, neglects supply capacity constraints and implies an infinite elasticity of supply with respect to demand. This is particularly problematic given the frequent short-term focus of IO studies.

We should point out that there are also supply-driven IO models building upon Ghosh's model (Ghosh 1958) that assumes exogenous primary factors and derives final consumption values endogenously. Ghosh's model does not comply with a Leontief production function and has been heavily criticized amongst other aspects for the assumption of perfect demand elasticities and perfect substitution of inputs (Oosterhaven 1988, Gruver 1989, De Mesnard 2009). The IO inoperability model (Haimes & Jiang 2001, Santos & Haimes 2004) is another notable supply-side focused model.

Dietzenbacher & Miller (2015) show, however, that it closely corresponds to conventional IO modeling approaches.

Neither the demand- nor the supply-driven specification of IO models are able to incorporate supply and demand constraints at the same time. A potential remedy is the mixed endogenous/exogenous model (MEEM), which has been applied in several studies including Steinback (2004), Kerschner & Hubacek (2009) and Arto et al. (2015). The MEEM acknowledges that not only final demand is constrained but that for some industries supply constraints are more severe and therefore binding. A difficulty in empirical analysis is to define which industries are supply and which are demand constrained. In some cases there might be "natural" distinctions. For example, Arto et al. (2015) study global supply chain disruption effects of the 2011 Tōhoku earthquake by solving the MEEM where only the Japanese transport equipment industry is supply constrained, whereas for all other industries gross output is endogenous. In case of simultaneous (severe) supply and demand shocks the line of distinction will be more blurred.

Note that the MEEM incorporates both supply and demand shocks, but not *simultaneously* for a single industry. Instead, an industry is either supply or demand constrained. As a consequence, solutions to the MEEM might not comply with exogenous industry-level constraints to supply and demand, i.e. they can be economically *infeasible*. As is shown in detail in Appendix D.2, this is exactly what happens if we apply the MEEM to the pandemic shocks discussed in the previous section. For all three countries, the MEEM yields final consumption values that are either negative or larger than the exogenous constraints.

4.4 Propagation of simultaneous supply and demand shocks

The industry-specific effects to supply and demand during the Covid-19 pandemic and the difficulty of incorporating them in the existing models motivated us to explore possible extensions of the IO modeling framework. Rather than using more complicated models that can deal with simultaneous supply and demand shocks, such as computable general equilibrium (CGE) models or those discussed in the introduction, we intentionally keep the modeling approach simple. We remain close to the Leontief framework, the "work horse" of many more advanced IO-based models. This allows us to better isolate the basic mechanisms of shock propagation that are easily conflated with other effects in more sophisticated models.

To determine lower bounds of shock propagation with respect to output and consumption, we study the idealized case of a social planner who allocates goods to maximize either total gross output or total final consumption within the given economic constraints. This will give us a best case scenario of negative economic impacts, i.e. the minimal decrease in total output and final consumption necessary to arrive within the set of feasible solutions given the exogenous constraints to supply and demand. Of course, alternative objectives could be optimized as well (e.g. employment).

While deriving a lower bound is valuable, it is unlikely to be realistic. We therefore consider a second approach by comparing alternative rationing algorithms. If an industry is supply constrained, it will not be able to satisfy all of its demand and thus needs to make a decision which customer to serve and to what extent. We implement several decision rules and investigate how this choice influences the estimated economic impact. In contrast to the optimization method or the MEEM, which simply compute an equilibrium, this approach explicitly computes the transient dynamics that lead to a new equilibrium.

4.4.1 Feasible market allocations

Given exogenous constraints to supply and demand, what is the feasible market allocation that maximizes final consumption and/or total output? The solution needs to lie within exogenous bounds on supply and demand and also needs to satisfy the assumption of Leontief production, Eq. (4.4). We seek market allocations $\{\boldsymbol{x}^*, \boldsymbol{f}^*\}$ that (a) respect given production recipes $\boldsymbol{x}^* = \mathbf{L}\boldsymbol{f}^*$ and (b) satisfy basic output and demand constraints $\boldsymbol{x}^* \in [\mathbf{0}, \boldsymbol{x}^{\max}]$ and $\boldsymbol{f}^* \in [\mathbf{0}, \boldsymbol{f}^{\max}]$.

We follow a mathematical optimization procedure to map out the solution space of feasible market allocations. As a first case, we determine the market allocation that maximizes gross output under the assumptions specified. Large levels of gross output indicate high economic activity, which in turn entail high levels of primary factors such as labor compensation. As a second case we look at market allocations that maximize final consumption given current production capacities. Due to the linearity of the Leontief framework, the problems boil down to linear programming exercises.

Maximizing gross output:

$$\begin{aligned} \max_{\boldsymbol{f} \in [\mathbf{0}, \boldsymbol{f}^{\max}]} \quad & \mathbf{1}^\top (\mathbb{I} - \mathbf{A})^{-1} \boldsymbol{f}, \\ \text{subject to} \quad & (\mathbb{I} - \mathbf{A})^{-1} \boldsymbol{f} \in [\mathbf{0}, \boldsymbol{x}^{\max}]. \end{aligned} \tag{4.5}$$

Maximizing final consumption:

$$\begin{aligned} \max_{\boldsymbol{x}\in[\mathbf{0},\boldsymbol{x}^{\max}]} \quad & \mathbf{1}^\top(\mathbb{I}-\mathbf{A})\boldsymbol{x}, \\ \text{subject to} \quad & (\mathbb{I}-\mathbf{A})\boldsymbol{x} \in [\mathbf{0}, \boldsymbol{f}^{\max}]. \end{aligned} \tag{4.6}$$

To maximize gross output of the economy, $\sum_i x_i^*$, requires us to the find the vector of final consumption $\boldsymbol{f}^*$. The constraint $\mathbf{L}\boldsymbol{f} \in [\mathbf{0}, \boldsymbol{x}^{\max}]$ ensures that industry output levels lie within the respective production capacities. The problem is similar when maximizing final consumption where a vector of output levels x^* is chosen to maximize final consumption, $\sum_i f_i^*$. The auxiliary constraint enforces that final consumption levels do not exceed given demand. The optimization problem always admits a solution since the trivial allocation of a full collapse $\{\boldsymbol{x}^*, \boldsymbol{f}^*\} = \{\mathbf{0}, \mathbf{0}\}$ always exists, although we expect positive values for realistic input data.

Note that the market allocations are "optimal" in a mathematical sense. That is, they represent maximum values of output and consumption given the exogenous constraints to supply and demand. Or stated differently, they determine the minimum level of shock propagation measured in aggregate final consumption and gross output. Thus, this optimization method is not intended to give realistic or necessarily desirable economic outcomes, but rather probes the system boundaries by mapping out the solution space. Any feasible solution, under the above-specified assumptions, has to lie within this space.

4.4.2 Input bottlenecks and rationing variations

As our second method we implement different rationing schemes for output constrained industries. In contrast to the optimization methods, this represents a bottom-up approach for finding feasible market allocations. Industries place orders to their suppliers based on incoming demand. Since suppliers can be output constrained, they might not be able to satisfy demand fully. A supplier therefore needs to make a decision about how much of each customer's demand it serves. Intermediate consumers transform inputs to outputs based on fixed production recipes. Thus, if a customer receives less inputs than she asked for, she faces an input bottleneck further constraining her production. As a consequence, the customer reduces her demand for other inputs as they are not further needed under limited productive capacities. We iterate this procedure forward until the algorithm converges.

We run this algorithm with four alternative rationing rules: (a) strict proportional rationing, (b) proportional rationing to intermediate demand but priority of intermediate over final demand, (c) priority rationing serving largest customers first and (d)

random rationing, where customers are served based on a random order. We then compare the results obtained from the four competing behavioral rules.

These rationing approaches are frequently applied in the literature, although there are differences in the exact specifications. In general it is hard to calibrate how firms distribute output in case of supply constraints to empirical data and so the rationing choice is often ad-hoc. Here, we apply all four rules within a consistent dynamic framework to better understand how alternative behavioral assumptions affect impact estimates. Strict proportional rationing is frequently assumed in the literature (Henriet et al. 2012, Hallegatte 2014, Guan et al. 2020, Mandel & Veetil 2020*a*, Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2020) and also mixed proportional/priority rationing has been considered (Battiston et al. 2007, Hallegatte 2008, Li et al. 2013, Diem et al. 2021). The agent-based framework of Inoue & Todo (2019) and Inoue & Todo (2020) assumes a variation of priority rationing while the random matching of suppliers and customers in the agent-based model proposed by Poledna et al. (2018) and Poledna et al. (2019) most closely resembles a random rationing approach.

(a) Strict proportional rationing. If industries are unable to satisfy total incoming demand completely, they distribute output proportional to their customers' demand, where no distinction is made between intermediate and final customers. More specifically, if an industry's output, x_i, is smaller than incoming demand, d_i, it will supply $Z_{ij} = O_{ij}\frac{x_i}{d_i}$ to customer j, where O_{ij} denotes the demand from customer j to industry i. We implement the rationing algorithm in the following way: First, industries determine their total demand as if there were no supply-side constraints, i.e. $\boldsymbol{d} = \mathbf{L}\boldsymbol{f}^{\max}$. Industries then evaluate if they are able to satisfy demand given their constrained production capacities. If an industry i can satisfy demand only partially, it will create a bottleneck of size $r_i = \frac{x_i^{\max}}{d_i}$ to other industries due to proportional rationing. Since industries produce according to fixed Leontief input recipes, their largest input bottleneck, $s_i = \min_{j:A_{ji}>0}\{r_j, 1\}$ will be the binding constraint in production. Thus, in case of input bottlenecks where $s_i < 1$, production of i reduces to $x_i = \min_{j:A_{ji}>0}\left\{x_i^{\max}, \frac{s_i A_{ji} d_i}{A_{ji}}\right\} = \min\{x_i^{\max}, s_i d_i\} < d_i$. This in turn reduces the amount of goods delivered to the final consumer $f_i = \min\{x_i - \sum_j A_{ij}x_j, 0\}$. The new final demand vector $\boldsymbol{f}$ now implies a new, lower level of aggregate demand, $\boldsymbol{d} = \mathbf{L}\boldsymbol{f}$, and we again let industries evaluate whether they can satisfy total demand within given production constraints. We iterate this procedure forward until all demand is met

and no input bottleneck further constrains production. We can write the proportional rationing algorithm more compactly as follows:

Algorithm 1 *Proportional rationing; industries are not prioritized over the final consumer. Take an initial demand vector* $\boldsymbol{f}[0] = \boldsymbol{f}^{max}$ *as given, implying an initial aggregated demand vector* $\boldsymbol{d}[1] = \mathbf{L}\boldsymbol{f}[0]$*. By looping over the index* $t = \{1, 2, ...\}$*, the following system is iterated forward:*

$$r_i[t] = \frac{x_i^{max}}{d_i[t]}, \tag{4.7}$$

$$s_i[t] = \min_{j:A_{ji}>0}\{r_j[t], 1\}, \tag{4.8}$$

$$x_i[t] = \min\{x_i^{max}, s_i[t]d_i[t]\}, \tag{4.9}$$

$$f_i[t] = \max\left\{x_i[t] - \sum_j A_{ij}x_j[t], 0\right\}, \tag{4.10}$$

$$d_i[t+1] = \sum_j L_{ij}f_j[t]. \tag{4.11}$$

The algorithm converges to a new feasible economic allocation if $d_i[t+1] = d_i[t]$ *for all* i*. In this case output and final consumption levels are given as* $x_i = d_i[t+1] = x_i[t+1]$ *and* $f_i = f_i[t+1]$*, respectively.*

Eq. (4.7) indicates whether an industry is output constrained, where r_i is the share of demand that can be met given existing productive capacities. If $r_i \geq 1$, industry i is able to meet demand completely, whereas demand can only be partially satisfied if $r_i < 1$. If any supplier of industry i (which is the set $\{j : A_{ji} > 0\}$) is sufficiently output constrained, industry i faces an input bottleneck according to Eq. (4.8). Due to perfect complementarity of inputs prescribed by the Leontief production function, industry i can only produce a fraction of total demand as indicated in Eq. (4.9), reducing its delivery to final consumers as specified in Eq. (4.10). The new total demand to an industry i is then again derived through the weighted sum of final demand values where the weights are obtained from the Leontief inverse, Eq. (4.11).

(b) Mixed proportional/priority rationing. It has been argued that firm-firm relationships are stronger than firm-household ties, so that intermediate demand should be prioritized over final demand (Hallegatte 2008, Inoue & Todo 2019). While this assumption might make sense for households, this is not necessarily the case for other final demand categories. Final demand categories in national IO tables can include demand by governments and non-profit organizations, exports and investments

and it is debatable whether intermediate demand should be prioritized over these categories.

To quantify the effect of this assumption on the amplification of initial shocks, we implement a mixed proportional/priority rationing algorithm. Here, industries ration intermediate demand proportional analogously to the proportional rationing algorithm (a), but prioritize intermediate demand over final demand. We outline this algorithm below.

Algorithm 2 *Proportional rationing among industries; industries are prioritized over final consumers. Take an initial demand vector* $\boldsymbol{f}[0] = \boldsymbol{f}^{max}$ *as given, implying an initial aggregated demand vector* $\boldsymbol{d}[1] = \mathbf{L}\boldsymbol{f}[0]$*. By looping over the index* $t = \{1, 2, ...\}$*, the following system is iterated forward:*

$$r_i[t] = \frac{x_i^{max}}{\sum_j A_{ij} d_j[t]}, \tag{4.12}$$

$$s_i[t] = \min_{j:A_{ji}>0} \{r_j[t], 1\}, \tag{4.13}$$

$$x_i[t] = \min\{x_i^{max}, s_i[t] d_i[t]\}, \tag{4.14}$$

$$f_i[t] = \max\left\{x_i[t] - \sum_j A_{ij} x_j[t], 0\right\}, \tag{4.15}$$

$$d_i[t+1] = \sum_j L_{ij} f_j[t]. \tag{4.16}$$

The algorithm converges to a new feasible economic allocation if $d_i[t+1] = d_i[t]$ *for all* i*. In this case output and final consumption levels are given as* $x_i = d_i[t+1] = x_i[t+1]$ *and* $f_i = f_i[t+1]$*, respectively.*

The algorithm is similar to the proportional rationing algorithm (a) but differs mainly in one aspect. Only intermediate demand affects the extent of an industry's output constraints, Eq. (4.12). As a consequence, final demand does not play a role in the creation of input bottlenecks, Eq. (4.13).

(c) Priority rationing ("largest first"). Since it is not obvious that industries should pass on their output proportionally in case they are not able to meet demand fully, we next consider a type of priority rationing. Industries rank their customers based on demand magnitude and serve larger customers before smaller customers. In the proportional rationing setting an output constrained supplier affects all customers in the same way. Under a priority rationing scheme, however, supply shocks propagate downstream heterogeneously. For example, if a supplier cannot meet demand by only

a small margin, most customers will not be affected by the priority rationing scheme. Only the smallest customers will face input bottlenecks, whereas every customer would experience the same small shock in the proportional rationing setup.

Intuitively, a priority rationing rule could make sense, as firms might have an interest in serving more important (large) customers fully, or at least as well as possible, before focusing on less important customers. It also seems closer to practice that firms process orders one-by-one instead of working through all orders simultaneously and leaving them incomplete to the same degree. While a priority rationing scheme could be plausible on the basis of firms or single transactions, it might be less so for more aggregate industry-level data. A link between two industries in IO tables corresponds to many firm/establishment level transactions and so it is not clear if large inter-industry links are due to a few big orders or many small orders. This becomes particularly evident when considering final consumers. Several industries face large demand from private consumers which effectively is the sum of many small orders (e.g. restaurants, grocery, theaters). Since we only consider an aggregate of final demand representing several distinct categories such as private consumers, government and investments, we exclude final demand from the priority rationing scheme. Thus, in the same manner as in the mixed prop./prior. rationing algorithm (b), we assume that intermediate demand is always prioritized over final demand. By adopting this convention, we can formulate the priority rationing algorithm as follows.

Algorithm 3 *Largest first rationing; industries are prioritized over the final consumer. Take an initial demand vector* $\boldsymbol{f}[0] = \boldsymbol{f}^{max}$ *as given, implying an initial aggregate demand vector* $\boldsymbol{d}[1] = \mathbf{L}\boldsymbol{f}[0]$. *Every industry* i *ranks each customers* j *based on initial demand size:* $h_{ij} = \{k_{(1)}, k_{(2)}, ..., k_{(j)} \ : \ A_{ik_{(1)}} d_{k_{(1)}}[1] \geq A_{ik_{(2)}} d_{k_{(2)}}[1] \geq ... \geq A_{ik_{(j)}} d_{k_{(j)}}[1]\}$. *By looping over the index* $t = \{1, 2, ...\}$, *the following system is iterated forward:*

$$r_{ij}[t] = \frac{x_i^{max}}{\sum_{N \in h_{ij}} A_{in_{(j)}} d_{n_{(j)}}[t]}, \tag{4.17}$$

$$s_i[t] = \min_{j: A_{ji} > 0} \{r_{ji}[t], 1\}, \tag{4.18}$$

$$x_i[t] = \min\{x_i^{max}, s_i[t] d_i[t]\}, \tag{4.19}$$

$$f_i[t] = \max\left\{x_i[t] - \sum_j A_{ij} x_j[t], 0\right\}, \tag{4.20}$$

$$d_i[t+1] = \sum_j L_{ij} f_j[t]. \tag{4.21}$$

The algorithm converges to a new feasible economic allocation if $d_i[t+1] = d_i[t]$ *for all* i*. In this case output and final consumption levels are given as* $x_i = d_i[t+1] = x_i[t+1]$ *and* $f_i = f_i[t+1]$*, respectively.*

(d) Random rationing. As our final case, we again consider priority rationing. Rather than using a fixed ordering scheme based on demand magnitude, we use random priority. Industries rank their customers randomly and serve customers based on their position in the ranking. While largest-first rationing makes intuitive sense in some cases, it is unlikely to be a good approximation for all real-world settings. In practice, it is likely that other factors such as timing of orders matter. In that case industries could adopt a first-come-first-serve principle to process orders. Since IO data rarely comes with granular time information, we adopt a random ordering of incoming orders that mimics a first-come-first-serve principle under a uniform prior of which orders are coming in first. The algorithm is presented below.

Algorithm 4 *Random rationing; industries are prioritized over the final consumer. The algorithm is identical to Algorithm 3, except that the ranking of customer* j *by industries* i*,* $h_{ij} = \{k_{(1)}, k_{(2)}, ..., k_{(j)}\}$*, is randomly drawn.*

While these algorithms are not guaranteed to converge to a steady-state equilibrium, we observe convergence in the vast majority of our simulation experiments. If the algorithm converges, the economic allocations obtained are automatically feasible. This means that no industry has negative output or produces more than its productive capacities allow, final consumption is non-negative and below given exogenous maximum consumption levels, and there are no input bottlenecks left that further constrain production of downstream industries. The Leontief equation $\boldsymbol{x} = \mathbf{L}\boldsymbol{f}$ holds, which implies that all sales add up to total output.

4.4.3 Results

We initialize these four rationing algorithms with the supply and demand shocks and IO data of Germany, Italy and Spain. We then compare the steady-state equilibrium economic outcomes predicted by the rationing algorithms with the optimization results and the direct shock computations (Appendix D.1, Tables D.1 and D.2).

Fig. 4.1 summarizes the main result visually, where a diamond indicates the aggregate final consumption and gross output levels predicted by the different approaches. Note that a data point in the top-right corner indicates high gross output and final

consumption values, i.e. limited negative impacts, whereas diamonds in the bottom-left corner represent extremely adversely impacted economies. As already reported in Table 4.1, for all three countries supply shocks are substantially larger than final consumption shocks, pushing the *Direct shock* diamond substantially below the 45 degree line. Note that the direct shock market allocations are not feasible, as they ignore higher-order effects, such that a direct supply shock to an industry reduces the inputs of downstream industries and thus reduces their output too.

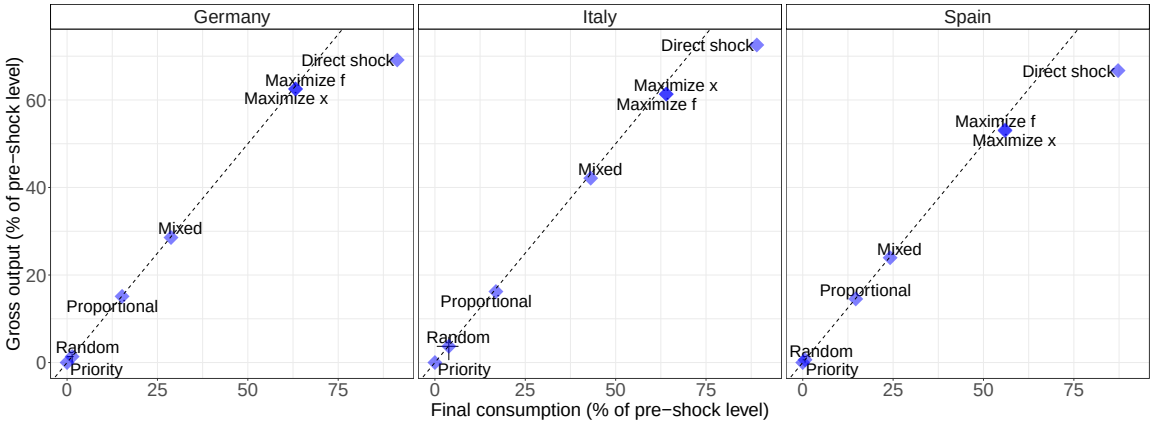

Figure 4.1: **Comparison of different shock propagation mechanisms.** The y-axis denotes aggregate gross output levels normalized by pre-shock levels as predicted by the rationing and optimization methods. The x-axis shows the same for aggregate final consumption levels. The *Random rationing* diamond represents the average taken from 100 samples and the error bars indicate the interquartile range.

The best case feasible market allocations are given by the two optimization methods (maximize output vs. consumption), which yield exactly the same predictions on the aggregate and the industry level. This is true even though the two optimization methods are not equivalent.[5] The two optimization methods have the highest output, but they still amplify the initial supply shocks to reduce the output from about 69% to 63% in Germany and from 67% to 53% in Spain.

All the other feasible market allocations, obtained from the rationing algorithms, lie substantially below the best case scenarios. The wide range of predicted outcomes is the most striking aspect of Fig. 4.1. Both the random rationing scheme and the priority rationing scheme essentially collapse the entire economy. Proportional rationing substantially collapses the economy (with an output below 20% of normal for all countries) and the mixed scheme reduces output for Germany and Spain by more

[5]When perturbing the economic systems we can find cases where the two optimization methods do not yield exactly the same results. Practically, we find that aggregate predictions are always similar for the two methods although industry-level results can sometimes differ significantly.

than 70% and about 60% for Italy. Interestingly, all feasible market allocations lie close to the identity line, suggesting similar impacts to output and consumption.

On a sectoral level, we find the overall orderings of economic impacts are highly correlated across methods and countries, despite a few pronounced differences. An interesting example is Forestry (A02) which represents the only industry that experiences less adverse impacts under proportional rationing compared to mixed rationing. Due to being non-essential and having a low Remote Labor Index (Appendix D.1, Table D.2), forestry is the industry with the largest supply shocks. Thus, it virtually always faces stronger supply constraints than demand constraints. In case of no prioritization of industries, Forestry receives less demand and thus creates smaller input bottlenecks for downstream industries. This, in turn, leads to smaller input bottlenecks for Forestry as its downstream industries can be found upstream as well.

There is also substantial variation in how close industries are to their theoretical maximum value as determined by the optimization method. When considering the proportional rationing algorithm in Italy, for example, the output levels of the best faring industries (such as Education (P85) or Telecommunications (J61)) reach around 30% of their theoretical maximum. On the other hand, the hardest hit industry, Accommodation-Food (I), is only at a 5% level of what is possible. Thus, Accommodation-Food is not only hit extremely hard by direct shocks, but also experiences large higher-order effects through the rationing dynamics.

Fig. 4.1 makes it clear that the behavioral assumptions imposed on suppliers matter enormously for economic impact predictions. In fact, results vary more across alternative assumptions than across countries.

In the absence of further modeling refinements (such as inventories, adaptive behavior, substitution effects), shock amplification is always pronounced if basic national accounting identities are required to hold. But the actual extent of shock amplification depends strongly on how input bottlenecks are created and passed on downstream in the supply chain.

4.4.3.1 Shock magnitude effects

We next investigate the sensitivity of economic impact predictions with respect to shock magnitude and network connectedness. While we have seen in Section 4.4.3 that different assumptions on rationing behavior can lead to entirely different estimates of impact, it is not clear whether these results are specific to the three datasets considered. We therefore conduct a series of simulation experiments to gain a better understanding of the generality of the results.

First, we investigate how the estimates of the various methods depend on the magnitude of shocks. To do this, we rescale supply and demand shocks and we apply the optimization methods and the rationing algorithms to the new shock data. We then redo this analysis for various shock scales. To better differentiate between the qualitative effects of demand and supply shock propagation, we allow for different scaling factors for demand and supply constraints, i.e.

$$x_i^{\max} = (1 - \alpha^S \epsilon_i^S) x_{i,0}, \tag{4.22}$$

$$f_i^{\max} = (1 - \alpha^D \epsilon_i^D) f_{i,0}, \tag{4.23}$$

where $\alpha^S, \alpha^D \in [0, 1]$.

Fig. 4.2(a) shows aggregate output levels for all cases when only demand shocks are scaled between zero and one and when there are no further supply constraints being present ($\alpha^S = 0, \alpha^D \in [0, 1]$).[6] It becomes evident that predictions made by the rationing algorithms and the optimization methods are identical and scale linearly with the demand shock magnitude. Thus, the rationing algorithms do not differ with respect to upstream shock propagation and always arrive at the optimal solution in absence of further supply side constrictions.

The picture becomes very different when rescaling only supply shocks and turning off demand shocks ($\alpha^D = 0, \alpha^S \in [0, 1]$) as demonstrated in Fig. 4.2(b). Here, the direct impact reduces gross output linearly as the shock magnitude is increased and puts an upper bound to the solutions of the methods considered here. When there are no shocks, $\alpha^S = 0$, all methods recover the empirical IO data as expected. For small shocks we observe that estimated impacts are very similar across different methods, but the proportional rationing algorithm consistently returns the smallest output values. The results of the other algorithms (mixed, priority, random) lie between those proportional rationing and optimization predictions, and are identical for small shocks up to about $\alpha^S = 0.5$. As α^S is increased further they deviate dramatically. Under the priority algorithm even a small increase in α^S causes the economy to collapse. The random algorithm also causes the economy to collapse, though more slowly, and the mixed algorithm is substantially better.

Fig. 4.2 thus makes clear that the ranking of the impact assessment methods as seen in Fig. 4.1 is not generic, but strongly depends on the shock magnitude. If there are only small supply shocks present, a priority rationing rule is better than

[6] Since results for aggregate final consumption and aggregate gross output are very similar, we only present figures of the latter in this section. We show results for scaling supply and demand shocks concurrently ($\alpha^S = \alpha^D \in [0, 1]$) in Appendix D.4.

proportional rationing. In this case most of the demand of downstream industries can be met and small shocks only propagate to few customers. In a proportional rationing setup, on the other hand, shocks are passed on to every customer. Despite small shock magnitudes, the wider breadth of shock propagation leads to comparatively larger aggregate impacts.

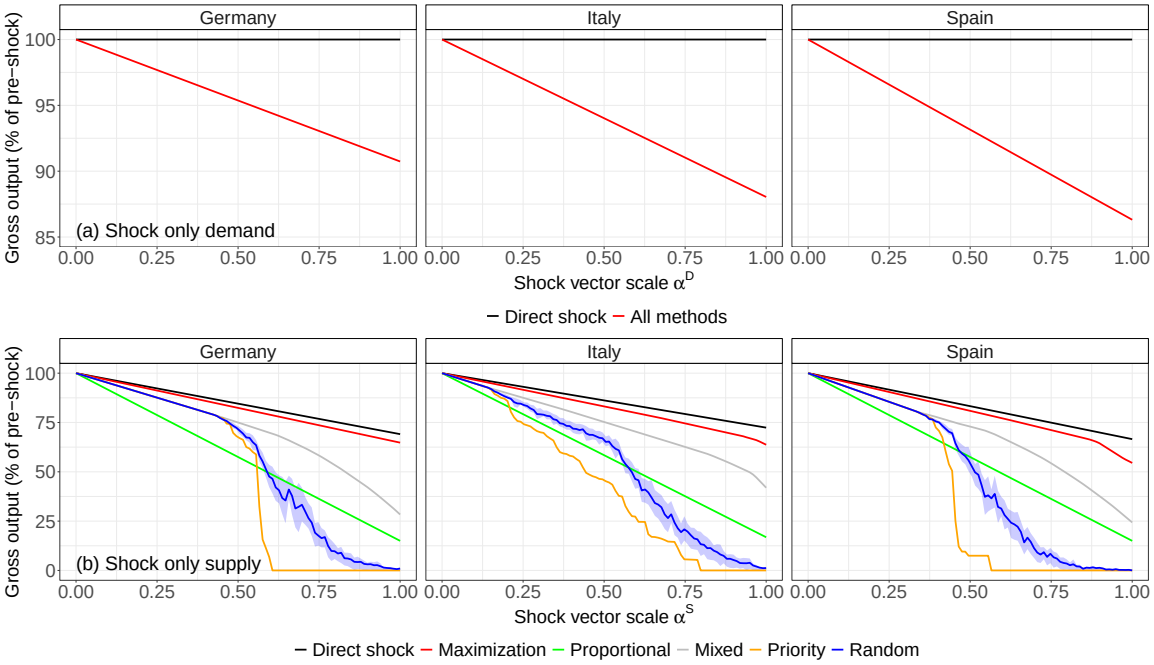

Figure 4.2: **Economic impact as a function of shock magnitude.** (a) Aggregate gross output levels as a function of scaling only demand shocks between zero and one ($\alpha^S = 0, \alpha^D \in [0,1]$). All methods yield the same economic impact estimates which scale linearly with α^S. The black horizontal line indicates that there is no adverse economic impact from supply side constraints. (b) Aggregate gross output levels as a function of scaling only supply shocks between zero and one ($\alpha^D = 0, \alpha^S \in [0,1]$). The *Random rationing* line represents the average taken from 100 samples and the shades indicate the interquartile range. We do not distinguish further between the two maximization methods as results are almost identical.

In contrast, priority rationing exacerbates shock propagation compared to proportional rationing in case of large supply shocks. If industries are severely output constrained and ration based on a priority rule, the demand of several customers' might not be satisfied at all. The Leontief production function, in turn, implies a complete shutdown of these downstream industries due to the input bottlenecks created. More input bottlenecks will be created in further rounds of shock propagation, potentially causing massive collapses. If supply shocks are that large, shock amplification is milder if passed on proportionally to customers. While here every customer experiences some shocks from a constrained supplier, this effect is outweighed by the

fact that proportional rationing avoids the creation of even larger idiosyncratic input bottlenecks.

4.4.3.2 Network density

Intuitively, the extent of shock propagation not only depends on the size of initial shocks but also on how industries are connected with each other. If the production network is dense, i.e. most of the potential links are present, idiosyncratic shocks will spread out very quickly to many other industries in the network. In contrast, if the network is sparse, shock propagation might be more local, at least initially, and takes more steps to spread out in the network.

We therefore conduct an experiment where we again apply the different shock propagation mechanisms to the data, but control for the IO network density. We do this in the following way. First, we randomly eliminate a given number of links in the intermediate consumption matrix $\mathbf{Z}$.[7] We only consider deleting links instead of adding links since the aggregate IO networks we are using are highly dense ($\sum_{ij} \mathbb{1}_{\{Z_{ij}>0\}}/N^2 > 99\%$). Note that setting a link $Z_{ij} > 0$ equal to zero without changing final consumption values reduces total output of supplier i, since the accounting identity $x_i = \sum_j Z_{ij} + f_i$ has to hold. The output of customer j will not be affected if we assume that the reduction in intermediate consumption is absorbed by a respective increase of j's value added (which does not affect the simulations). After deleting nodes we thus rebalance the economic system by adjusting gross output of the relevant industries such that the basic national accounting identity is satisfied. Next, we recompute the technical coefficient matrix $\mathbf{A}$, the Leontief matrix $\mathbf{L}$ and the vector of productive constraints $\boldsymbol{x}^{\max}$ to initialize the optimization and rationing simulations. We repeat this procedure 50 times for a given level of density and explore the whole range of possible density values.

We visualize aggregate output levels predicted by the various impact assessment methods as a function of network density in Fig. 4.3. The plot again confirms that the observed ranking of methods in Fig. 4.1 is not generic but is strongly affected by the network density. It becomes clear that shock amplification is comparatively small if the network is very sparse. On the other hand, if the network is very dense, as the empirical data would suggest, economic impacts are substantial for all cases. In particular, the gap between the optimal solution and the bottom-up rationing

[7] We also experimented with eliminating smallest links first instead of randomly selection. Results are similar to the ones presented here and shown in the Appendix D.3.

approaches widens with higher network density. Interestingly, the minimal amplification of direct shocks also increases with higher density values, although the size of this effect is rather small.

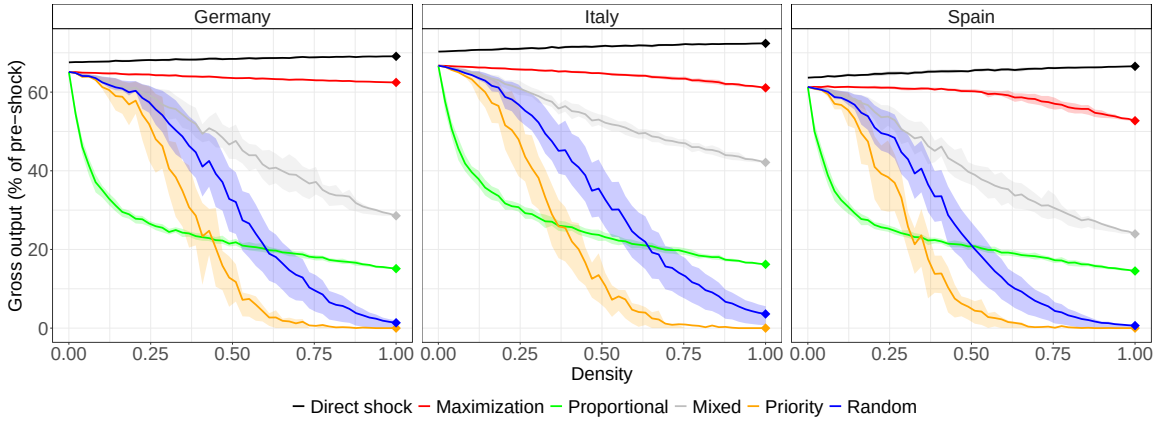

Figure 4.3: **Economic impact as a function of network density.** The network density is changed by randomly eliminating links in the IO network. Solid lines are predicted output values obtained from averaging over 50 samples and shades indicate the interquartile range. Since random rationing is stochastic by itself, we apply this algorithm 50 times to every network sample and take averages and quantiles over the full (pooled) density-specific sample. The actual data corresponds to the very right of the x-axis, as indicated by the diamonds.

If industries have only few suppliers and customers, proportional rationing yields by far the most pessimistic predictions of aggregate impact. If there are many connections among industries, on the contrary, proportional rationing mitigates shock propagation dynamics compared to priority and random rationing. Shock amplification is always mildest if industries use a mixed proportional/priority rationing rule, although even in this case economic impacts increase substantially with higher density.

These results indicate that estimates of economic impact are highly sensitive with respect to the network structure which, in turn, often depends on the level of data aggregation. More aggregated data necessarily implies higher levels of density. For example, industry-level links are the result of pooling many firm links. Firm-level production networks, on the other hand, are an aggregation of many individual contractual relationships and payments. Imagine applying the rationing mechanisms to a firm-level production network in two ways: first to the actual firm-level data and second to an industry aggregate of the same data. Our results suggest that predicted economic impacts could be substantially larger in the second case, despite using the same underlying data.

Our findings also point out that economic impact predictions can be sensitive with respect to data quality. Many disaggregate firm-level production network datasets are substantially biased, as they frequently include only specific supplier-customer relationships which are subject to specific reporting rules. Thus, we would expect real world production networks to be substantially denser than what the data suggests. Our analysis indicates that such biases could have important consequences for impact assessment.

4.5 Discussion

We have shown that existing IO models have difficulties in dealing with simultaneous supply and demand shocks. However, both types of shocks play an important role in situations such as natural disasters and during a pandemic. We have introduced a simple optimization method that allows us to find best case market allocations that are consistent with the exogenous shocks. To obtain more realistic, bottom-up impact estimates, we studied alternative rationing dynamics which differ in how suppliers serve customers in case of output constraints. Using IO data for Germany, Italy and Spain, we found that these bottom-up approaches lead to substantial amplifications of initial shocks, which are much higher than optimal solutions would suggest. We further established that different rationing assumptions can lead to dramatically different economic impact predictions. Moreover, these predictions are highly sensitive with respect to the magnitude of first-order shocks and the production network structure.

It is clear that adequate macroeconomic predictions of pandemic impacts require more sophisticated modeling techniques than the ones studied here. Yet the downside of more complicated models is that the underlying mechanisms of predictions can be difficult to isolate. We therefore studied relatively simple economic models to better carve out key mechanisms of shock propagation in twofold constrained production networks. The choice for simplicity in this study also entails limitations which are important to bear in mind. We did not depart from the assumption of industry-specific Leontief production functions throughout our analysis, although the choice of production function has been shown to be a key variable for impact assessment (Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer 2020). While input substitutions might be limited in the short-run, fixed production recipes are nevertheless a strong assumption, in particular for the aggregate data considered here.

It is also important to stress that we did not take any adaptive behavior of economic agents into account and did not allow for the possibility of inventory buildup and depletion. By imposing feasible market solutions, we essentially forced the economy to converge into a new equilibrium. It is not clear, however, that a perturbed complex system such as national economies would quickly approach a new equilibrium state instead of following transient paths for an extended period. In general, we refrained from making the time dimension explicit, although we acknowledge its relevance for shock propagation dynamics.

Despite the caveats mentioned, our analysis makes clear that level of aggregation and data quality play an important role in aggregate predictions of shock amplification. Detailed and high-quality production network data are rare, necessitating researchers to study biased or aggregate data. Inevitably, this will affect the connectedness of economic agents and thus influence shock propagation mechanisms. For example, studies based on large-scale firm level databases report network density values well below 0.01% (Kumar et al. 2021, Borsos & Stancsics 2020, Peydró et al. 2020, Demir et al. 2020, Huneeus 2018, Spray 2017). We should also mention that estimates of direct shocks to supply and demand are subject to large uncertainties. This could have important consequences for model predictions, due to the sensitivity of shock propagation mechanisms with respect to first-order shock magnitudes.

We find that the number of links between industries strongly influences how different behavior rules amplify direct shocks. The level of connectedness is a very simple aggregate network measure, and we would expect that further aspects, such as community structure, degree/strength heterogeneity or (dis-) assortative mixing, play an important role too. Understanding how alternative rationing assumptions interact with structural properties of complex networks would require more detailed production network data, e.g. at the firm level, and could be an interesting avenue for future research.

We conclude by stressing that our analysis is not constrained to pandemic shocks. While the Covid-19 pandemic is a main motivation of our study, simultaneous supply and demand constraints are ubiquitous features of any economy, in particular in the short run. Supply shocks are a prominent characteristic of many natural hazards (floods, earthquakes, hurricanes) and tools of mostly demand-driven IO analysis are frequently applied for impact assessment. Our results indicate that adequately modeling shock propagation in production networks will require a better integration of both types of economic constraints.

Conclusion

In this thesis we have investigated network-dependent dynamics of innovation and production systems. We have used historical records of patents and their technological classifications to sketch out long-term inventive dynamics of various technological categories. Using both, conventional time series approaches and more sophisticated network models, we have analyzed technology specific patenting rates and have tested the predictive power of these techniques. We have found that rather simple time series models yield relatively good predictions of future patenting dynamics. Reducing the number of model parameters to fewer general system parameters can further improve predictions, given the strong common trends in the evolution of technological classes.

Based on our observation of network-dependent patent growth correlations, we have introduced a network model of knowledge creation. We have shown that the model can give interesting insights when solving it for the steady state. We have derived an estimator to calibrate the model to the data and used it for making predictions of future patenting rates. We have found superior predictive power of the novel network model compared to conventional time series model benchmarks.

Motivated by the sudden onset of the Covid-19 pandemic, we have developed a data-driven model of the UK economy tailored to address the specific features of the pandemic. At the core of the model lies the input-output network where industries dynamically use intermediate inputs to produce goods. We have shown that the model yielded remarkably accurate predictions during the early phase of the pandemic and is able to reproduce sectoral and aggregate economic dynamics very well. By varying model assumption, we have established a number of interesting theoretical insights. We have shown that inventory levels and alternative production function specifications can strongly influence how shocks propagate through the network and that the first-order shock size is not necessarily predictive for overall economic impacts. When shocking industries one-by-one, we have found that neither upstreamness nor downstreamness of an industry fully explains the model results.

We have looked closer into key mechanisms of shock propagation in production networks by studying a simpler framework that is closer to traditional input-output models. We have shown that existing models have difficulties dealing with supply and demand shocks simultaneously. Consequently, we have proposed a simple optimization framework that allows us to derive lower bounds on economic impacts conditional on both shock types. To study more realistic shock propagation scenarios, we have followed a bottom-up approach where we have equipped industries with different rationing rules. Input bottlenecks in the production system arise very differently under various rationing assumptions. Thus, different rationing rules lead to very different estimated economic impacts, making them a key assumption in disaster modeling.

This thesis closes some research gaps in dynamically modeling innovation and production networks but leaves ample room for future research. For example, patents are known to strongly interact with scientific papers (Ahmadpoor & Jones 2017, Hötte et al. 2021). It remains to be shown whether our understanding of patenting growth dynamics can be improved by complementing our analysis with data on scientific research output. Similarly, R&D spending is a major driver of technological progress and thus, we should expect improved results from incorporating detailed data on research investments in different technological categories.

It should be pointed out that we have studied innovation and production networks independently from each other. However, we would expect that there are strong interactions between inventive and productive systems. Here, we have quantified the extent of mutual influence in technological invention dynamics. Technological improvements, however, will also affect the mode of production, for example in terms of productivity improvements. Given the interconnected nature of firms and products, we would expect these improvements to percolate through the production network (McNerney et al. 2018). Future research could aim at combining data on patent and production networks to actually quantify these effects. Surana et al. (2020) and Hötte (2021) are two notable recent efforts toward this direction.

A promising future research agenda is to place innovation and production networks at the center in economic models of climate change and the "green transition" (Balint et al. 2017). The dynamic input-output model proposed here could be a starting point for modeling adverse climate shocks. Extending the model with innovation, capital and financial modules to accommodate for economic growth and structural change dynamics, could be a step towards a data-driven economic model of the green transition. This type of research could benefit from several further extensions such

as fine-graining the model on material flow data and firm-level data to improve our representations of intermediate consumption relationships and production functions.

Appendix A

Appendix to Chapter 1

A.1 Data sources and preprocessing

Our database contains all US utility patents granted between 1836 and 2017, except for rare exceptions. We focus on technological classes defined by the Cooperative Patent Classification (CPC) system which represents the current classification scheme used by the USPTO to categorize patents. Our final data set covers 10.1M unique patents and 14.4M patent-technology (on the 4-digit CPC level) observations. The reason why there is more patent-technology observation is due to co-classification, i.e. a patent can be assigned to multiple CPC codes. CPC classifications are retrieved from the Master Classification File available from the USPTO/Reed Tech (version 01/04/2019).

We map patents onto their CPC codes and eliminate duplicates if the same patent was classified multiple times within the same category. We also eliminate all Y-codes, which are tags rather than separate technology classes, and the "2000 series" related to indexing schemes codes (EPO & USPTO 2017).

Note that Fig. 1.1 in the main text uses a different dataset. To ensure better comparability, we have used the original dataset provided by Acemoglu et al. (2016) which we discuss in more depth in Appendix A.4.

A.2 Principal component analysis

To get a better understanding of the common underlying trends, we have conducted a principal component analysis (PCA) on the growth rates over various aggregation levels. We restrict the patent data sample to the time horizon 1867–2017, to avoid conflating the results with the strong upward trend in patenting in the early period (Fig. 1.2) that is driven by relatively few patents. We conduct the PCA on the basis of

correlation instead of covariance matrices, since we found results to be quite sensitive with respect to very small, noisy categories when using covariance matrices.[1] More formally, we represent the technology specific growth rates as the N-dimensional random vector $\boldsymbol{g} = (g_1, ..., g_N)$ which we normalize by the corresponding standard deviations σ_i to obtain $\tilde{\boldsymbol{g}} = \mathbf{diag}(\boldsymbol{\sigma})^{-1}\boldsymbol{g}$. We let $\boldsymbol{\Sigma}$ denote the covariance matrix of $\tilde{\boldsymbol{g}}$ (or equivalently the correlation matrix of $\boldsymbol{g}$). The i^{th} principal component is then given as $\boldsymbol{z}_i = \boldsymbol{e}_i^\top \boldsymbol{g}$ where $\boldsymbol{e}_i$ is the eigenvector corresponding to the i^{th} largest eigenvalue λ_i of $\boldsymbol{\Sigma}$ (for details see Jolliffe (2002)). A key insight of PCA is that we have

$$\frac{\text{Var}(z_i)}{\sum_i \text{Var}(g_i)} = \frac{\lambda_i}{\sum_i \lambda_i}. \tag{A.1}$$

Thus, the ratios between the eigenvalues of $\boldsymbol{\Sigma}$ and $N = \sum_i \lambda_i$ yields the share of variance explained by a given principal component. Table A.2 shows that on the most aggregate level of eight technological classes the first principal component $\boldsymbol{z}_1$ explains more than three quarters of the variance. Using more detailed technology classifications on the 3-digit (123 technologies) and 4-digit (626 technologies) levels, this number reduces to around 30% and 13%, respectively.

To test whether the first principal component corresponds to the overall system mode, we rotate the data by the first eigenvector

$$\boldsymbol{g}^{\text{rotated}} = \mathbf{G}^\top \boldsymbol{e}_1, \tag{A.2}$$

where $\mathbf{G}$ is the $N \times T$ matrix with elements $g_{i,t}$. We then regress the rotated growth data onto the yearly cross sectional averages

$$g_t^{\text{rotated}} = \beta_0 + \beta_1 \bar{g}_t + \zeta_t, \tag{A.3}$$

and show the results in Table A.1. The regression confirms that the first component closely corresponds to the system average growth rate, i.e. represents a system component analogous to a market component in financial time series analysis (Tsay 2010). These results suggest that there is a strong common trend across growth rates of various technological classes underlying the patenting system, although the effect is smaller for more fine-grained classifications. Centering the growth rates by their cross-sectional averages substantially reduces the relative size of the principal component for all classification levels.

[1] Whenever our correlation computation yielded a non positive-semidefinite correlation matrix due to numerical instabilities, we use its positive-semidefinite approximation based on Higham (2002) and obtained via the *R Matrix* package (Bates & Maechler 2019).

	Dependent variable: Centered growth rates g_t^*		
	Section (1-dig)	Class (3-dig)	Subclass (4-dig)
Rotated growth rates g_t^{rotated}	-0.355^{***}	-0.095^{***}	-0.043^{***}
	(0.001)	(0.001)	(0.0004)
Constant	0.0004^{**}	0.001^{*}	0.003^{**}
	(0.0002)	(0.001)	(0.001)
Observations	151	151	151
R^2	1	0.994	0.988
Note:		*p<0.1; **p<0.05; ***p< 10^{-16}	

Table A.1: **Regressing the cross sectional averages against the rotated first component.** Regressions give extremely good fits, suggesting that the principal components closely correspond to the average growth rate in the patent system.

Analogously to the industry component in financial time series (Tsay 2010), we further test whether aggregate technology groupings explain growth rate variations on the more disaggregate level additionally to the first principal component. We do this by taking advantage of the nested nature of the CPC classification scheme. Instead of computing the variance explained by the first component on the raw growth data $g_{i,t}$, we also perform the PCA on the centered data where we use different within-group averages for centering. For example, using the most granular 4-digit technology classes, we center the growth rates by their corresponding 1- and 3-digit category averages, as well as with the overall cross sectional mean $\bar{g}_{i,t}$. We then apply the PCA to the centered data and compute the variance explained by the first leading principal component. Table A.2 shows the percentage of variation explained for different centering approaches. Removing the cross-sectional averages from the actual growth rates reduces the variance explained by the first component from 77% to 24% for the 1-digit categories and from 13% to slightly above 2% for the 4-digit case. This confirms our observation that the overall system mode is an important factor of overall growth rate variations. Interestingly, when subtracting aggregate technology specific averages from the granular technology growth rates, the size of the first principal component is not further removed; see last two rows of Table A.2 and Fig A.1. Thus, we do not find any evidence for an "industry" mode that explains patenting growth rates beyond the strong "system" mode identified by the first principal component.

	Section 1-dig	Class 3-dig	Subclass 4-dig
Nr categories	8	123	626
PC actual	77.1 %	30.2 %	13.3 %
PC centered	23.6 %	5.3 %	2.4 %
PC centered 1-dig		6.4 %	2.5 %
PC centered 3-dig			2.6 %

Table A.2: **Share of variance explained by the leading principal components for different aggregation levels.** The first row indicates the number of separate categories per aggregation level. PC actual refers to the share of variance explained by the principal component (leading eigenvalue divided by the sum of all eigenvalues of the correlation matrix) based on the actual growth rates. PC centered is the same but based on centered, i.e. cross-sectionally demeaned, growth rates. For the rows PC centered 1-dig and 3-dig growth rates have been centered by 1- and 3-digit cross-sectional averages, respectively. On the subclass level we had to exclude five categories to obtain a well-defined correlation matrix.

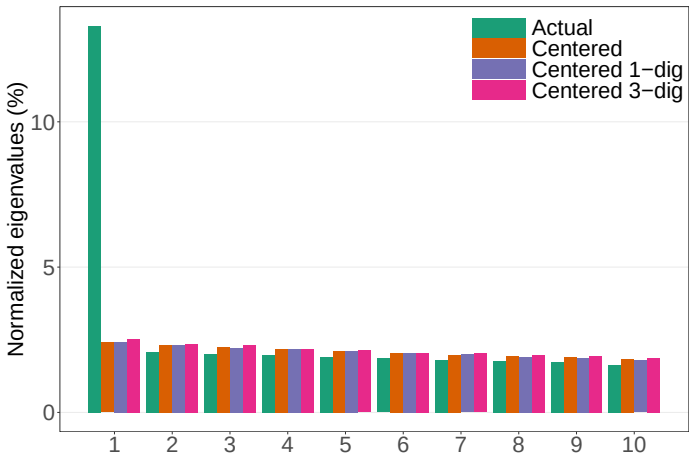

Figure A.1: **Eigenvalue spectrum (scree plot) based on 4-digit categories.** Centering the data strongly reduces the size of the leading eigenvalue. Centering the 4-digit growth data with 1- and 3-digit does not reduce the leading eigenvalue any further. Note that the x-axis is cut off on the right hand side.

A.3 Further results

Details on IMA(1,1) parameters. We further investigate how the estimated parameters of Eq. (1.1) relate to each other and further possibly relevant aspects, such as technological class size or the age of a technology. Fig. A.2 shows bivariate scatter plots of estimated drift $\hat{\mu}_i$, noise autocorrelation $\hat{\theta}_i$ and noise standard deviation $\hat{s}_i$. There are only small, weakly significant positive relationships between drift and both autocorrelations and standard deviations. We find, however, some strong negative

relationship between the standard deviation of the noise and the autocorrelation parameter (Pearson corr. of -0.53).

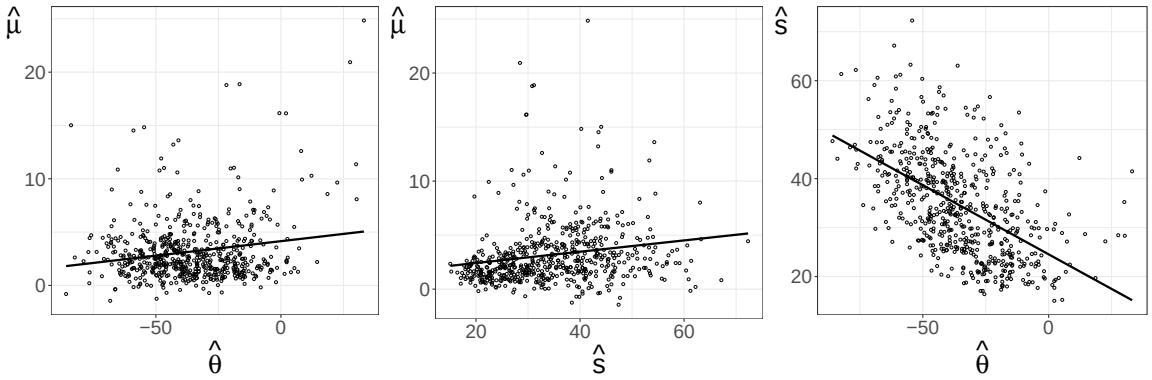

Figure A.2: **Bivariate parameter scatter plots**. All values are expressed in %. Solid lines represent ordinary least square regression fits.

Fig. A.3 shows a scatter plot of the estimated noise standard deviation $\hat{s}_i$ and the size of a technological class, measured in the average yearly number of patents $\sum_{t=1}^{T} p_{i,t}/T$. The figure shows that larger technological classes are substantially less volatile. Computing the Pearson correlation on the log-transformed values yields -0.59.

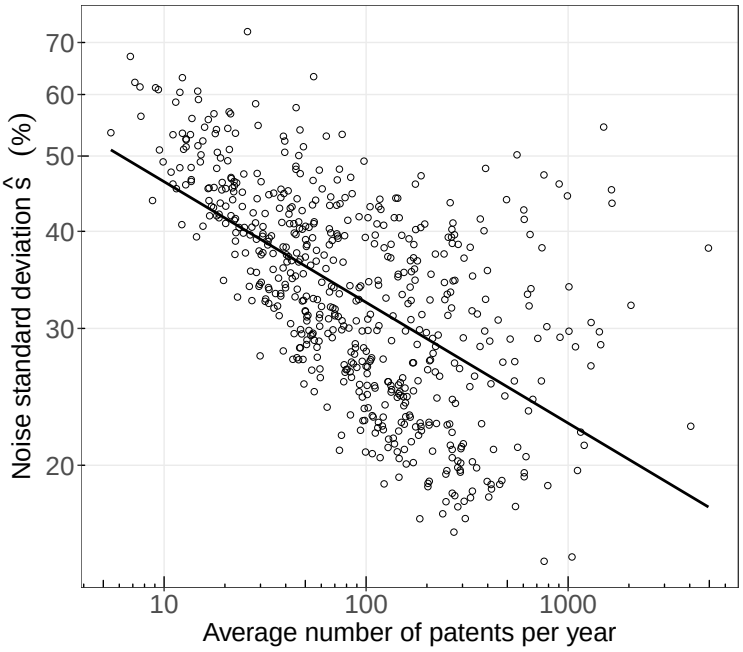

Figure A.3: **Larger patent classes grow less noisily**. The x-axis denotes the average number of yearly patents of a given class i, once class i has been established. The solid line represents the ordinary least square regression fit. Both axes are in log-scales.

We next discuss whether the age of a given technological class, i.e. the number of years since a patent first appeared with this classification, affects its growth rate. Fig. A.4 underlines our observation from Table 1.1 that it is mostly young technological categories exhibiting large growth rates. Overall, we observe a strong negative relationship between the estimated drift parameters and the technological lifetime (Pearson corr. of -0.71). Note, however, that the technological age distribution is strongly left-skewed, as shown in the lower panel of Fig. A.4.

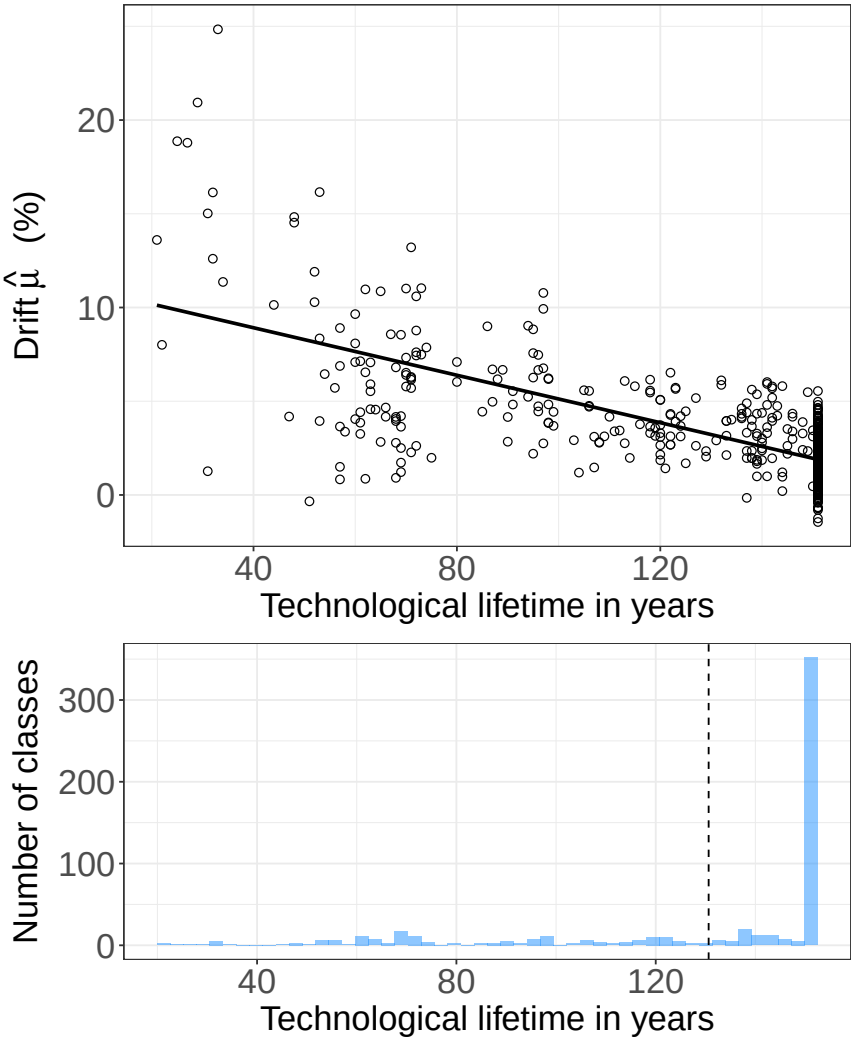

Figure A.4: **High drift values are only found for relatively young technologies**. Technological lifetime refers to the number of years since a given class has been established (time series length). The solid line in the upper panel represents the ordinary least square regression fit.

Details on the hindcasting appraoch. The number of forecasts we can make with our hindcasting approach depends on the choice of m and τ. As Fig. A.5

indicates, we can make more than 30 thousand forecasts for $m = 10$ and $\tau = 1$, but substantially less with larger estimation windows such as $m = 50$.

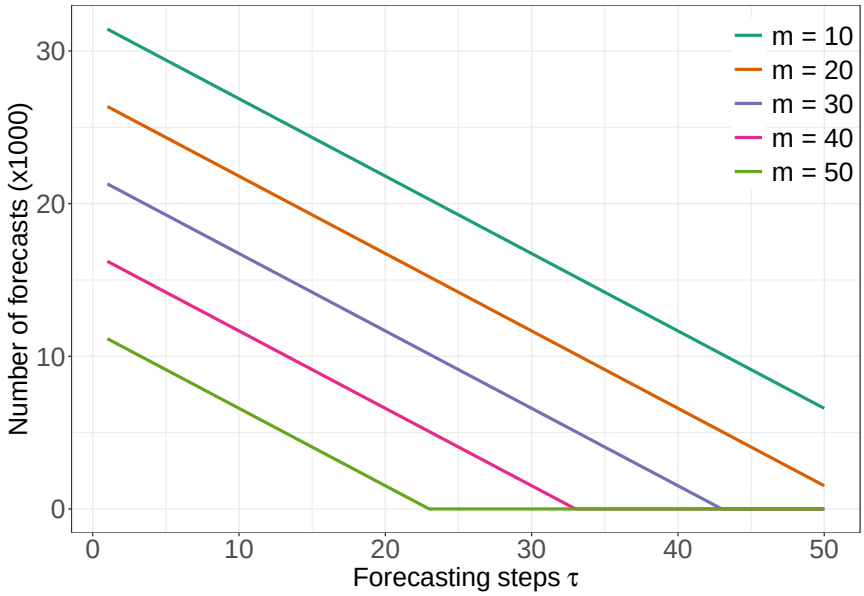

Figure A.5: **Number of forecasts for given combination of estimation window m and forecasting steps τ.** Note that the y-axis in thousands.

The upper panels of Fig. A.6 depicts the mean normalized forecast errors, analogously to its squared counterpart shown in Fig. 1.7. We observe that the length of the estimation and forecast horizons affect the overall bias, but there is only small differences between the competing model specifications. The center and lower panels of Fig. A.6 reproduce the left and right panels of Fig. 1.6, but for further choices of estimation windows m.

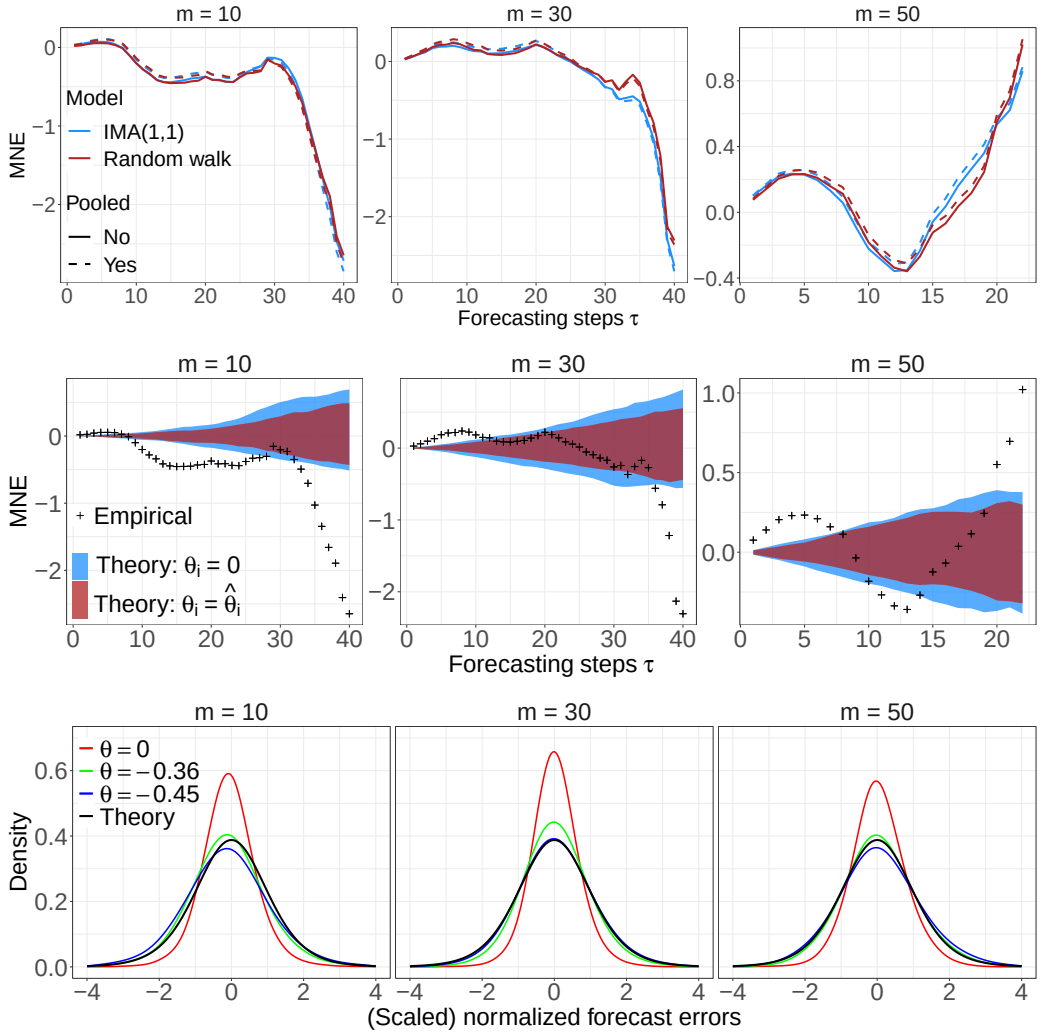

Figure A.6: Upper panels: Comparison of mean normalized forecast errors across alternative model specifications. Center panels: Mean normalized forecast errors obtained from random walk predictions on empirical, synthetic random walk and synthetic IMA(1,1) data. Lower panels: Forecast errors normalized with varying choices of the autocorrelation parameter θ.

A.4 Comparing the innovation network with the random walk

Fig. 1.1 of the main text indicates that the random walk model outperforms the network model introduced by Acemoglu et al. (2016) in predicting future patenting dynamics. We adopt the terminology used by Acemoglu et al. (2016) who called the model the *innovation network*. In this appendix we explain the innovation network model in more depth, provide some additional context and show further quantitative comparisons between this model and the geometric random walk.

Data. When comparing the time series models with the innovation network, we use the smaller dataset from Acemoglu et al. (2016). This dataset, together with other reproduction files, are kindly provided by the authors at `https://economics.mit.edu/faculty/acemoglu/data`. This data contains 1.8 million US utility patents applied for between 1975 and 2004 with at least one inventor living in a US metropolitan area. Fig. A.7 visualizes how the data set differs from the total number of patent applications. The blue line shows the annual number of patent applications in the US according to the data provided by the USPTO (Marco et al. 2015). The red line shows the number of applications per year in the data set considered here. The data of Acemoglu et al. (2016) encompasses between 27-44% of all US patents, depending on the year. The patenting time series of the smaller data set differ substantially from the overall patenting activity. In particular, there is a substantial decline in patent applications in the last four years. The total number of applications drops by more than 20% between 2001 and 2004, while we do not observe anything similar in overall US patent applications.

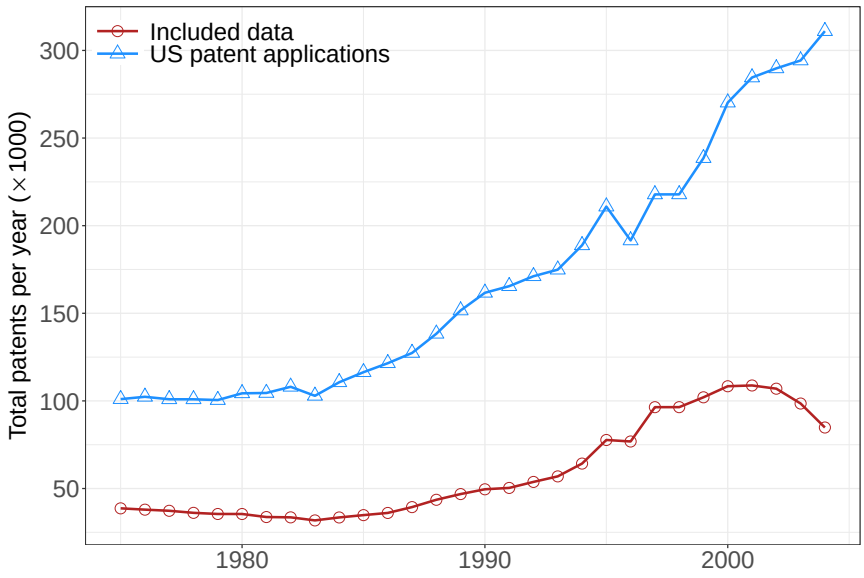

Figure A.7: Total US utility patents applied in a given year following the data of Marco et al. (2015) and number of patents included in Acemoglu et al. (2016).

Here, patents are grouped into various technological categories based on 353 USPTO main classes (USPC), and their grouping into 36 subcategories (SCAT) by the seminal NBER study of Hall et al. (2001). Note that the USPC is no longer maintained and updated by the USPTO. Another qualitative difference between the two

datasets is that the data from Pichler, Lafond & Farmer (2020) uses date of granting as time index, while Acemoglu et al. (2016) use application dates.

Acemoglu et al. (2016) use patent citations to construct the network model. Before calibrating the innovation network, they conduct some cleaning on the citation data. To be included in the citation network, *cited* patents have to be applied for between 1975 and 1993 and *citing* patents between 1976 and 1994. The total number of citations S_p (see Eq. (A.5) below) is then computed by considering only citation pairs with time lags ranging from one to ten. After obtaining S_p, all citations where the *cited* patent was applied for before 1985 are excluded. The final cleaned dataset contains slightly more than 1 million citations.

The original work was carried out in STATA and the analysis presented here was mostly done in R. We are able to reproduce the original results exactly for the SCAT and almost exactly for the USPC codes, where we get a Pearson correlation of 0.995 between the original predictions and ours. Although the differences are minor, we use the original (rather then reproduced) predictions in the regression analysis presented below. Since there are different reporting defaults in R and STATA, we redid all regressions in STATA and report the STATA results for better comparison.

The innovation network model. Using matrix notation, the innovation network can be formulated as

$$\hat{\mathbf{p}}_t = \sum_{l=1}^{10} \mathbf{M}_l \mathbf{p}_{t-l}, \tag{A.4}$$

where the i^{th} element of the column vector $\hat{\mathbf{p}}_t$ is the predicted number of patents in class i at time t and $\mathbf{p}_{t-l}$ is the observed number at $t-l$. The matrix $\mathbf{M}_l$ represents the innovation network, which captures knowledge spillovers from a technological class in column j to another in row i at a given time lag l. To define the innovation network $\mathbf{M}_l$, let $H_{pq,l}$ be 1 if there is a citation from patent p to patent q at time lag l and 0 otherwise. The patent-class relationship is defined by the bipartite network $\mathbf{B}_l$, with elements $B_{pi,l} = 1$ if patent p is associated with technology class i and 0 otherwise. A link of the innovation network is then given by

$$M_{ij,l} \equiv \frac{\sum_p \sum_q B_{pi,l} B_{qj,l} \frac{H_{pq,l}}{S_p}}{p_i^{75-84}}, \tag{A.5}$$

where p_i^{75-84} is the total number of patents of domain i applied for in the years 1975 to 1984 and S_p is the total number of citations made by patent p. Eq. (A.5) can be understood as a projection of the patent citation network onto technological classes,

normalized twice: first, by the total number of citations each patent makes, and second, by the total number of patents per technological domain.

Comparing predictions. Acemoglu et al. (2016) construct the matrix $\mathbf{M}_l$ by using citation data from 1975 to 1994 and predict patent applications for the years 1995 to 2004 based on Eq. (A.4). Note that the patenting forecast for a given year uses patenting levels of the preceding ten years. In other words, although the network is calibrated to data prior to 1994, all predictions are essentially one-year ahead. For example, for predicting patenting in 2004, the data used is the 1975–1994 innovation network, and patenting levels between 1994 and 2003.

Since for one-year forecasts the various time series models yield similar performances (see Fig. 1.7), we simply use the geometric random walk to benchmark the innovation network model forecasts with. Analogously, we only use data from 1975 to 1994 to estimate the parameter μ_i, and make one-year ahead predictions in the horizon 1995-2004 while holding μ_i fixed.

Similar to Fig. 1.1 in the main text, Fig. A.8 shows the predicted patenting against the observed patenting for the full network, the external network, and the random walk on a double logarithmic scale. As in the original work, we have made predictions using the *full* (left panel) and *external* (center panel) network. Full network refers to the network model as specified in Eq. (A.4). External network is the same but with the diagonal equal to zero: $M_{ii,l} = 0$ for all i. Thus, in this case patents of a given class i are predicted using patenting data only of other classes $j \neq i$. The center panel reproduces Fig. 7A of the supplementary information in Acemoglu et al. (2016), except that we show technology-specific regression lines instead of a single line obtained from a panel regression (as in Fig. 1.1 in the main text).

Compared to the network models, the random walk predictions lie closer to the 45 degree line, and exhibit less dispersion. The regression lines of the right panel tend to have intercepts closer to zero, slopes closer to one, and a higher R^2, indicating a lower bias and higher precision in the random walk predictions as compared to the network models.

As in Fig. A.8, Fig. A.9 shows actual against predicted patenting activity, but when using the more fine-grained USPC instead of the SCAT scheme. Results are qualitatively similar with the random walk forecasts being more aligned with the 45 degree line and less spread out than the innovation network forecasts. We show further tests on the predictive performance of the models below.

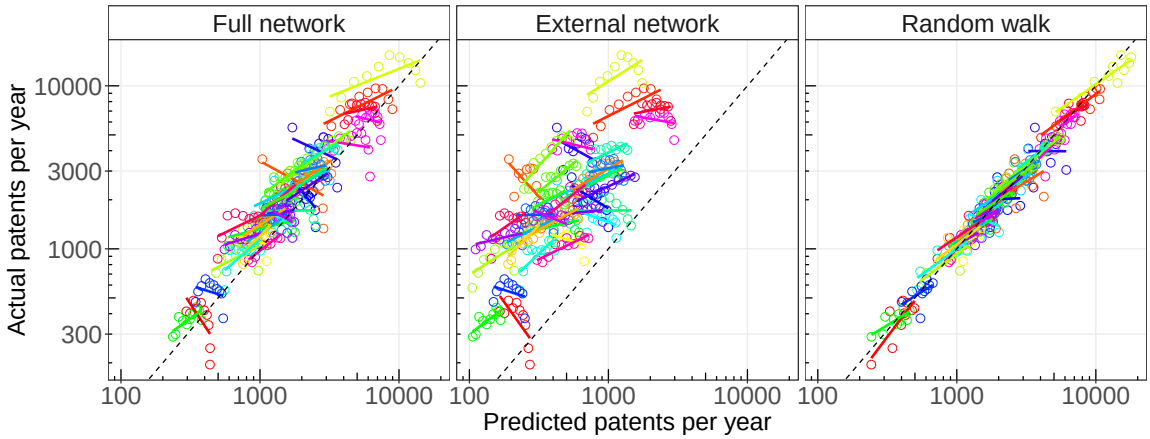

Figure A.8: **Predicted vs. actual annual patenting activity based on SCAT.** The left panel shows predictions obtained from the innovation network as specified in Eq. (A.4). The center panel is the same, but with zero diagonal. The right panel shows predictions from a geometric random walk; Eq. (1.2). Colors correspond to the 36 SCAT technology classes. The dashed line in each panel is the 45 degree line, corresponding to a perfect prediction, and the solid lines are technology-specific regression lines obtained from regressing predicted against observed values.

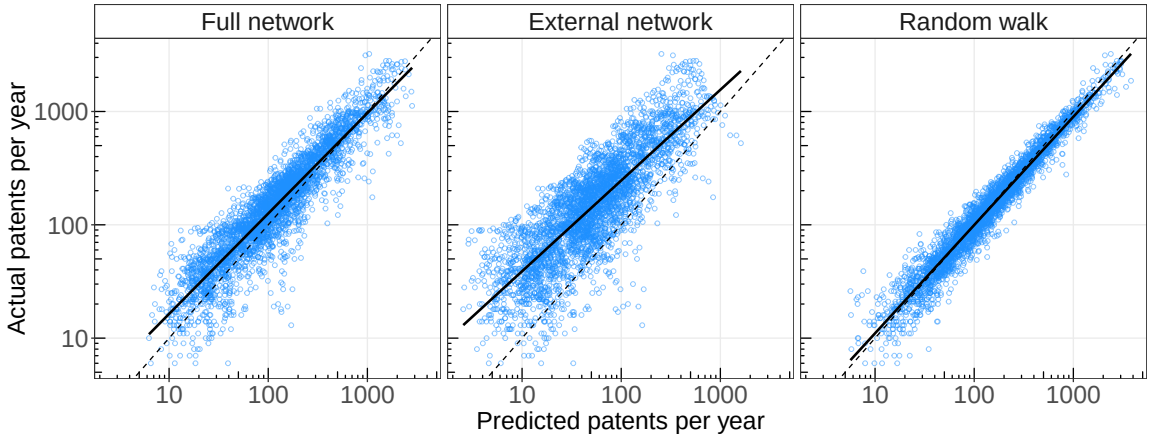

Figure A.9: **Predicted vs. actual annual patenting activity based on USPC.** Actual against predicted patent levels for the full innovation network (left panel), the external innovation network (center panel) and the random walk (right panel) forecasts. The dashed lines depict the 45 degree line. The solid lines are obtained from regressing predicted against observed patenting activity with a panel model.

Prediction bias in the innovation network. An issue when applying the innovation network is that the overall level of the forecasts is affected by two somewhat arbitrary choices: the number of lags, l, in Eq. (A.4), and the normalization constant p_i^{75-94} in Eq. (A.5). Fig. A.10 illustrates how predicted levels change when these choices are made differently.

Fig. A.10(a) shows the total number of predicted patents for varying numbers of lags. Obviously, predicted levels increase if more lags are included and decrease if less lags are used. We observe that adding two more lags has less impact on the predictions

than removing two lags. This has two reasons. First, the number of citations decreases for larger lags. Second, the upward time trend of patent applications implies that adding an extra patenting vector $\mathbf{p}_{t-l}$ for large l in the predictions has comparatively less impact than removing smaller lags.

In Fig. A.10(b) we see how a change in the normalization constant impacts the total number of predicted patent applications. The blue line depicts the baseline scenario where the network is normalized by all patents applied for between 1975 and 1984. The red and green line show how the predictions change if only the first 5 and the first 15 years are used instead, respectively. Unsurprisingly, the predictions increase if less years are used for the normalization and decrease in the other case.

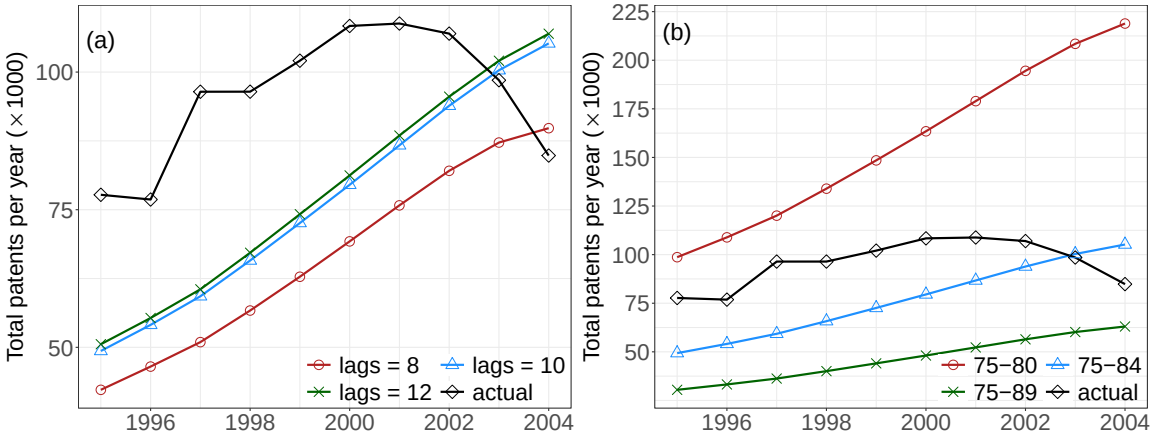

Figure A.10: **Total number of patent applications per year.** The black line indicates the actual number of patents in the data set. (a) The red, blue and the green lines show the total number of predicted patents for 8, 10 and 12 lags, respectively. Note that the blue line is the baseline case presented in Acemoglu et al. (2016). (b) Total number of predicted patents for varying network normalizations: red (all patents in 1975–1980), blue (1975–1984, the baseline) and green (1975–1989). Total numbers of predicted patents are extrapolated from using SCAT forecasts.

Regressions. In Acemoglu et al. (2016) the prediction performance of the network model is evaluated by regressing the logarithm of the observed values on the logarithm of the predictions. This is done both in a pooled model and in a twoway fixed effects

panel model as follows[2]:

$$y_{i,t} = b_0 + b_1 \hat{y}_{i,t} + \epsilon_{i,t}, \tag{A.6}$$

$$y_{i,t} = a_i + \tau_t + b_1 \hat{y}_{i,t} + \epsilon_{i,t}. \tag{A.7}$$

The pooled model is a classical forecast evaluation regression (Mincer & Zarnowitz 1969), where the hypothesis of forecast efficiency can be tested via the joint hypothesis $b_0 = 0, b_1 = 1$. Recognizing that classes are of different sizes and the yearly number of patents may grow in time, Acemoglu et al. (2016) suggest the second specification with individual and time fixed-effects, and focus on the hypothesis $b_1 = 1$.

In the original paper, the pooled regression is estimated via ordinary least squares (OLS) and the twoway model with weighted least squares (WLS) with technology-specific weights, $\sum_{t=1985}^{1994} p_{i,t}$. To enhance comparability, the results below are presented using the same estimators as the original paper.

Table A.3 summarizes the regression results for predicting patenting based on the SCAT (a) and USPC (b) schemes. Columns *Full* and *External* reproduce the results of the original paper based on the full and external network model forecasts. In agreement with Acemoglu et al. (2016), we obtain large positive coefficients b_1 for the pooled regression in Eq. (A.6), ostensibly providing support for network effects. Yet when looking at the regression coefficient of the random walk forecasts, column *GRW*, we obtain a coefficient even closer to 1, explaining almost all the variance. Recall that the random walk assumes that every time series evolves independently, ignoring all network effects. This observation thus calls into question whether the forecasts in Acemoglu et al. (2016) really capture network effects and instead suggests that they may be driven by the persistence of patenting growth.

For the twoway fixed effects specification, Eq. (A.7), the coefficient is substantially lower for the full network model and somewhat larger for the external network model. The random walk coefficient takes on an intermediate value, but is within one standard error of the external network coefficient. The fixed effect results are sensitive to the specification; for example, if the same estimation is done using the USPC classification with 353 categories, or using OLS instead of weighted least squares,

[2]The coefficients from the network model predictions are fairly sensitive with respect to using the log transformation in the regressions. For example, the pooled external network regression yields $b_1 = 3$ in case the data is not log-transformed. In the twoway specification the coefficient is insignificant. When using USPC codes instead of SCAT, coefficients turn negative and insignificant for the twoway model. Random walk forecasts are less sensitive. Coefficients from the random walk regressions are always significant and lie between 0.46 and 0.92 if the data is not log transformed.

the random walk coefficient is larger. In all cases the random walk has a substantially higher R^2 value.[3] We also computed alternative standard metrics for assessing predictions which are based on forecast errors, such as mean absolute (percentage) errors etc. All these metrics strongly favor the random walk over the network model forecasts.

(a) SCAT	*Dependent variable:* $\log(p_{i,t})$					
	Eq. (A.6)			Eq. (A.7)		
	Full	External	GRW	Full	External	GRW
b_1	0.937***	0.786***	0.966***	0.534***	0.851***	0.732***
	(0.040)	(0.102)	(0.012)	(0.110)	(0.163)	(0.051)
b_0	0.687**	2.701***	0.253***			
	(0.293)	(0.645)	(0.087)			
a_i, τ_t	No	No	No	Yes	Yes	Yes
weights	No	No	No	Yes	Yes	Yes
R^2 (within)				0.569	0.571	0.746
R^2 (overall)	0.870	0.559	0.967	0.913	0.606	0.980
Obs.	360	360	360	360	360	360

(b) USPC	*Dependent variable:* $\log(p_{i,t})$					
	Eq. (A.6)			Eq. (A.7)		
	Full	External	GRW	Full	External	GRW
b_1	0.886***	0.801***	0.958***	0.414***	0.305***	0.680***
	(0.016)	(0.026)	(0.004)	(0.085)	(0.088)	(0.029)
b_0	0.755***	1.817***	0.211***			
	(0.081)	(0.112)	(0.021)			
a_i, τ_t	No	No	No	Yes	Yes	Yes
weighted	No	No	No	Yes	Yes	Yes
R^2 (within)				0.290	0.261	0.549
R^2 (overall)	0.846	0.677	0.954	0.855	0.674	0.958
Obs.	3,530	3,530	3,530	3,530	3,530	3,530
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.3: Regression tables for Eqs. (A.6) and (A.7) when using the SCAT classification into 36 categories (a) and when using the USPC classification into 353 categories (b). Columns *Full* refer to full network predictions, *External* to external network predictions and *GRW* to random walk predictions. Time horizon ranges from 1995 to 2004. Standard errors are based on heteroskedasticity and autocorrelation consistent estimation of the covariance matrix.

[3]For the twoway panel model we also report *within* besides *overall* R^2 (for a discussion on goodness-of-fit with panel data see e.g. Verbeek (2017) p.395ff). For all regressions the R^2 (within, between, overall) is largest for the random walk and smallest for the external network predictions.

Forecast encompassing. The results so far show that the network model does not systematically outperform the geometric random walk in predicting patenting levels. However, the network model predictions might contain *some* information that is not captured by the random walk forecasts. A common way to test for relative information contents of competing forecasts is to test for forecast encompassing, typically by regressing the actual values on the values predicted by each of the competing methods. In our case,

$$\log(p_{i,t}) = a_i + \tau_t + \beta_N \log(\hat{p}^N_{i,t}) + \beta_W \log(\hat{p}^W_{i,t}) + \epsilon_{i,t}, \tag{A.8}$$

where $\hat{p}^N_{i,t}$ and $\hat{p}^W_{i,t}$ are the innovation network and the random walk predictions, respectively. The random walk encompasses the network model if $\beta_W = 1$ and $\beta_N = 0$. If both parameters are different from zero, the two forecasts contain complementary information, suggesting that a combination of both forecast methods could improve predictions.

Table A.4 summarizes the results of the forecast encompassing regressions. In each classification system we find a strongly positive coefficient for the random walk model forecasts. For the network model forecasts, in contrast, the coefficients are negative and statistically insignificant in three out of four cases. The coefficient is positive only for the external network using the SCAT classification, and is significant only at one standard deviation.

It is noteworthy that the coefficients β_W of the forecast encompassing test are very similar to the coefficients of the random walk regressions reported in Table A.3. This shows that the random walk coefficient is highly robust against adding the network model forecasts as additional regressor. As indicated by negligible differences between R^2 values of the random walk and the forecast encompassing regressions, adding the network forecasts as regressor does not improve the precision.

	Dependent variable: $\log(p_{i,t})$			
	SCAT		USPC	
Network	Full	External	Full	External
β_N	-0.063	0.212^{*}	-0.060	-0.026
	(0.074)	(0.116)	(0.043)	(0.043)
β_W	0.764^{***}	0.678^{***}	0.699^{***}	0.685^{***}
	(0.069)	(0.065)	(0.032)	(0.030)
a_i, τ_t	Yes	Yes	Yes	Yes
weighted	Yes	Yes	Yes	Yes
R^2 (within)	0.747	0.751	0.550	0.549
R^2 (overall)	0.979	0.967	0.955	0.957
Obs.	360	360	3,530	3,530
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.4: Coefficients obtained from forecast encompassing. Columns *Full*: full network versus random walk predictions. Columns *External*: external network versus random walk predictions. Time horizon ranges from 1995 to 2004. Standard errors are based on heteroskedasticity and autocorrelation consistent estimation of the covariance matrix.

Network randomization. Acemoglu et al. (2016) found that upstream patenting is more predictive than downstream patenting, indicating that the network specification in Eq. (A.4) is important. To test whether the empirical network is important for obtaining the positive relationship between observed and predicted values (Table A.3), we redo the predictions with a randomized version of the network model.

We apply the following randomization procedure: First, for a row i in the innovation network matrix $\mathbf{M}_l$, Eq. (A.4), we sample all off-diagonal elements with equal probability and without replacement. This keeps the total link weight (strength) constant but randomizes the upstream neighborhood of class i. We repeat this for every row in the matrix and all time lags l. The randomized network is then used to predict patenting applications according to Eq. (A.4)). To test the predictive performance of the randomized network model we again run the regressions shown in Eqs. (A.6) and (A.7).

We repeat this procedure 10,000 times and show the histogram of the estimated coefficients of interest, $\hat{b}_1$, in Fig. A.11. For the pooled regression models, we find that the empirically estimated coefficients lie in the upper tail of the randomized network coefficient histogram. But it is also evident that even for the randomized innovation networks, the pooled regression always yields large coefficients which are

fairly close to one. For the full network the mean coefficient for the randomized networks is roughly 0.925, whereas the coefficient with the true network is only 0.937. Similarly, for the external network the mean coefficient for the randomized networks is 0.786, whereas the coefficient for the true network is 0.778. Thus while the increase is statistically significant, it is not economically significant. This result suggests that, given the persistent time trends in patent levels of technology classes, the use of pooled regressions for evaluating predictive power is questionable.

For the twoway specification using the full network the coefficient for the true network is in the middle of the distribution of that of the random networks, indicating no predictive power. In contrast, for the external network the behavior changes dramatically. In this case the estimated coefficients are distributed across a far wider range, varying roughly from -3.5 to 2.8. The estimated coefficient of the true network lies in the right hand tail (98%-quantile), suggesting that there is a statistically and economically significant effect.

Applying the same procedure to the USPC classification does not yield any further support for the network forecasts. Here, the empirical estimates even lie to the left of the histograms, except for the external network-twoway specification where the estimate is again in the upper tail (96%-quantile).

These results suggest that the coefficient of interest b_1 is not very sensitive with respect to the network structure, except for the twoway specification when evaluating the external network forecasts. Here, we find the empirical estimate to be in the upper tail of the coefficient distribution of randomized network, potentially indicating some signal in the network model forecasts.

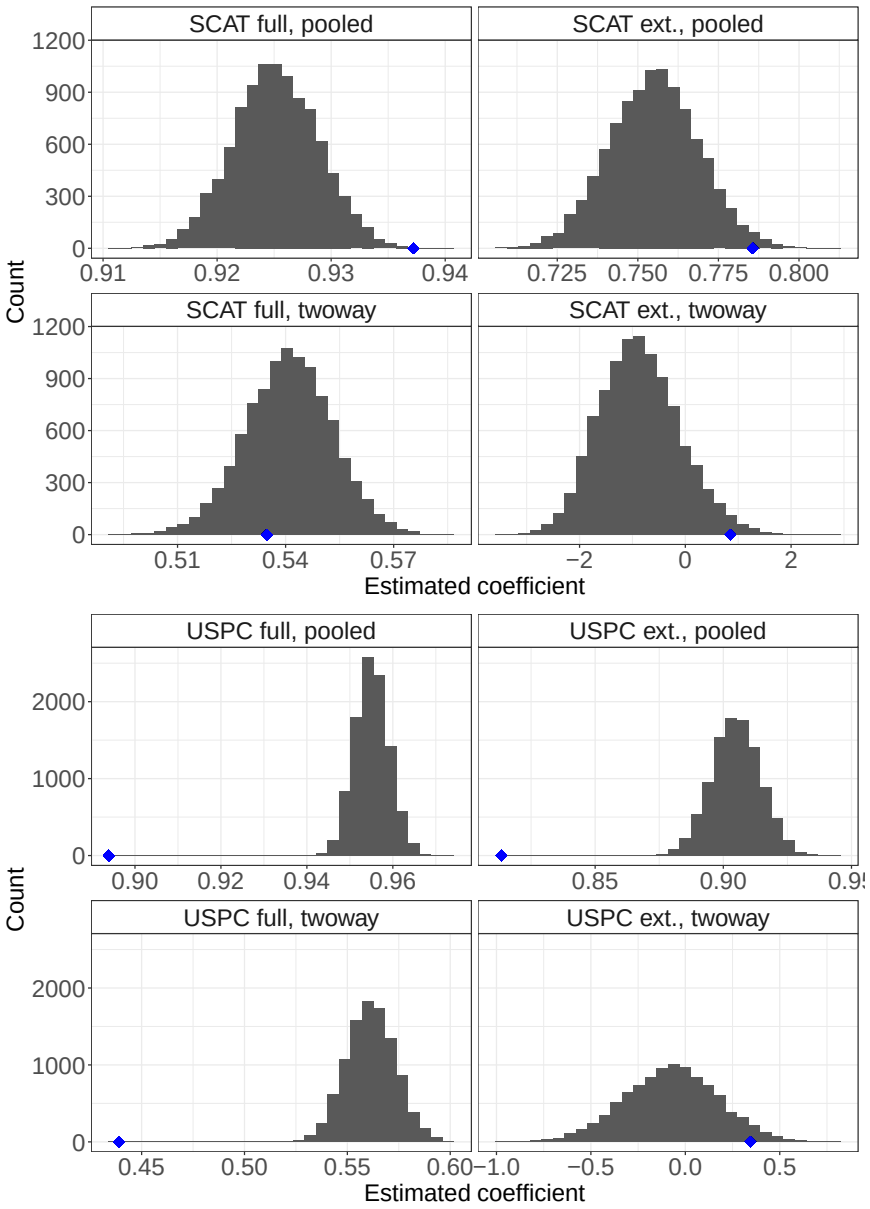

Figure A.11: Histograms of regression coefficients for 10,000 network randomizations based on the SCAT (upper four panels) and USPC (bottom four panels) classification. *Pooled* refers to Eq. (A.6), *twoway* to Eq. (A.7). *Full* refers to predictions which include the diagonal elements, whereas *ext.* refers to predictions with zero diagonal. As in Table A.3, regression coefficients are obtained by ordinary least squares for the pooled model and by weighted least squares for the twoway specifications. The blue diamonds represent the values as reported in Acemoglu et al. (2016).

Appendix B

Appendix to Chapter 2

B.1 Data sources and preprocessing

Chapter 2 uses the same database as described in Appendix A.1, but also includes patent citation data. Moreover, we show further robustness tests below that use the International Patent Classification (IPC). (US Patents are classified in two different classification systems: CPC and IPC, and up to 2015, they were also classified in the US Patent Classification System (USPC)).

IPC classifications and patent dates are retrieved from Google Patents Public Data, provided by IFI CLAIMS Patent Services, (accessed June 2018). For the citation data, we merge the citation database of Kogan et al. (2017) with the citation data collected by Lafond & Kim (2019) and with the database provided by the USPTO (patentsview.org/download, accessed: 14/05/2019).

Fig. B.1a) shows total patenting rates in the US from 1836 to 2019 for distinct classification codes of patents. While we find that CPC codes for almost every patent back to 1836 (the solid and the dotted line are almost indistinguishable in the plot), reclassification of IPC codes seems to be less rigorous for patents before 1920. For 2018 and 2019 the data is not yet entirely updated, indicated by the sharp drop in patenting rates at the end of the time series. We therefore exclude these years in further analysis. Fig. B.1b) shows the citations counts per year. The figure illustrates that reliable data on citations made is available only from 1947 on, but we have information on citations received by earlier patents.

We map patents onto their CPC and IPC codes and eliminate duplicates if the same patent was classified multiple times with the same 3- or 4-digits code. We only include technological classes which were at least ten times assigned to patents between 1836 and 2000. This removes no class at all in the CPC 3-digits case and only a single class in the CPC 4-digits case (A61P with only a single patent up to 2000).

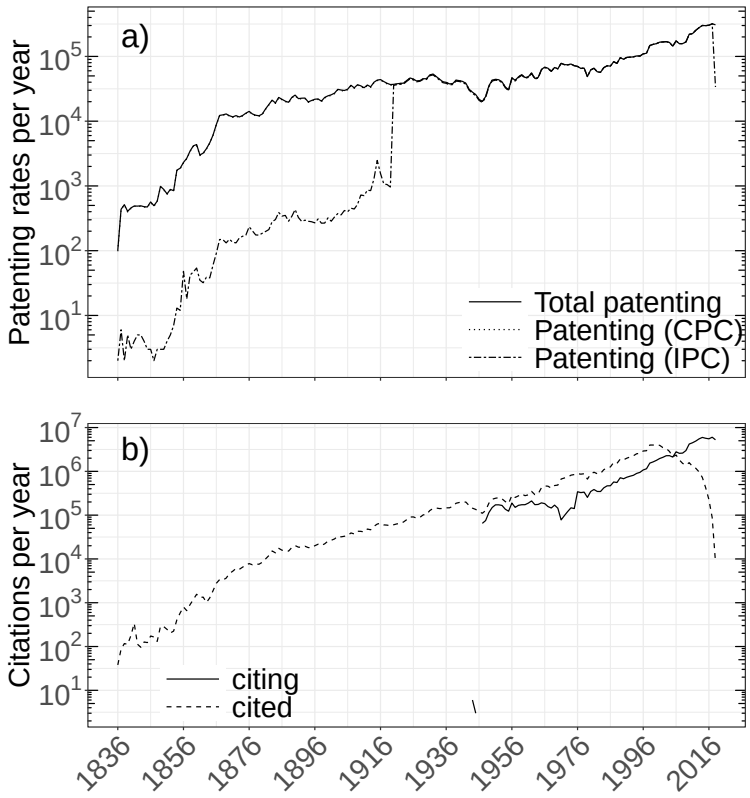

Figure B.1: a) Yearly time series of US patents granted between 1836 and 2019. The solid line is the number of total granted patents, the dotted line is the number of patents with CPC attributes and the dotdashed line the number of patents having IPC attributes. b) Time series of citations per year in the data set. *Citing* refers to the number of citations made in a given year and *cited* refers to the number of citations received.

The cleaning is more effective in the IPC regimes where there are several classes which are very infrequently used. Introducing this threshold eliminates 143 IPC 3-digits and 389 IPC 4-digits classes. For the CPC system, we also eliminate all Y-codes, which are tags rather than separate technology classes, and the "2000 series" related to indexing schemes codes (EPO & USPTO 2017). Table B.1 summarizes some basic statistics of the used data, after excluding the years 2018 and 2019.

Classification	# classes	# observations	# patents
CPC 3-digits	123	12,912,549	9,774,895
CPC 4-digits	630	14,422,099	9,774,895
IPC 3-digits	121	12,000,519	8,449,988
IPC 4-digits	633	13,697,093	8,449,720

Table B.1: Descriptive statistics of used data after initial cleaning.

B.2 Significance sampling

We obtain a temporal network for each of the four classification schemes. To avoid links between technological classes which one would expect by chance, we apply the significance sampling procedure (Alstott et al. 2017) described in Chapter 2. In Fig. B.2 we show the evolution of network density, i.e. the number of existing links divided by the number of all possible links (N^2), It becomes evident that the density is reduced substantially after applying the significance sampling procedure, but is also increasing in time. Unsurprisingly, the network density is higher when using the more aggregate 3-digits classification.

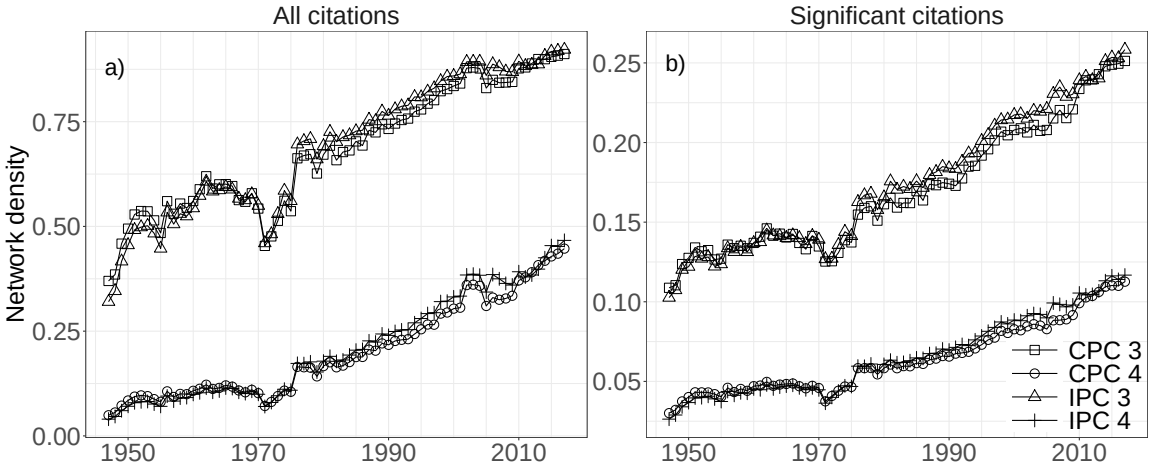

Figure B.2: a) The network densities, number of existing links divided by the number of all possible links, over time for different classification schemes when all citations are included. b) The network densities after removing non-significant links.

B.3 Temporal network stability

As a measure of temporal stability, Fig. B.3 shows the average cosine similarity between a row i of the technology matrix at t and $t-l$, or in more mathematical terms

$$\text{similarity}^{\text{av.}}_{t,t-l} = \frac{1}{N}\sum_{i=1}^{N} \frac{\sum_j W_{ij,t} W_{ij,t-l}}{\sqrt{\sum_j W^2_{ij,t}}\sqrt{\sum_j W^2_{ij,t-l}}}. \tag{B.1}$$

We observe that the network tends to be fairly stable in time, i.e. the future technological dependence of a given technology is likely to be similar to its current one. After the temporal stability of the network was increasing for most parts of the second half of the last century, it became weaker in the last 20 years. As one would expect,

we find that the average similarity decreases for higher lags, although there is still a pronounced positive relationship between technological dependences over a 10-year horizon. A qualitative difference between 3-digits and 4-digits classification schemes is apparent. The networks based on the more aggregated 3-digits classifications exhibit higher correlations in time for all shown lag choices. Moreover, the differences in temporal correlations between different time lags become less pronounced. This suggests that it is important to use more aggregated technology classification if the research is interested on the temporal evolution and structural change of the technological ecosystem.

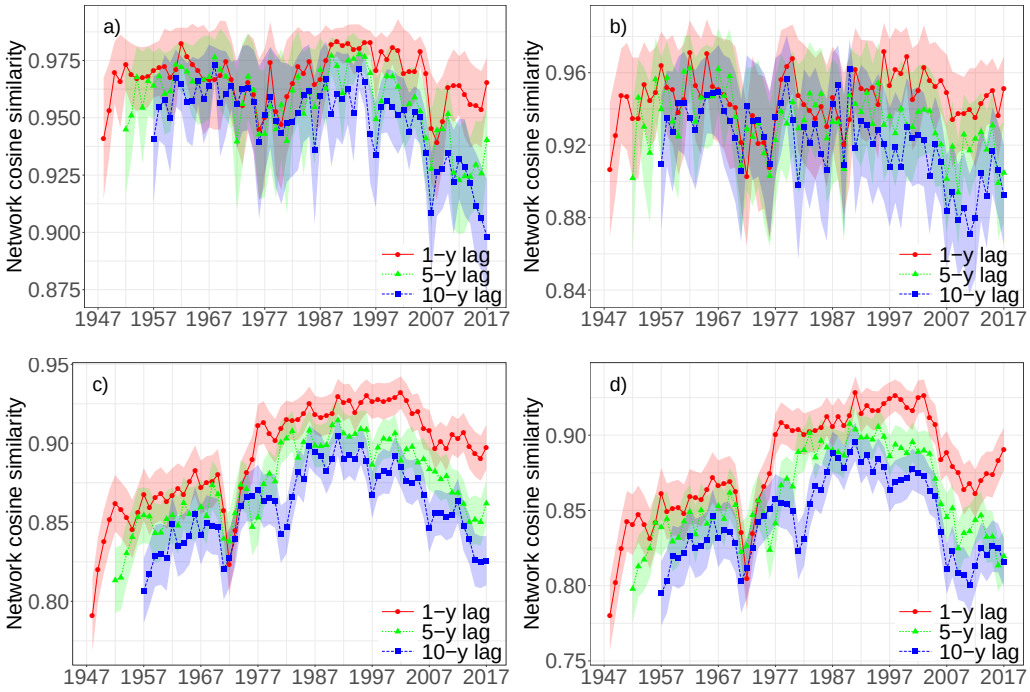

Figure B.3: Average temporal cosine similarity of technology networks over time. The shaded area shows the range of the average plus/minus twice the standard errors. The results are shown for networks based on the a) 3-digits CPC codes, b) 3-digits IPC codes and c) 4-digits CPC codes and d) 4-digits IPC codes.

B.4 Nearest neighbor growth correlations

In Chapter 2 we have shown that the technology network is highly assortative with respect to growth rates, suggesting that patenting growth rates are similar if the technologies are connected. Fig. B.4 plots the average nearest neighbor growth (ANNG) rates correlation over time for the differently aggregated networks. The positive assortativity pattern is more noisy in the more aggregated networks, panel a) and b),

but still pronounced for most of the cases. The average nearest neighbor growth rates correlation is relatively insensitive with respect to different growth rate lengths choices for the more aggregated networks, again pointing at the importance of using more granular networks for discerning temporal effects in the technological evolution.

Table B.2 summarizes the nearest neighbor growth correlations for different classifications and and lags. We see that the correlation magnitude is similar for different classification schemes, but systematically less noisy when using the more granular 4-digits classification. Data aggregation often reduces noise by averaging out extrema. Thus, it is not obvious a priori that the more fine-grained network representations exhibit less variation in their average nearest neighbor growth rates correlations.

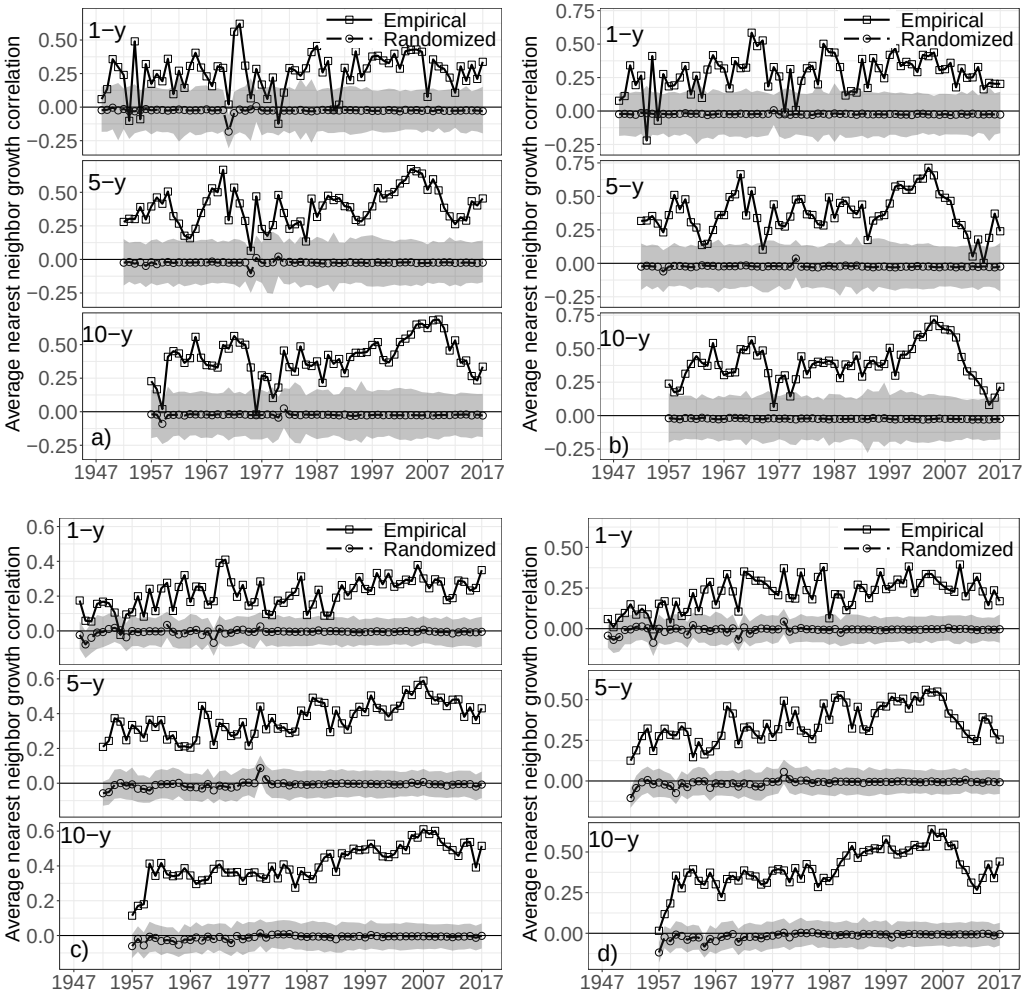

Figure B.4: Average nearest neighbor growth rates correlation across time for networks based on the a) 3-digits CPC codes, b) 3-digits IPC codes, c) 4-digits CPC codes and d) 4-digits IPC codes. The shaded area is the 5-95% inter-quantile range of results based on 1,000 randomized networks (described in Chapter 2).

	Lag	CPC 3	CPC 4	IPC 3	IPC 4
Time average	1-year	0.26	0.21	0.28	0.22
	5-year	0.39	0.37	0.38	0.35
	10-year	0.40	0.41	0.39	0.40
Time std. deviation	1-year	0.15	0.09	0.14	0.10
	5-year	0.14	0.10	0.15	0.11
	10-year	0.15	0.10	0.15	0.12

Table B.2: Average nearest neighbor growth rate correlations. Time average and time standard deviation of the average nearest neighbor growth rates correlation for distinct technological classification schemes.

B.5 Econometric estimation

B.5.1 Econometric network model and SAR

In Chapter 2 we have presented results based on the econometric model

$$g_{i,t} = \frac{a_i}{1 - \beta W_{ii,t}} + \frac{\beta}{1 - \beta W_{ii,t}} \sum_{j=1}^{N} W_{ij,t} g_{j,t} (1 - \delta_{ij}) + \epsilon_{i,t}, \tag{B.2}$$

where $\epsilon_{i,t} \sim N(0, \sigma^2)$. Note that the econometric model is a generalization of the spatial autoregressive (SAR) model for panel data. In case of a zero diagonal matrix, $W_{ii} = 0$, the model reduces to the fixed-effects SAR with time-varying spatial component

$$g_{i,t} = a_i + \beta \sum_{j=1}^{N} W_{ij,t} g_{j,t} + \epsilon_{i,t}, \tag{B.3}$$

discussed by Wang & Yu (2015). Moreover, if the network is fixed in time, $\mathbf{W}_t = \mathbf{W}$, this model reduces to the standard fixed-effects SAR which can be estimated as outlined in Elhorst (2014).

B.5.2 Time-varying network model

To test the robustness of the results in Chapter 2, we estimate different variations of these econometric models. First, we calibrate the model as discussed in Chapter 2 to all four classification schemes and report the results in Table B.3. The results are qualitatively similar for different aggregations.

The econometric estimation also allows use to make the magnitude of network effects in patenting growth rates explicit. Let us define the average direct and average

total impacts of research effort growth on patenting growth by

$$\bar{I}_{direct,t} := \alpha \frac{\sum_{i=1}^{N} L_{ii,t}}{N}, \tag{B.4}$$

$$\bar{I}_{total,t} := \alpha \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} L_{ij,t}}{N}, \tag{B.5}$$

respectively, yielding the relative network impact of

$$I_{network,t} := 1 - \frac{\bar{I}_{direct}}{\bar{I}_{total}} = 1 - \frac{\sum_{i=1}^{N} L_{ii,t}}{\sum_{i=1}^{N} \sum_{j=1}^{N} L_{ij,t}}. \tag{B.6}$$

Fig. B.5 shows the relative network impacts $I_{network,t}$ over time for all four classifications. The network effects are growing in time which is not surprising given the increase in network density we have observed. The figure shows that we would estimate higher network impacts in a more aggregated network, i.e. when based on 3-digits classifications. We also find that network effects are slightly larger for IPC codes than for CPC codes, if compared on the same digit-level. But the general trend of rising relative network impacts over time holds for all four classification schemes.

Dependent variable: patenting growth

	CPC3			*CPC4*			*IPC3*			*IPC4*		
	(1-y)	(5-y)	(10-y)	(1-y)	(5-y)	(10-y)	(1-y)	(5-y)	(10-y)	(1-y)	(5-y)	(10-y)
average $\hat{a}_i$	0.01	0.03	0.05	0.01	0.03	0.04	0.01	0.02	0.05	0.01	0.03	0.05
$\hat{\beta}$	0.92***	0.94***	0.94***	0.85***	0.93***	0.94***	0.92***	0.94***	0.92***	0.87***	0.93***	0.93***
	(0.006)	(0.010)	(0.015)	(0.006)	(0.006)	(0.008)	(0.006)	(0.011)	(0.019)	(0.006)	(0.007)	(0.009)
$\hat{\sigma}^2$	0.03	0.08	0.13	0.11	0.20	0.30	0.03	0.09	0.14	0.11	0.23	0.31
Log-likelihood	2,525	-277	-364	-13,387	-5,474	-3,625	2,867	-384	-395	-13,431	-6,049	-3,677
Observations	8,610	1,720	858	43,651	8,734	4,348	8,349	1,669	834	42,688	8,528	4,427

Note: ***p< 10^{-16}

Table B.3: Results from estimating the econometric network model presented in Chapter 2 (Eq. (B.2)) for different classification schemes. The results are shown for 1-, 5- and 10-year growth rates.

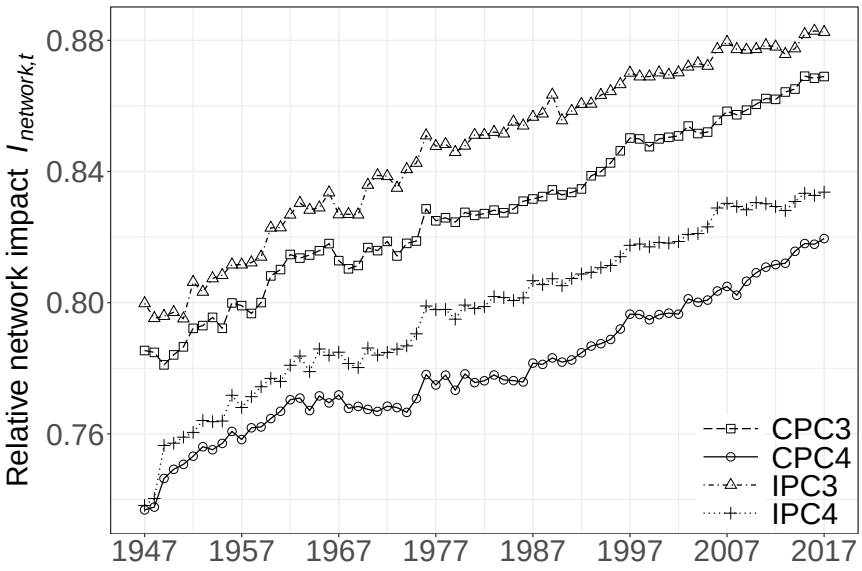

Figure B.5: Relative network impacts over time for different classification schemes. Results are based on 1-year growth rates.

B.5.3 Static network model

We next investigate how the estimated parameters change if we use fixed network specifications. Patent networks are naturally dynamic, but the vast majority of the spatial econometrics literature does not allow for time-varying networks. We thus test whether our results also hold if we keep the technology network fixed.

We do this by estimating the model for every network snapshot separately, i.e. by setting $\mathbf{W}_t = \mathbf{W}$ in Eq. (B.2) where $\mathbf{W}$ represents a particular network in the sample. We then check how the parameter β changes if we use a different network snapshot instead. It should be pointed out that there might be an endogeneity issue if more recent networks are included in the model, since the network formation itself could depend on the growth of technological classes. Moreover, this estimation uses future information on the right-hand side of Eq. (B.2) to explain past left-hand side variables, for example, when the network of 2017 is used as $\mathbf{W}$. Nevertheless, we estimate the model with a fixed, potentially future, network to see how results change from the time-varying to the static specification.

In Fig. B.6 we plot the coefficient β as a function of the used network $\mathbf{W}_t$. We show the results for CPC 3-digits and CPC 4-digits classes only, since results are similar for the IPC counterparts. The figure depicts qualitatively different results for the two aggregation levels. For the CPC 3-digits case, we find that β exhibits a slight downward trend as a function of the time index. In the CPC 4-digits case, on the

other hand, the coefficient is smaller in earlier and larger for more recent networks. Additionally, from the mid-70s on the parameter seems to stabilize around the value of 0.84 which we have estimated in the original time-varying specification.

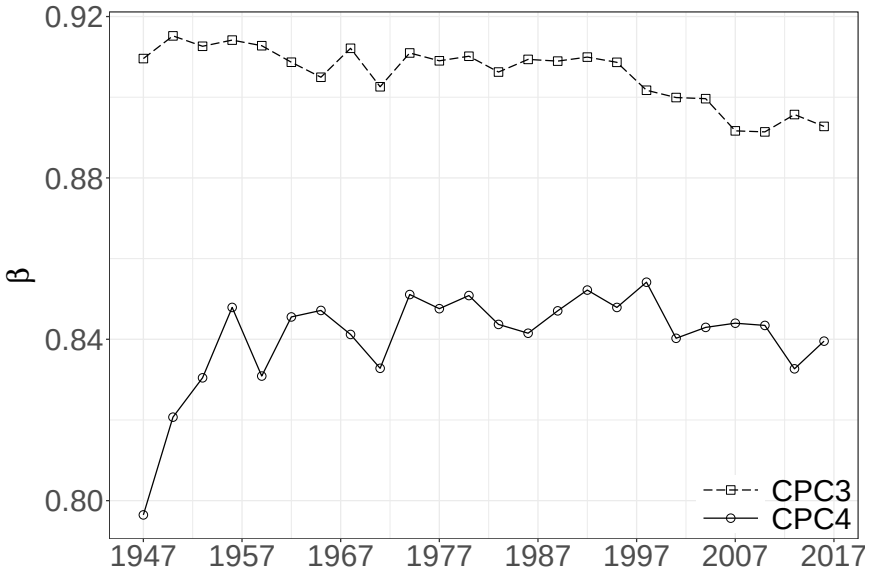

Figure B.6: Parameter β obtained from estimating Eq. (B.2), but with a fixed instead of a time-varying network. The x-axis indicates the year of which the network snapshot was obtained from. For example, at 1947 the model was estimated with using $\mathbf{W}_t = \mathbf{W}_{1947}$ for all t. Results are based on 1-year growth rates.

B.5.4 Time-varying spatial autoregressive model

For comparison we also estimated a standard SAR model (zero-diagonal) with time-varying network component (Eq. (B.3)). The SAR model is convenient for demonstration purposes, since extensive literature renders it a "standard" model for studying network effects. Due to its simpler form, it is also straightforward to include further regressors to check for robustness. We therefore estimate Eq. (B.3) also with additional regressors such as time effects and technological class sizes.

The estimation of the SAR model simplifies in comparison to the full network model of Eq. (B.2). To estimate Eq. (B.3) including exogenous regressors, we stack all elements $g_{i,t}$ into the NT-dimensional vector $\bar{\boldsymbol{g}} := (g_{1,1}, g_{2,1}, ..., g_{N,1}, g_{1,2}, ..., g_{N,T})$. In the same manner we obtain the vector $\bar{\boldsymbol{\epsilon}}$. By defining the T-dimensional vector of ones $\mathbf{1}$ and the block-diagonal matrix $\mathbf{W} := diag(\mathbf{W}_1, ..., \mathbf{W}_T)$, we can rewrite Eq. (B.3) into matrix notation

$$\bar{\boldsymbol{g}} = (\mathbf{1} \otimes \mathbb{I}_N)\boldsymbol{a} + \beta \mathbf{W}\bar{\boldsymbol{g}} + \bar{\mathbf{X}}\boldsymbol{\delta} + \bar{\boldsymbol{\epsilon}}, \tag{B.7}$$

where $\bar{\mathbf{X}}$ is the $NT \times K$ matrix of exogenous regressors and $\boldsymbol{\delta}$ the corresponding K-dimensional vector of coefficients. After subtracting the time average, Eq. (B.7) simplifies to

$$\boldsymbol{g} = \beta \mathbf{W}\boldsymbol{g} + \mathbf{X}\boldsymbol{\delta} + \boldsymbol{\epsilon}, \tag{B.8}$$

where $\boldsymbol{g}$, $\mathbf{X}$ and $\boldsymbol{\epsilon}$ are the time-demeaned versions of $\bar{\boldsymbol{g}}$, $\bar{\mathbf{X}}$ and $\bar{\boldsymbol{\epsilon}}$, respectively. The log-likelihood of the time-demeaned model then reads

$$\text{LogL} = -\frac{NT}{2}\ln(2\pi\sigma^2) + \sum_{t=1}^{T} \ln|\mathbb{I} - \beta \mathbf{W}_t| - \frac{1}{2\sigma^2}\boldsymbol{\epsilon}^\top\boldsymbol{\epsilon}, \tag{B.9}$$

with $\boldsymbol{\epsilon} = (\mathbb{I}_{NT} - \beta\mathbf{W})\boldsymbol{g} - \mathbf{X}\boldsymbol{\delta}$. The parameter β is obtained by solving

$$\begin{aligned}\underset{\beta}{\text{argmax}} \quad & constant + \sum_{k=1}^{NT}\ln(1-\beta e_k) \\ & - \frac{NT}{2}\ln(\boldsymbol{g}^\top\boldsymbol{g} - 2\beta\boldsymbol{g}^\top\mathbf{W}\boldsymbol{g} - \beta^2\boldsymbol{g}^\top\mathbf{W}^\top\mathbf{W}\boldsymbol{g} - \mathbf{A}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{A}\boldsymbol{g},\end{aligned}$$

where $\mathbf{A} := (\mathbb{I}_{NT} - \beta\mathbf{W})$ and e_k represents eigenvalues of matrix $\mathbf{W}$. Under standard regularity assumptions, we derive the variance-covariance matrix by inverting the Fisher information matrix which yields

$$var(\sigma^2, \beta, \boldsymbol{\delta}) = \begin{bmatrix} \frac{NT}{\sigma^4} & -\mathbb{E}\left[\frac{\partial^2 LogL}{\partial\sigma^2\partial\beta}\right] & \mathbf{0}_K^\top \\ . & -\mathbb{E}\left[\frac{\partial^2 LogL}{\partial\beta^2}\right] & \frac{1}{\sigma^2}\mathbf{X}^\top\mathbf{W}\mathbf{A}^{-1}\mathbf{X}\boldsymbol{\delta} \\ . & . & \frac{1}{\sigma^2}\mathbf{X}^\top\mathbf{X} \end{bmatrix}^{-1}, \tag{B.10}$$

where

$$-\mathbb{E}\left[\frac{\partial^2 LogL}{\partial\delta^2}\right] = \text{trace}(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}\mathbf{W}) + \tag{B.11}$$

$$\frac{1}{2\sigma^2}\text{trace}\left([\mathbf{A}^{-1}\mathbf{X}\boldsymbol{\delta}\boldsymbol{\delta}^\top\mathbf{X}^\top(\mathbf{A}^\top)^{-1} + \sigma^2\mathbf{A}^{-1}(\mathbf{A}^\top)^{-1}]\mathbf{W}^\top\mathbf{W}\right) \tag{B.12}$$

and

$$-\mathbb{E}\left[\frac{\partial^2 LogL}{\partial\sigma^2\partial\beta}\right] = \frac{1}{\sigma^2}\text{trace}\left(\boldsymbol{\delta}^\top\mathbf{X}^\top\mathbf{W}\mathbf{A}^{-1}\mathbf{X}\boldsymbol{\delta}\right) + \tag{B.13}$$

$$\text{trace}\left(\frac{1}{\sigma^2}\mathbf{A}^\top\mathbf{W}[\mathbf{A}^{-1}\mathbf{X}\boldsymbol{\delta}\boldsymbol{\delta}^\top\mathbf{X}^\top(\mathbf{A}^\top)^{-1} + (\mathbf{A}^\top)^{-1}(\mathbf{A}^\top)^{-1}]\right). \tag{B.14}$$

$\mathbf{0}_K$ denotes a K-dimensional vector of zeros. Matrix elements indicated by the dots are not reported since they are trivially filled due to symmetry.

In Table B.4 we show the results of estimating the SAR with time-varying network component for the CPC 3-digits, columns (a) to (e), and CPC 4-digits aggregation,

columns (f) to (j). We find that the network coefficient β is estimated to be slightly larger, if the more aggregated networks are used. We also include time effects in the regression which reduces the network parameter in the CPC 4-digits case. When estimating the model for higher growth rate lags (not shown here), including time effects has less impact on the coefficient. The coefficient is fairly robust against adding further regressors such as including logged patenting rates, columns (c) and (h), logged cumulative patent stock, (d) and (i), or logged discounted cumulative patent stock, (e) and (j), where we have used a 15% yearly discount factor as suggested by Hall et al. (2005). The exogenous regressors are highly significant and have a negative sign, indicating that higher patenting rates and larger knowledge stocks at time point t come with smaller patenting rate growth in the following year. The average magnitude of the fixed effects increases as a consequence of including exogenous regressors.

Dependent variable: patenting growth

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
	CPC3	CPC3	CPC3	CPC3	CPC3	CPC4	CPC4	CPC4	CPC4	CPC4
average $\hat{a}_i$	0.010	0.023	0.435	0.407	0.176	0.015	0.024	0.571	0.366	0.131
$\hat{\beta}$	0.830***	0.492***	0.496***	0.497***	0.494***	0.755***	0.471***	0.464***	0.471***	0.470***
	(0.009)	(0.018)	(0.018)	(0.018)	(0.018)	(0.008)	(0.109)	(0.010)	(0.108)	(0.108)
Patent rates			−0.070***					−0.130***		
			(0.003)					(0.002)		
Patent stock				-0.039***					−0.042***	
				(0.004)					(0.003)	
Discount. pat. stock					−0.033***					−0.036***
					(0.004)					(0.003)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time effects	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Log-likelihood	2,502	2,699	2,890	2,742	2,735	-12,945	-12,493	-11,027	-12,423	-12,433
Observations	8,461	8,461	8,461	8,461	8,461	42,403	42,403	42,403	42,403	42,403

Note: ***p< 10^{-16}

Table B.4: Results from estimating the SAR with time-varying network component and additional exogenous regressors (Eq. (B.7)) for different classification schemes. The first five columns (a-e) show the estimated parameters for the CPC 3-digits aggregation and the last five columns (f-j) for the CPC 4-digits aggregation. The results are based on 1-year growth rates.

B.6 Predicting innovation dynamics

In Chapter 2 we have shown that predictions of patenting rates in technological classes can be significantly improved if network effects are taken into account. In this appendix we present further details on predicting patenting rates with the introduced model. Results shown in Chapter 2 are based on the CPC 4-digits codes. Here, we show how the model predictions change when using alternative aggregation

levels. Moreover, we discuss further aspects of the presented results and investigate alternative forecast benchmarks.

B.6.1 ARIMA forecasts

We first focus on the forecasts obtained from the ARIMA models which we state again for the ease of reference:

$$g_{i,t} = \nu_i + \sum_{s=0}^{p} \phi_{i,s} g_{i,t-s} + \sum_{s=0}^{q} \psi_{i,s} u_{i,t-s} + u_{i,t}, \tag{B.15}$$

where $\phi_{i,0} = \psi_{i,0} = 0$. As indicated in Chapter 2, we have fitted every (p, q) combination up to order (5,5) of the ARIMA(p,1,q) model for each (logged) time series in the training set (1947-1987) and used the parameter fits to forecast the time series in the validation set (1988-2002). We then refitted the ARIMA(p,1,q) which yielded the smallest median absolute percentage error in the validation set to the data up to 2002 to forecast the test set from 2003 to 2017. Table B.5 gives an overview of the best (p,q) combinations for different aggregation methods. In the CPC 4-digit case discussed in Chapter 2, almost 25% of model specifications use less than two AR and MA terms, with the the geometric random walk ($p = q = 0$) alone yielding the best forecasts already in 12% percent of all cases. But we also find time series where higher-order models are preferred. Similarly for alternative aggregations, the geometric random walk model is the most frequent choice.

CPC3		MA order q					
		0	1	2	3	4	5
AR order p	0	0.13	0.02	0.02	0.03	0.03	0.02
	1	0.04	0.00	0.01	0.00	0.02	0.02
	2	0.05	0.01	0.02	0.02	0.01	0.06
	3	0.02	0.02	0.04	0.02	0.03	0.02
	4	0.03	0.02	0.01	0.02	0.02	0.06
	5	0.02	0.00	0.02	0.03	0.05	0.04

IPC3		MA order q					
		0	1	2	3	4	5
AR order p	0	0.10	0.03	0.04	0.03	0.03	0.02
	1	0.01	0.03	0.02	0.00	0.03	0.03
	2	0.08	0.00	0.02	0.00	0.01	0.04
	3	0.03	0.00	0.02	0.04	0.02	0.03
	4	0.04	0.03	0.01	0.03	0.03	0.03
	5	0.01	0.05	0.02	0.02	0.07	0.05

CPC4		MA order q					
		0	1	2	3	4	5
AR order p	0	0.12	0.04	0.03	0.03	0.02	0.03
	1	0.06	0.02	0.01	0.02	0.02	0.04
	2	0.02	0.02	0.02	0.02	0.02	0.03
	3	0.03	0.01	0.02	0.01	0.03	0.03
	4	0.02	0.02	0.02	0.02	0.03	0.03
	5	0.02	0.04	0.02	0.02	0.03	0.02

IPC4		MA order q					
		0	1	2	3	4	5
AR order p	0	0.13	0.04	0.02	0.02	0.02	0.04
	1	0.05	0.02	0.01	0.02	0.02	0.04
	2	0.02	0.02	0.02	0.02	0.03	0.03
	3	0.02	0.01	0.01	0.03	0.02	0.03
	4	0.03	0.01	0.01	0.02	0.03	0.03
	5	0.04	0.04	0.02	0.03	0.04	0.03

Table B.5: Distribution of (p, q) pairs which minimize the median absolute percentage error in the validation set for the ARIMA(p,1,q) forecasts. p is the number of autoregressive terms and q the number of moving average terms in the ARIMA model. For example, in the CPC3 case we have 0.15 for the combination $(p, q) = (0, 0)$, meaning that for 13% of all cases the geometric random walk with drift yielded the best predictions in the validation set.

B.6.2 Network model forecasts

To obtain the unconditional forecasts, a similar model selection procedure is applied. Here, we first calibrate the network model to the data in the training set and choose, as discussed in Chapter 2, the parameter k' such that the median absolute percentage error is minimized in the validation set. The model is then refitted to the data, including training and validation set (1947-2002), to predict patenting in the test set from 2003 to 2017 with given k'. When doing this for the four different technology aggregations, we obtain the following results: $k' = 9$ for CPC 3-digits, $k' = 3$ for CPC 4-digits, $k' = 6$ for IPC 3-digits and $k' = 2$ for IPC 4-digits. Thus, the applied model selection procedure always prefers $k' > 0$, i.e. the procedure always suggests to include network effects in the predictions.

B.6.3 Performance evaluation

As in Chapter 2, we compare forecasts based on the predictability gain measure

$$PG1_{i,t} = \frac{|P_{i,t} - \hat{P}_{i,t}^{ARIMA}| - |P_{i,t} - \hat{P}_{i,t}^{network}|}{|P_{i,t}|}, \tag{B.16}$$

i.e. taking the difference between ARIMA and network model absolute percentage errors (APE) for every predicted value. We also check the predictability gains based on the absolute errors (AE) instead of the absolute percentage errors by computing

$$PG2_{i,t} = |P_{i,t} - \hat{P}_{i,t}^{ARIMA}| - |P_{i,t} - \hat{P}_{i,t}^{network}|. \tag{B.17}$$

We then take averages, standard deviations, medians and interquartile ranges for every year and plot the results for all four aggregation schemes in Figs. B.7 to B.10. In every figure, panel a) shows the average predictability gains based on APE, Eq. (B.16). Panel b) gives the corresponding yearly median values and interquartile ranges of the APE-based predictability gains. Averages and standard errors of AE predictability gains, Eq. (B.17), are shown in panel c) and the corresponding medians and interquartile ranges are visualized in panel d). We discuss the results for each aggregation scheme separately.

CPC 3-digit codes (Fig. B.7): Similarly as the results shown in Chapter 2, predictions can be substantially improved in the conditional forecasts where the growth rates of a focal technology's neighborhood is known. This holds true for averages and medians, regardless of the considered performance metric. In contrast to the result in

Chapter 2, we find that the unconditional forecasts are not able to beat the ARIMA model systematically. The main reason why in this scenario the unconditional predictions underperform is that the optimal tuning parameter k' is not stable between the validation and test set. While a large $k' = 9$ is chosen based on the validation set forecasts, this choice is suboptimal when predicting patenting rates in the test set. This is largely a result of temporal instability of coefficients. Estimating the model from earlier years yields higher fixed effects and a slightly lower β than when estimating it from more recent data, leading to an overprediction of patenting levels in the test set. A potential way to tackle this problem in future research is to allow for time-varying coefficients.

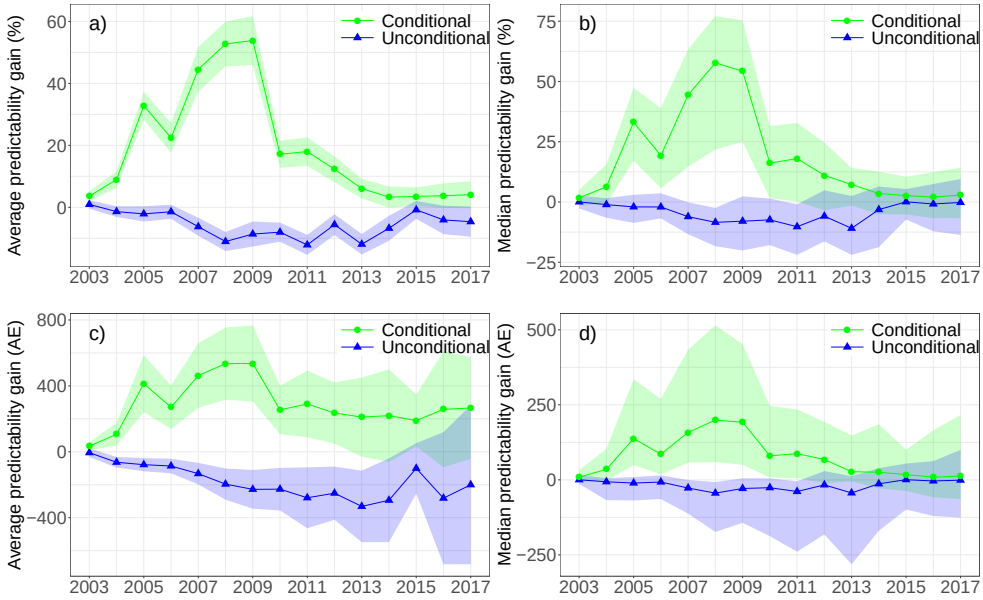

Figure B.7: Prediction results based on **CPC 3-digit codes**. a) Average predictability gain based on $PG1$, Eq. (B.16). Shaded areas indicate twice the standard error. b) Median predictability gain based on $PG1$. Shaded areas indicate the interquartile range. c) Average predictability gain based on $PG2$, Eq. (B.17). Shaded areas indicate twice the standard error. d) Median predictability gain based on $PG2$. Shaded areas indicate the interquartile range.

CPC 4-digit codes (Fig. B.8): In all four panels predictability gains of conditional forecasts are substantial. There is more variation in the unconditional forecasts. Panel a) is the same plot as presented in Chapter 2 which plots average $PG1$ (solid lines) and twice the standard errors (shaded areas) over time. In both prediction exercises, the network model comes with substantial positive predictability gains. In panel b) the median predictability gains, as well as the interquartile range are shown. In the unconditional forecasts, we see positive median predictability gains in the first

eleven years, ranging from 1.3% to 30% and negative predictability gains in the last four years (−5% to −13%) when using $PG1$. The conditional forecasts always exhibit positive median predictability gains. Panels c) shows the average predictability gains based on absolute errors. On average, the network model outperforms the ARIMA forecasts in both prediction exercises, with the conditional forecasts yielding positive prediction gains in every year and the unconditional forecasts in 11 out of 15 years. Panel d) plots the median predictability gains based on the corresponding interquartile range for $PG2$ which yields qualitatively similar results as panel b).

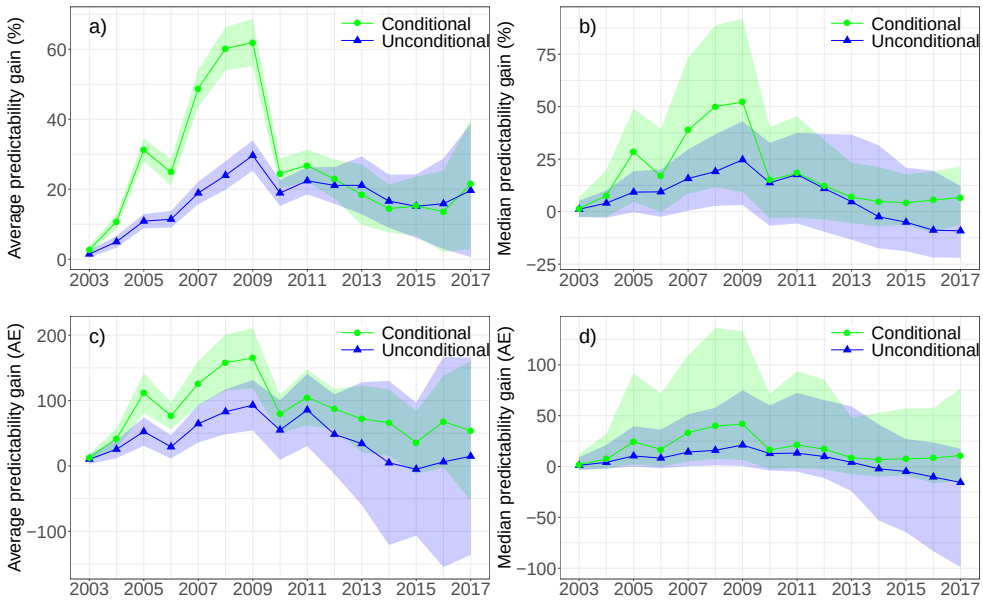

Figure B.8: Prediction results based on **CPC 4-digit codes**. a) Average predictability gain based on $PG1$, Eq. (B.16). Shaded areas indicate twice the standard error. b) Median predictability gain based on $PG1$. Shaded areas indicate the interquartile range. c) Average predictability gain based on $PG2$, Eq. (B.17). Shaded areas indicate twice the standard error. d) Median predictability gain based on $PG2$. Shaded areas indicate the interquartile range.

IPC 3-digit codes (Fig. B.9): The average predictability gains over the ARIMA models in the conditional forecast setting reach almost 60% in 2009 when measured in percentages, panel a), or roughly 800 patents when looking at absolute errors, panel b). In the unconditional scenario, average predictability gains based on $PG1$ are less pronounced (although positive in every single year), but are substantial when considering the absolute error based $PG2$ measure. The median predictability gain based on the absolute percentage error in the conditional forecasts scenario lies between 1.6% and 55% as can be seen in panel c). In the unconditional scenario, the median gains are smaller, but always positive. Panel d) reveals similar patterns as the ones

observed in panel c).

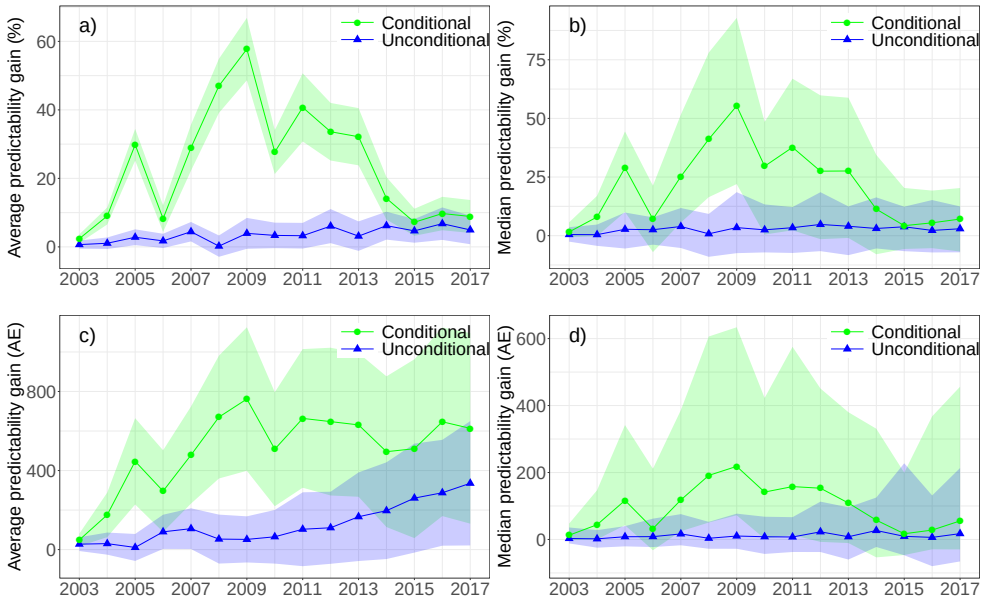

Figure B.9: Prediction results based on **IPC 3-digit codes**. a) Average predictability gain based on $PG1$, Eq. (B.16). Shaded areas indicate twice the standard error. b) Median predictability gain based on $PG1$. Shaded areas indicate the interquartile range. c) Average predictability gain based on $PG2$, Eq. (B.17). Shaded areas indicate twice the standard error. d) Median predictability gain based on $PG2$. Shaded areas indicate the interquartile range.

IPC 4-digit codes (Fig. B.10): For the IPC 4-digit codes we find average percentage predictability gains between 4% in 2003 and 74% in 2013 when considering the conditional network forecasts; see panel a). The predictability gains for the unconditional forecasts can get as high as 29% in the year 2016. Panel b) shows that the median predictability gains are always positive for the conditional forecasts (ranging from 5% to 68%) and positive for the unconditional forecasts in 13 out of 15 forecasts. Only in the last two years, the median predictability gains are slightly negative (less than one percent). When benchmarking the forecasts with absolute errors, Eq. (B.17), we find that predictability gains are always positive on average and tend to increase in time for both prediction scenarios; panel c). Qualitatively, panel d) resembles panel b) where median predictability gains are always positive, except for the last two years in the unconditional forecasts.

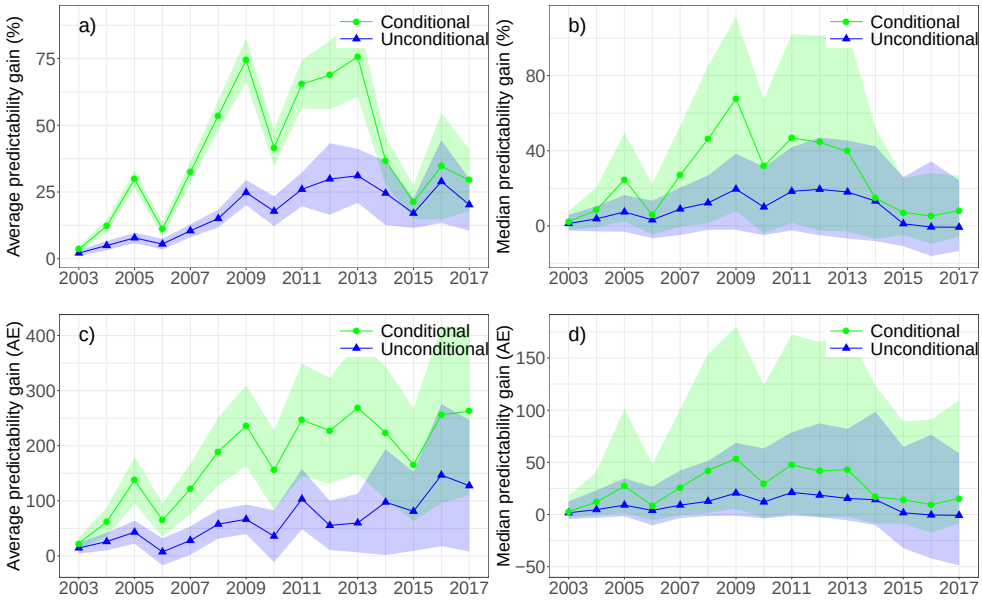

Figure B.10: Prediction results based on **IPC 4-digit codes**. a) Average predictability gain based on $PG1$, Eq. (B.16). Shaded areas indicate twice the standard error. b) Median predictability gain based on $PG1$. Shaded areas indicate the interquartile range. c) Average predictability gain based on $PG2$, Eq. (B.17). Shaded areas indicate twice the standard error. d) Median predictability gain based on $PG2$. Shaded areas indicate the interquartile range.

Appendix C

Appendix to Chapter 3

C.1 First-order supply and demand shocks

C.1.1 Supply shocks

Due to the Covid-19 pandemic, industries experience supply-side reductions due to the closure of non-essential industries, workers not being able to perform their activities at home, and difficulties adapting to social distancing measures. Many industries also face substantial reductions in demand. del Rio-Chanona et al. (2020) provide quantitative predictions of these first-order supply and demand shocks for the U.S. economy. To calculate supply-side predictions, del Rio-Chanona et al. (2020) constructed a Remote Labor Index, which measures the ability of different occupations to work from home, and scored industries according to their essentialness based on the Italian government regulations. We use these estimates, adjusted for the UK context, as our first-order supply shocks.

NAICS to WIOD mapping. Since supply shock estimates are based on indexes and scores available for industries in the NAICS 4-digit classification system, we need to map them into the WIOD industry classification system. We do this by building a crosswalk from the NAICS 4-digit industry classification to the classification system used in WIOD, which is a mix of ISIC 2-digit and 1-digit codes. We make this crosswalk using the NAICS to ISIC 2-digit crosswalk from the European Commission and then aggregating the 2-digit codes presented as 1-digit in the WIOD classification system. We then do an employment-weighted aggregation of the index or score in consideration from the 277 industries at the NAICS 4-digit classification level to the 55 industries in the WIOD classification. Some of the 4-digit NAICS industries map into more than one WIOD industry classification. When this happens, we assume

employment is split uniformly among the WIOD industries the NAICS industry maps into. (The mapping from NAICS to WIOD has primarily be done by R. Maria del Rio-Chanona.)

Remote Labor Index. To estimate the fraction of workers that could work remotely, we use the Remote Labor Index from del Rio-Chanona et al. (2020). Under the assumption that the distribution of occupations across industries and that the percentage of essential workers within an industry is the same for the US and the UK, we can map the Remote Labor Index by del Rio-Chanona et al. (2020) into the UK economy following the mapping methodology explained in the previous paragraph. The WIOD industry sector T ("Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use") only maps into one NAICS code for which we do not have a Remote Labor Index. Since we consider sector T to be similar to the R_S sector ("Other service activities") we use the R_S Remote Labor Index also for the T sector.

Essential Score. The supply shocks at the NAICS level depend on the list of industries' Essential Score at the NAICS 4-digit level provided by del Rio-Chanona et al. (2020). It is important to note that, although the list of industries' Essential Score provided by del Rio-Chanona et al. (2020) is based on the Italian list of essential industries, these lists are based on different industry classification systems (NAICS and NACE, respectively) and do not have a one-to-one correspondence. To derive the Essential Score of industries at the NAICS level, del Rio-Chanona et al. (2020) followed three steps. (i) They considered a 6-digit NAICS industry as essential if the industry had correspondence with at least one essential NACE industry. (ii) They aggregated the 6-digit NAICS essential lists into the 4-digit level taking into account the fraction of NAICS 6-digit subcategories that were considered essential. (iii) They revised each 4-digit NAICS industry's essential score to check for implausible classifications and reclassified ten industries whose original essential score seemed implausible.

Computing supply shocks. Using the Essential Score and the Remote Labor Index, supply shocks are computed following Eq. (3.21). An exception is the sector Real Estate. This sector includes imputed rents, which account for 69% of the monetary value of the sector.[1] Because we think applying a supply shock to imputed rent does

[1] `https://www.ons.gov.uk/economy/grossvalueaddedgva/datasets/nominalandrealregionalgrossvalueaddedbalancedbyindustry`, Table 1B.

not make sense, we compute that the supply shock derived from the Remote Labor Index and Essential Score (which is around 50%) only affects 31% of the sector, leading to a 15% final supply shock to Real Estate (due to an error, our original work used a value of 4.7%). If we had used the correct real estate shock we would have predicted a 22.1% reduction, exactly like in the data. In the main text, we still describe our prediction as having been a 21.5% contraction, as explicitly stated in our original paper.

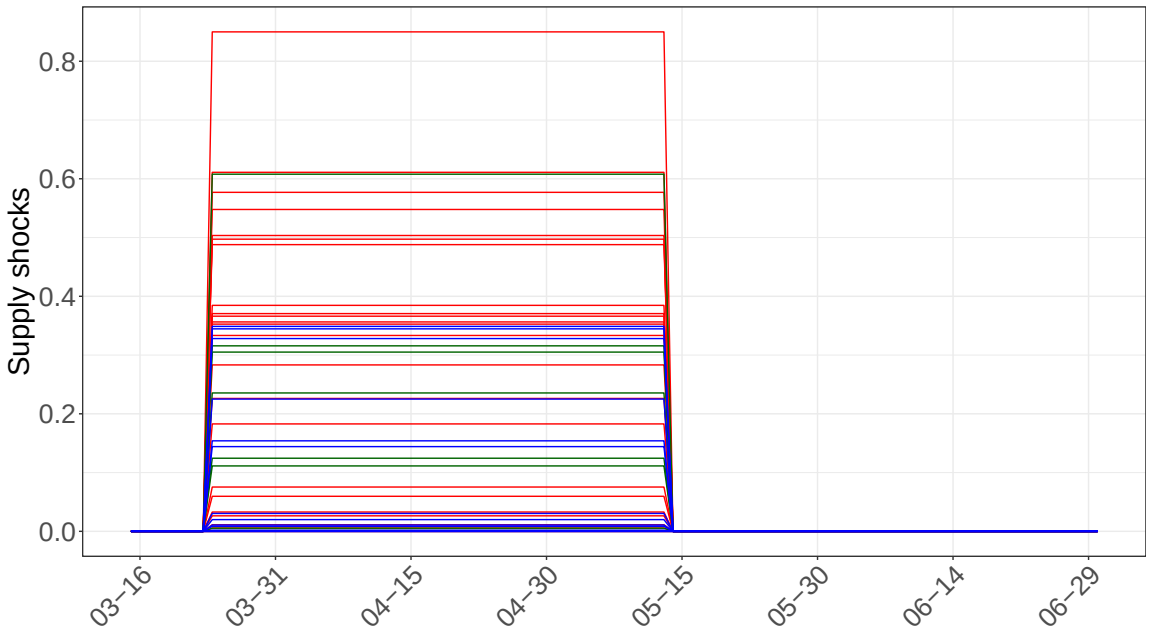

Figure C.1: **Supply shocks $\epsilon^S_{i,t}$ over time and for every industry.** The coloring of the lines is based on the same code as in Fig. 3.1.

C.1.2 Demand shocks

For calibrating consumption demand shocks, we use the same data as del Rio-Chanona et al. (2020) which are based on the Congressional Budget Office (2006) estimates. These estimates are available only at the more aggregate 2-digit NAICS level and map into WIOD ISIC categories without complications. To give a more detailed estimate of consumption demand shocks, we also link manufacturing sectors to the closure of certain non-essential shops as follows.

- Consumption demand shock to C13-C15 (Manufacture of textiles, wearing apparel, leather and other related products). Four subsectors (4751, 4771, 4772, 4782) selling goods produced by this manufacturing sector were mandated to

close, while one subsector was permitted to remain open as it sells homeware goods (4753). Lacking more detailed information about the share of C13-C15 products that these subsectors sell, we simply give equal shares to all subsectors, leading to an 80% consumption demand shock to C13-C15.

- Consumption demand shock to C20 (Manufacture of chemicals and chemical products). Three subsectors (4752, 4773, 4774) selling homeware and medical goods were considered essential, while subsector 4775, selling cosmetic and toilet articles, was mandated to close. Using the same assumptions as above, we get a 25% consumption demand shock for this sector.

The same procedure leads to consumption demand shocks for all other manufacturing subsectors.

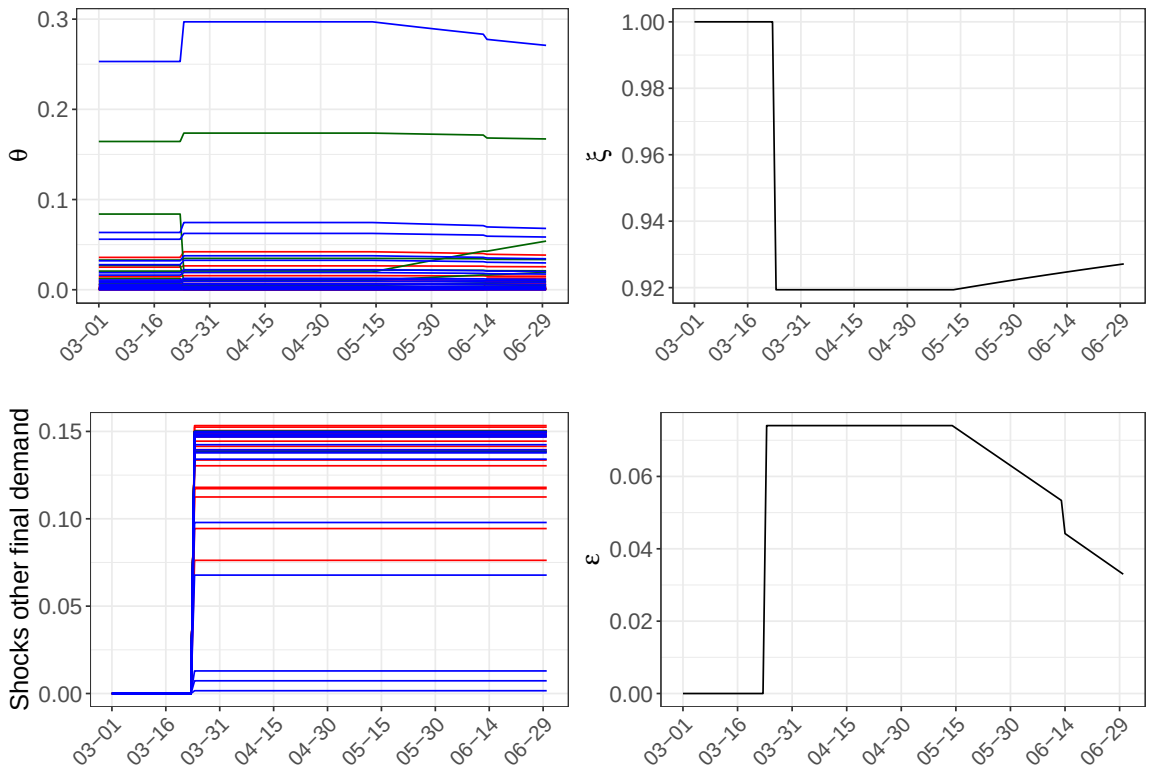

Figure C.2: **Demand shocks over time.** The left upper panel shows the preferences $\theta_{i,t}$ over time. The lower left panel shows industry-specific shocks to other final demand categories. The panel at the top-right shows demand shocks due to fear of unemployment ξ_t. Recall that $\xi_t = 1$ is the pre-lockdown state indicating no shock. The panel at the bottom-right is the aggregate demand shock $\tilde{\epsilon}_t$ taking the savings rate of 50% into account. The coloring of the lines for industry-specific results follows the same code as in Fig. 3.1.

C.2 Inventory data and calibration

Note that the model calibration to inventory data has mostly been achieved by Marco Pangallo. For completeness details on inventory data is still shown below. To calibrate the inventory target parameters n_j in Eq. (3.2), we used the UK Annual Business Survey (ABS). The ABS is the main structural business survey conducted by the ONS.[2] It is sampled from all non-farm, non-financial private businesses in the UK (about two-thirds of the UK economy); data are available up to the 4-digit NACE level, but for our purposes 2-digit NACE industries are sufficient. The survey asks for information on a number of variables, including turnover and inventory stocks (at the beginning and end of each year). Data are available from 2008 to 2018. They show a general increase in turnover and inventory stocks (consistent with growth of the UK economy in the same period), with moderate year-on-year fluctuations.

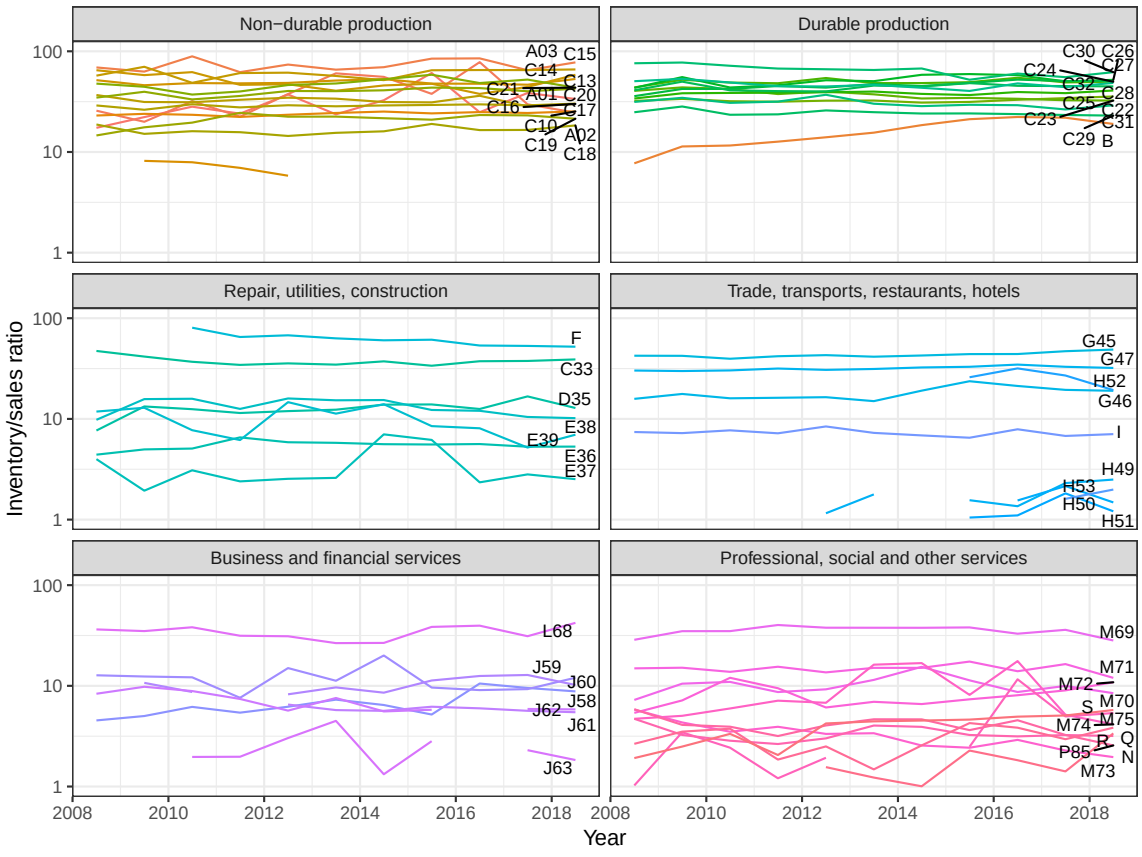

Figure C.3: **Inventory/sales ratios over time.**

To proxy inventory levels available in February 2020, we proceed as follows, for

[2] https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/methodologies/annualbusinesssurveyabs

each industry: (i) we take the simple average between beginning- and end-of year inventory stock levels; (ii) we calculate the ratio between this average and yearly turnover, and multiply this number by 365, because we consider a daily timescale: this inventory-to-turnover ratio is our proxy of n_j (as can be seen in Fig. C.3, the inventory/sales ratios are remarkably constant over time, suggesting that inventory stocks at the beginning of the Covid-19 pandemic could reliably be estimated from past data); (iii) we consider a weighted average between inventory-to-turnover ratios across all years between 2008 and 2018, giving higher weight to more recent years;[3] (iv) we aggregate 2-digit NACE industries to WIOD sectors (for example, we aggregate C10, C11 and C12 to C10_C12); (v) we fill in the missing sectors (K64, K65, K66, O84, T) by imputing the average inventory-to-turnover ratio across all service sectors.

The final inventory to sales ratios n_j are shown in Fig. C.4. As can be seen, inventories are much larger relative to sales in production, construction and trade, while they are generally lower in services, although there is considerable heterogeneity across sectors.

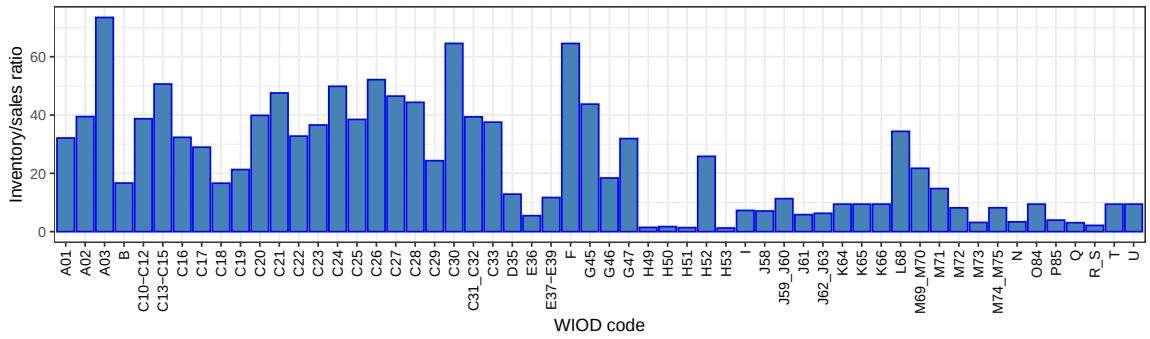

Figure C.4: **Inventory/sales ratios for all WIOD industries.**

C.3 Critical vs. non-critical inputs

A survey was designed to address the question when production can continue during a lockdown. For each industry, IHS Markit analysts were asked to rate every input of a given industry. The exact formulation of the question was as follows: "For each industry in WIOD, please rate whether each of its inputs are essential. We will present you with an industry X and ask you to rate each input Y. The key question is:

[3]More specifically, we consider exponential weights, such that weights from a given year X are proportional to $0.95^{(2018-X)}$. Some years are missing due to confidentiality problems, and data for some sectors have clear problems in some years. We deal with missing values by giving zero weights to years with missing values and renormalizing weights over the available years.

Can production continue in industry X if input Y is not available for two months?" Analysts could rate each input according to the following allowed answers:

- **0** – This input is *not* essential
- **1** – This input is essential
- **0.5** – This input is important but not essential
- **NA** – I have no idea

To avoid confusion with the unrelated definition of essential industries which we used to calibrate first-order supply shocks, we refer to inputs as *critical* and *non-critical* instead of *essential* and *non-essential*.

Analysts were provided with the share of each input in the expenses of the industry. It was also made explicit that the ratings assume no inventories such that a rating captures the effect on production if the input is not available.

Every industry was rated by one analyst, except for industries Mining and Quarrying (B) and Manufacture of Basic Metals (C24) which were rated by three analysts. In case there are several ratings we took the average of the ratings and rounded it to 1 if the average was at least 2/3 and 0 if the average was at most 1/3. Average input ratings lying between these boundaries are assigned the value 0.5.

The ratings for each industry and input are depicted in Fig. C.5. A column denotes an industry and the corresponding rows its inputs. Blue colors indicate *critical*, red *important, but not critical* and white *non-critical* inputs. Note that under the assumption of a Leontief production function every element would be considered to be critical, yielding a completely blue-colored matrix. The results shown here indicate that the majority of elements are non-critical inputs (2,338 ratings with score = 0), whereas only 477 industry-inputs are rates as critical. 365 inputs are rated as important, although not critical (score = 0.5) and *NA* was assigned eleven times.

The left panel of Fig. C.6 shows for each industry how often it was rated as critical input to other industries (x-axis) and how many critical inputs this industry relies on in its own production (y-axis). Electricity and Gas (D35) are rated most frequently as critical inputs in the production of other industries (score=1 for almost 60% of industries). Also frequently rated as critical are Land Transport (H49) and Telecommunications (J61). On the other hand, many manufacturing industries (ISIC codes starting with C) stand out as relying on a large number of critical inputs. For example, around 27% of inputs to Manufacture of Coke and Refined Petroleum Products (C19) as well as to Manufacture of Chemicals (C20) are rated as critical.

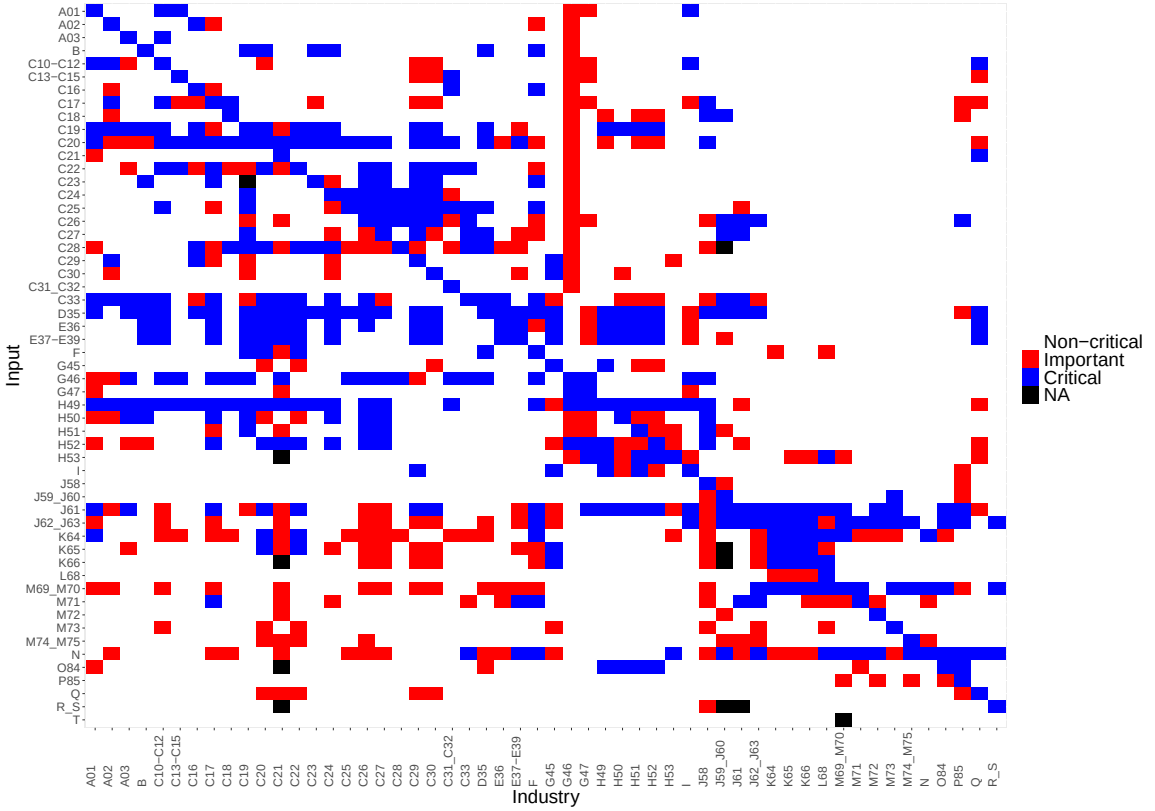

Figure C.5: Criticality scores from IHS Markit analysts. Rows are inputs (supply) and columns industries using these inputs (demand). The blue color indicates critical (score=1), red important (score=0.5) and white non-critical (score=0) inputs. Black denotes inputs which have been rated with NA. The diagonal elements are considered to be critical by definition. For industries with multiple input-ratings we took the average of all ratings and assigned a score=1 if the averaged score was at least 2/3 and a score=0 if the average was smaller or equal to 1/3.

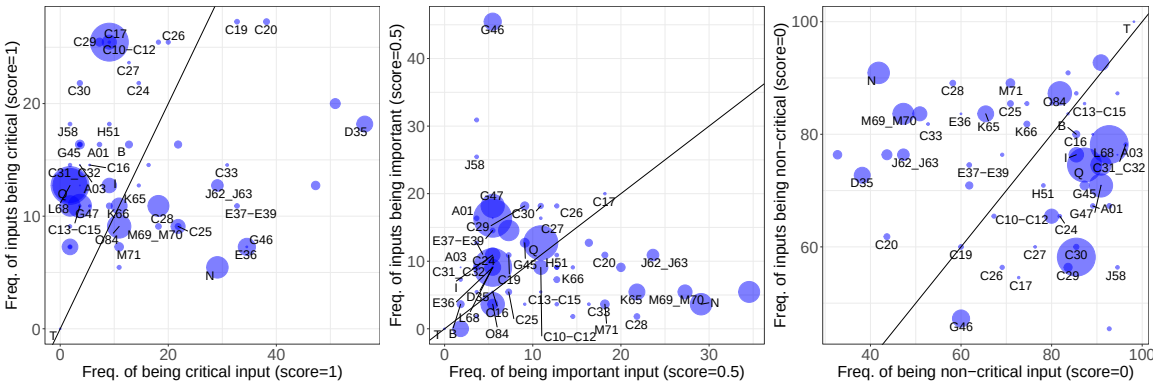

Figure C.6: (Left panel) The figure shows how often an industry is rated as a critical input to other industries (x-axis) against the share of critical inputs this industry is using. The center and right panel are the same as the left panel, except for using half-critical and non-critical scores, respectively. In each plot the identity line is shown. Point sizes are proportional to gross output.

The center panel of Fig. C.6 shows the equivalent plot for 0.5 ratings (important, but not critical inputs). Financial Services (K64) are most frequently rated as important inputs which do not necessarily stop the production of an industry if not available. Conversely, the industry relying on many important, but not binding inputs is Wholesale and Retail Trade (G46) of which almost half of its inputs got rated with a score = 0.5. This makes sense given that this industry heavily relies on all these inputs, but lacking one of these does not halt economic production. This case also illustrates that a Leontief production function could vastly overestimate input bottlenecks as Wholesale and Retail Trade would most likely still be able to realize output even if a several inputs were not available.

In the right panel of Fig. C.6 we show the same scatter plot but for non-critical inputs. 25 industries are rated to be non-critical inputs to other industries in 80% of all cases, with Household Activities (T) and Manufacture of Furniture (C31-32) being rated as non-critical in at least 96%. Industries like Other Services (R-S), Other Professional, Scientific and Technical Activities (M74-75) and Administrative Activities (N) rely on mostly non-critical inputs (>90%).

A detailed breakdown of the input- and industry-specific ratings are given in Table C.1.

ISIC	Sector (abbreviated)	Input-based rankings 1	0.5	0	NA	Industry-based rankings 1	0.5	0	NA	n
A01	Agriculture	4	2	49	0	9	9	37	0	1
A02	Foresty	2	3	50	0	7	9	39	0	1
A03	Fishing	2	1	52	0	8	5	42	0	1
B	Mining	7	1	47	0	9	2	44	0	3
C10-C12	Manuf. Food-Beverages	5	6	44	0	14	5	36	0	1
C13-C15	Manuf. Textiles	2	5	48	0	6	2	47	0	1
C16	Manuf. Wood	3	3	49	0	8	3	44	0	1
C17	Manuf. Paper	5	10	40	0	14	11	30	0	1
C18	Media print	3	6	46	0	6	3	46	0	1
C19	Manuf. Coke-Petroleum	18	4	33	0	15	6	33	2	1
C20	Manuf. Chemical	21	10	24	0	15	6	34	0	1
C21	Manuf. Pharmaceutical	2	2	51	0	9	17	25	7	1
C22	Manuf. Rubber-Plastics	11	7	37	0	14	5	36	0	1
C23	Manuf. Minerals	8	2	44	2	7	1	47	0	1
C24	Manuf. Metals-basic	8	2	45	0	12	7	36	0	3
C25	Manuf. Metals-fabricated	12	4	39	0	5	3	47	0	1
C26	Manuf. Electronic	10	7	38	0	14	10	31	0	1
C27	Manuf. Electric	7	6	42	0	13	9	33	0	1
C28	Manuf. Machinery	10	12	32	2	5	1	49	0	1
C29	Manuf. Vehicles	4	5	46	0	14	10	31	0	1
C30	Manuf. Transport-other	2	6	47	0	12	10	33	0	1
C31_C32	Manuf. Furniture	1	1	53	0	8	4	43	0	1
C33	Repair-Installation	17	9	29	0	8	2	45	0	1
D35	Electricity-Gas	31	3	21	0	10	5	40	0	1
E36	Water	19	3	33	0	4	5	46	0	1
E37-E39	Sewage	18	3	34	0	6	8	41	0	1
F	Construction	5	3	47	0	14	9	32	0	1
G45	Vehicle trade	2	5	48	0	9	7	39	0	1
G46	Wholesale	19	3	33	0	4	25	26	0	1
G47	Retail	2	3	50	0	6	10	39	0	1
H49	Land transport	28	3	24	0	11	2	42	0	1
H50	Water transport	9	8	38	0	8	5	42	0	1
H51	Air transport	5	7	43	0	10	6	39	0	1
H52	Warehousing	12	9	34	0	9	7	39	0	1
H53	Postal	6	7	41	2	3	5	47	0	1
I	Accommodation-Food	5	3	47	0	7	6	42	0	1
J58	Publishing	1	2	52	0	10	14	31	0	1
J59_J60	Video-Sound-Broadcasting	2	2	51	0	9	5	37	7	1
J61	Telecommunications	26	11	18	0	7	5	42	2	1
J62_J63	IT	16	13	26	0	7	6	42	0	1
K64	Finance	10	19	26	0	6	3	46	0	1
K65	Insurance	6	12	36	2	6	3	46	0	1
K66	Auxil. Finance-Insurance	5	7	41	4	6	4	45	0	1
L68	Real estate	1	3	51	0	7	5	43	0	1
M69_M70	Legal	12	15	28	0	5	3	46	2	1
M71	Architecture-Engineering	6	10	39	0	4	2	49	0	1
M72	R&D	1	2	52	0	4	3	48	0	1
M73	Advertising	1	7	47	0	5	2	48	0	1
M74_M75	Other Science	1	8	46	0	4	1	50	0	1
N	Private Administration	16	16	23	0	3	2	50	0	1
O84	Public Administration	6	3	45	2	5	2	48	0	1
P85	Education	1	4	50	0	6	8	41	0	1
Q	Health	1	6	48	0	7	7	41	0	1
R_S	Other Service	1	1	50	5	4	0	51	0	1
T	Household activities	0	0	54	2	0	0	55	0	0

Table C.1: Summary table of critical input ratings by IHS Markit analysts. Columns below *Input-based rankings* show how often an industry has been rated as critical (score=1), half-critical (score=0.5) or non-critical (score=0) input for other industries, or how often the input was rates as NA. Columns under *Industry-based rankings* give how often an input has been rated as with 1, 0.5, 0 or NA for any given industry. Column n indicates the number of analysts who have rated the inputs of any given industry. Industry T uses no inputs and is therefore not rated.

C.4 Model production function and CES

Here we show that the production functions used in the main text are highly related to nested CES production functions. Specifically, we consider a CES production function with three nests of the form (we suppress time indices for convenience)

$$x_i^{\text{inp}} = \left(a_i^{\text{C}} (z_i^{\text{C}})^{\beta} + a_i^{\text{IMP}} (z_i^{\text{IMP}})^{\beta} + (a_i^{\text{NC}})^{1-\beta} (z_i^{\text{NC}})^{\beta} \right)^{\frac{1}{\beta}}, \tag{C.1}$$

where β is the substitution parameter. Variables $a_i^{\text{C}} = \sum_{j \in \text{C}} A_{ji}$, $a_i^{\text{IMP}} = \sum_{j \in \text{IMP}} A_{ji}$ and $a_i^{\text{NC}} = \sum_{j \in \text{NC}} A_{ji}$ are the input shares (technical coefficients) for critical, important and non-critical inputs, respectively. To be consistent with the specifications of the main text, we do not consider labor inputs here and only focus on x_i^{inp}. Alternatively, we could include labor inputs in the set of critical inputs and derive the full production function in an analogous manner. z_i^{C}, z_i^{IMP} and z_i^{NC} are CES aggregates of critical, important and non-critical inputs for which we have

$$z_i^{\text{C}} = \left[\sum_{j \in \text{C}} A_{ji}^{1-\nu} S_{ji}^{\nu} \right]^{\frac{1}{\nu}}, \tag{C.2}$$

$$z_i^{\text{IMP}} = \frac{1}{2} \left[\sum_{j \in \text{IMP}} A_{ji}^{1-\psi} S_{ji}^{\psi} \right]^{\frac{1}{\psi}} + \frac{1}{2} x_i^{\text{cap}}, \tag{C.3}$$

$$z_i^{\text{NC}} = \left[\sum_{j \in \text{NC}} A_{ji}^{1-\zeta} S_{ji}^{\zeta} \right]^{\frac{1}{\zeta}}. \tag{C.4}$$

If we assume that every input is critical (i.e. the set of important and non-critical inputs is empty: $a_i^{\text{IMP}} = a_i^{\text{NC}} = 0$), and by taking the limits $\beta, \nu \to -\infty$ we recover the Leontief production function

$$x_i^{\text{inp}} = \min_j \left\{ \frac{S_{ji}}{A_{ji}} \right\}. \tag{C.5}$$

If we assume that there is no critical or important inputs at all ($a_i^{\text{C}} = a_i^{\text{IMP}} = 0$), taking the limits $\beta \to -\infty$ and $\zeta \to 1$ yields the linear production

$$x_i^{\text{inp}} = \frac{\sum_j S_{ji}}{a_i^{\text{NC}}}. \tag{C.6}$$

The IHS1 and IHS3 production functions treat important inputs either as critical or non-critical, i.e. the set of critical inputs is again empty ($a_i^{\text{IMP}} = 0$). We can

approximate these functions again by taking the limits $\beta, \nu \to -\infty$ and $\zeta \to 1$ which yields

$$x_i^{\text{inp}} = \min \left\{ \min_{j \in \text{C}} \left\{ \frac{S_{ji}}{A_{ji}} \right\}, \frac{\sum_{j \in \text{NC}} S_{ji}}{a_i^{\text{NC}}} \right\}. \tag{C.7}$$

Note that the difference between the two production functions lies in the different definition of critical and non-critical inputs. The functional form in Eq. (C.7) applies to both cases equivalently.

The IHS2 production function where the set of important inputs is non-empty can be similarly approximated. In addition to the limits above, we also take the limit $\psi \to -\infty$ to obtain

$$x_i^{\text{inp}} = \min \left\{ \min_{j \in \text{C}} \left\{ \frac{S_{ji}}{A_{ji}} \right\}, \frac{1}{2} \min_{j \in \text{IMP}} \left\{ \frac{S_{ji}}{A_{ji}} + x_i^{\text{cap}} \right\}, \frac{\sum_{j \in \text{NC}} S_{ji}}{a_i^{\text{NC}}} \right\}. \tag{C.8}$$

Eqs. (C.7) and (C.8) are not exactly identical to the IHS production functions used in the main text (Eqs. (3.8) – (3.10)), but very similar. The only difference is that in the IHS functions non-critical inputs do not play a role for production at all, whereas here they enter the equations as a linear term. However, simulations show that this difference is practically irrelevant. Simulating the model with the CES-derived functions instead of the IHS production functions yield exactly the same results, indicating that the linear term representing non-critical inputs is never a binding constraint.

C.5 Sensitivity analysis

To gain a better understanding of the model behavior, we investigate how economic impact estimates change under alternative parametrizations. In particular, we vary the parameters τ, γ_H, γ_F and Δs and consider their interactions with various production function specifications. Fig. C.7 summarizes the results with respect to different parametrizations of the inventory adjustment parameter τ (left column), hiring/separation parameters γ_H and γ_F (center left column), change in savings rate parameter Δs (center right column) and consumption adjustment speed ρ (right column) under alternative production function assumptions (panel rows).

Overall, we find that the model becomes more sensitive with respect to parameters the closer economic production follows a Leontief assumption. In contrast, results are extremely similar for the whole parameter range considered here under a linear production function. Even for the Leontief case, however, the model is not overly sensitive with respect to the parameters. We emphasize that the parameter sensitivity of

the model can be affected by further aspects, such as first-order shocks and inventory levels. These have been studied in Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2020) and Pichler, Pangallo, del Rio-Chanona, Lafond & Farmer (2021) to which we refer for further details.

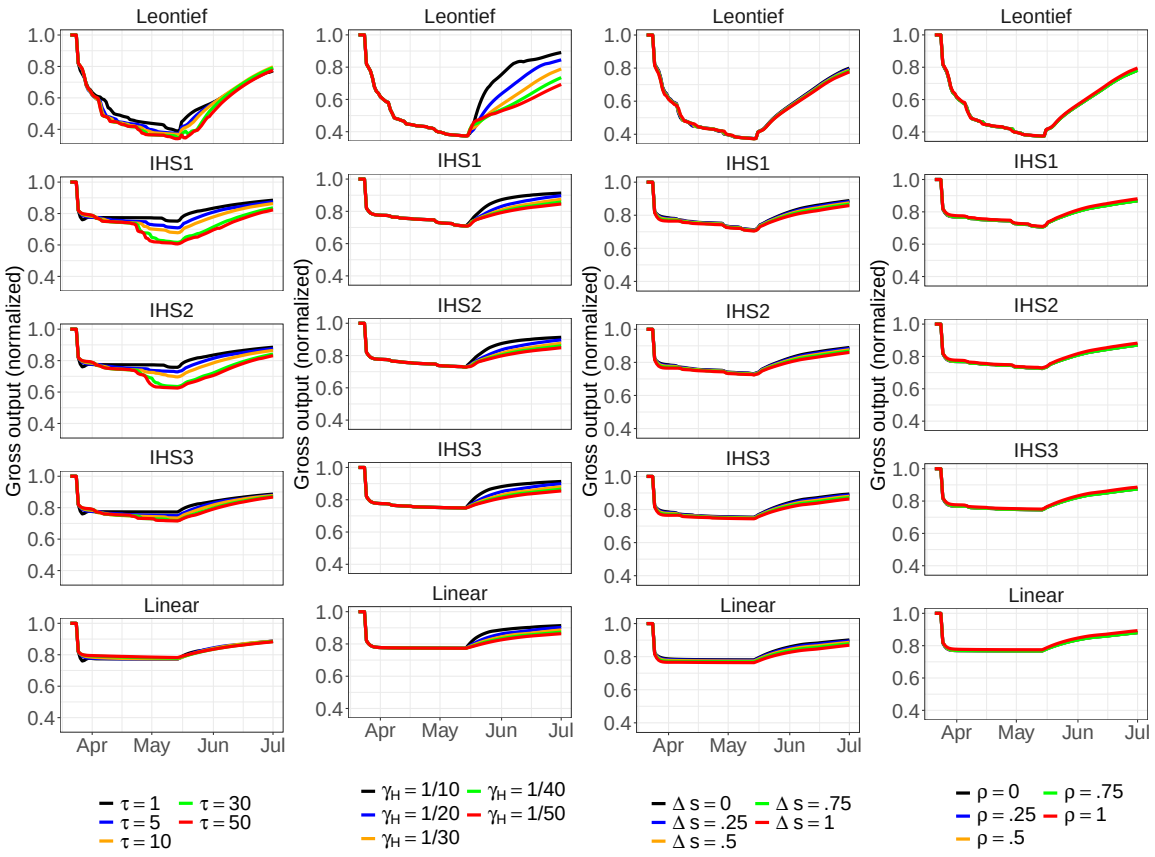

Figure C.7: **Results of sensitivity analysis for parameters τ, γ_H, Δs and ρ.** All plots show normalized gross output on the y-axis and days after the shocks are applied on the x-axis. Panel rows differ with respect to the production function assumption. Panel columns differ with respect to τ (left), γ_H (center left), Δs (center right) and ρ (right) as indicated in the legend below each column. Except for the four parameters and the production function, the model is initialized as in the baseline run of the main text. In all simulations we used $\gamma_F = 2\gamma_H$.

Appendix D

Appendix to Chapter 4

D.1 Details on first-order shocks to supply and demand

		f_i (%)			ϵ_i^D (%)		
ISIC	Industry	DEU	ESP	ITA	DEU	ESP	ITA
A01	Agriculture	0.5	1.4	1.0	10.0	9.9	10.0
A02	Forestry	0.1	0.0	0.1	10.0	8.6	3.8
A03	Fishing	0.0	0.1	0.0	10.0	10.0	10.0
B	Mining	0.3	0.5	0.4	10.0	10.0	10.0
C10-12	Manuf. Food-Beverages	4.1	5.0	4.0	10.0	10.0	10.0
C13-15	Manuf. Textiles	0.6	1.4	2.9	10.0	10.0	10.0
C16	Manuf. Wood	0.4	0.1	0.3	10.0	10.0	9.9
C17	Manuf. Paper	0.6	0.4	0.5	10.0	10.0	10.0
C18	Media print	0.1	0.1	0.1	10.0	9.9	10.0
C19	Manuf. Coke-Petroleum	1.7	2.6	1.4	10.0	10.0	10.0
C20	Manuf. Chemical	3.4	2.1	1.5	9.9	9.9	10.0
C21	Manuf. Pharmaceutical	1.1	1.0	1.2	8.1	9.3	9.8
C22	Manuf. Rubber-Plastics	1.4	0.7	1.0	10.0	9.9	10.0
C23	Manuf. Minerals	0.6	0.5	0.7	10.0	10.0	10.0
C24	Manuf. Metals-basic	1.5	1.3	1.4	10.0	10.0	10.0
C25	Manuf. Metals-fabricated	1.8	1.0	1.6	10.0	10.0	10.0
C26	Manuf. Electronic	2.0	0.4	0.9	9.9	9.9	10.0
C27	Manuf. Electric	2.3	0.8	1.4	10.0	10.0	10.0
C28	Manuf. Machinery	5.8	1.4	4.7	10.0	10.0	10.0
C29	Manuf. Vehicles	8.2	3.7	2.1	10.0	10.0	10.0
C30	Manuf. Transport-other	1.0	1.1	1.0	9.9	9.9	10.0
C31-32	Manuf. Furniture	1.4	0.6	1.4	9.9	9.9	10.0
C33	Repair-Installation	0.5	0.3	0.6	10.0	10.0	10.0
D35	Electricity-Gas	1.6	1.6	1.1	2.3	0.8	1.2
E36	Water	0.2	0.4	0.3	1.5	0.9	0.6
E37-39	Sewage	0.7	0.7	0.6	3.6	2.5	1.7
F	Construction	5.5	7.6	7.5	10.0	9.9	9.9
G45	Vehicle trade	0.9	1.7	1.4	10.0	10.0	10.0
G46	Wholesale	3.4	4.1	4.6	9.8	9.5	9.9
G47	Retail	3.4	5.5	5.7	9.5	9.5	9.8
H49	Land transport	0.8	1.7	1.8	56.9	38.8	55.2
H50	Water transport	0.7	0.2	0.5	11.4	27.5	47.3
H51	Air transport	0.6	0.6	0.4	50.3	24.7	47.3
H52	Warehousing	0.3	0.8	1.1	22.3	19.7	33.5
H53	Postal	0.1	0.0	0.1	3.4	2.2	3.4
I	Accommodation-Food	2.4	8.6	4.7	73.1	75.5	79.1
J58	Publishing	0.4	0.3	0.3	3.8	5.4	4.0
J59-60	Video-Sound-Broadcasting	0.6	0.5	0.3	4.3	3.8	3.5
J61	Telecommunications	0.9	1.3	1.2	1.1	1.6	3.0
J62-63	IT	1.5	1.6	1.1	8.8	9.5	8.5
K64	Finance	1.9	0.8	0.9	2.9	3.8	1.6
K65	Insurance	1.4	1.1	1.0	1.3	0.8	1.1
K66	Auxil. Finance-Insurance	0.0	0.2	0.1	2.0	1.9	4.9
L68	Real estate	7.2	7.8	9.8	0.2	0.1	0.5
M69-70	Legal	0.7	0.6	0.4	9.3	7.9	5.2
M71	Architecture-Engineering	1.0	1.0	0.2	9.4	9.3	8.4
M72	R&D	0.9	0.5	0.6	8.3	7.9	9.8
M73	Advertising	0.1	0.1	0.1	10.0	9.1	9.2
M74-75	Other Science	0.3	0.1	0.3	3.5	3.5	5.0
N	Private Administration	1.2	1.2	1.1	4.2	4.2	4.9
O84	Public Administration	6.1	6.3	7.1	0.2	0.7	0.0
P85	Education	4.1	5.1	3.9	0.7	1.0	0.0
Q	Health	8.3	7.2	7.5	0.1	0.1	0.1
R_S	Other Service	3.2	3.3	3.2	4.3	4.3	4.7
T	Household activities	0.2	0.8	1.1	0.0	-0.0	-0.0

Table D.1: **Industry-specific demand shock details.** f_i denotes final consumption per industry as fraction of aggregate final consumption. ϵ_i^D is the total demand shock per industry.

ISIC	Industry	x_i (%) DEU	ESP	ITA	ϵ_i^S (%) DEU	ESP	ITA	e_i (%) DEU	ESP	ITA	RLI_i (%)
A01	Agriculture	0.9	2.2	1.7	0.0	0.0	0.0	100.0	100.0	100.0	13.6
A02	Forestry	0.1	0.0	0.0	85.0	85.0	85.0	0.0	0.0	0.0	15.0
A03	Fishing	0.0	0.1	0.1	0.0	0.0	0.0	100.0	100.0	100.0	35.7
B	Mining	0.2	0.3	0.3	48.3	48.3	34.5	30.0	30.0	50.0	31.0
C10-12	Manuf. Food-Beverages	3.5	6.9	4.1	0.0	26.0	26.0	100.0	66.7	66.7	22.1
C13-15	Manuf. Textiles	0.4	1.0	2.6	68.5	68.5	59.4	0.0	0.0	13.3	31.5
C16	Manuf. Wood	0.5	0.3	0.5	73.1	73.1	60.6	0.0	0.0	17.0	26.9
C17	Manuf. Paper	0.7	0.6	0.7	34.3	0.0	48.7	50.0	100.0	29.0	31.5
C18	Media print	0.4	0.4	0.4	0.0	0.0	0.0	100.0	100.0	100.0	39.0
C19	Manuf. Coke-Petroleum	1.5	2.4	1.7	0.0	6.4	0.0	100.0	90.0	100.0	36.0
C20	Manuf. Chemical	2.6	2.5	1.6	19.0	52.5	8.2	70.0	17.0	87.0	36.7
C21	Manuf. Pharmaceutical	0.9	0.7	0.8	0.0	0.0	0.0	100.0	100.0	100.0	40.3
C22	Manuf. Rubber-Plastics	1.4	0.9	1.3	35.3	70.5	47.2	50.0	0.0	33.0	29.5
C23	Manuf. Minerals	0.8	0.8	1.0	63.9	63.9	61.4	0.0	0.0	4.0	36.1
C24	Manuf. Metals-basic	1.9	2.1	1.8	72.6	72.6	72.6	0.0	0.0	0.0	27.4
C25	Manuf. Metals-fabricated	2.4	1.4	2.6	66.3	66.3	66.3	0.0	0.0	0.0	33.7
C26	Manuf. Electronic	1.4	0.4	0.7	43.1	43.1	38.8	0.0	0.0	10.0	56.9
C27	Manuf. Electric	1.9	0.8	1.1	63.1	63.1	50.5	0.0	0.0	20.0	36.9
C28	Manuf. Machinery	4.5	1.2	3.6	61.8	61.8	47.0	0.0	0.0	24.0	38.2
C29	Manuf. Vehicles	6.3	2.5	1.5	69.7	69.7	69.7	0.0	0.0	0.0	30.3
C30	Manuf. Transport-other	0.8	0.8	0.7	59.7	59.7	59.7	0.0	0.0	0.0	40.3
C31-32	Manuf. Furniture	0.9	0.6	1.2	65.2	59.7	58.0	0.0	8.5	11.0	34.8
C33	Repair-Installation	0.7	0.5	0.6	60.6	60.6	34.0	0.0	0.0	44.0	39.4
D35	Electricity-Gas	2.4	4.7	2.7	0.0	5.8	0.0	100.0	90.0	100.0	41.6
E36	Water	0.2	0.5	0.3	0.0	0.0	0.0	100.0	100.0	100.0	33.5
E37-39	Sewage	0.9	0.8	1.3	0.0	0.0	0.0	100.0	100.0	100.0	29.8
F	Construction	5.2	6.4	6.7	71.6	71.6	49.6	0.0	0.0	30.7	28.4
G45	Vehicle trade	1.1	1.4	1.1	18.0	27.3	13.7	67.0	50.0	75.0	45.4
G46	Wholesale	3.9	4.9	5.3	0.0	30.0	33.5	100.0	40.0	33.0	50.1
G47	Retail	3.0	3.9	3.9	24.7	25.7	25.2	51.0	49.0	50.0	49.7
H49	Land transport	1.8	2.5	3.0	0.0	0.0	0.0	100.0	100.0	100.0	31.4
H50	Water transport	0.5	0.2	0.4	0.0	0.0	0.0	100.0	100.0	100.0	35.3
H51	Air transport	0.5	0.5	0.4	0.0	24.2	0.0	100.0	66.0	100.0	28.8
H52	Warehousing	2.3	2.1	2.1	0.0	0.0	0.0	100.0	100.0	100.0	29.6
H53	Postal	0.6	0.2	0.2	0.0	0.0	0.0	100.0	100.0	100.0	35.6
I	Accommodation-Food	1.6	5.8	3.3	64.6	64.6	56.6	0.0	0.0	12.5	35.4
J58	Publishing	0.6	0.4	0.3	0.0	7.6	0.0	100.0	75.0	100.0	69.8
J59-60	Video-Sound-Broadcasting	0.6	0.6	0.5	0.0	11.0	0.0	100.0	75.0	100.0	56.1
J61	Telecommunications	1.2	1.8	1.3	0.0	0.0	0.0	100.0	100.0	100.0	55.1
J62-63	IT	2.1	1.4	1.6	7.2	14.4	0.0	75.0	50.0	100.0	71.1
K64	Finance	2.7	2.1	2.9	0.0	0.0	0.0	100.0	100.0	100.0	71.4
K65	Insurance	1.4	1.0	0.8	0.0	0.0	0.0	100.0	100.0	100.0	71.3
K66	Auxil. Finance-Insurance	0.6	0.4	1.0	0.0	0.0	0.0	100.0	100.0	100.0	71.7
L68	Real estate	7.2	6.7	7.5	51.3	51.3	51.3	0.0	0.0	0.0	48.7
M69-70	Legal	2.5	1.5	2.3	36.3	18.1	0.0	0.0	50.0	100.0	63.7
M71	Architecture-Engineering	1.2	1.1	1.0	45.9	45.9	0.0	0.0	0.0	100.0	54.1
M72	R&D	0.6	0.3	0.4	41.1	41.1	0.0	0.0	0.0	100.0	58.9
M73	Advertising	0.4	0.5	0.5	39.7	39.7	39.7	0.0	0.0	0.0	60.3
M74-75	Other Science	0.4	0.3	0.7	19.6	19.6	0.0	50.0	50.0	100.0	60.8
N	Private Administration	3.9	2.6	3.0	42.7	54.3	41.6	33.3	15.2	35.0	36.0
O84	Public Administration	4.6	4.4	4.2	0.0	0.0	0.0	100.0	100.0	100.0	44.6
P85	Education	2.9	3.4	2.4	0.0	46.0	0.0	100.0	0.0	100.0	54.0
Q	Health	5.4	4.8	4.9	0.0	0.0	0.0	100.0	100.0	100.0	36.0
R_S	Other Service	2.8	2.7	2.7	61.2	56.0	49.2	0.0	8.6	19.6	38.8
T	Household activities	0.1	0.5	0.6	0.0	0.0	0.0	0.0	0.0	50.0	100.0

Table D.2: **Industry-specific supply shock details.** x_i denotes gross output per industry as share of aggregate output. ϵ_i^S is the total supply shock per industry. e_i is the extent to which an industry is considered as essential following government policies. RLI_i is the Remote Labor Index.

D.2 Mixed endogenous/exogenous modeling

A core motivation to study shock propagation based on alternative, bottom-up approaches was our observation that existing methods such as the mixed endogenous-

exogenous model (MEEM) yields infeasible solutions when applied to pandemic shocks. In the MEEM the N industries are divided into two groups. The first group is the set of N^s supply constrained industries and the second group the N^d demand-constrained industries (we have $N^s + N^d = N$). We can use this setup to partition Eq. (4.4) into

$$\begin{pmatrix} \boldsymbol{x}^s \\ \boldsymbol{x}^d \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{ss} & \mathbf{A}^{sd} \\ \mathbf{A}^{ds} & \mathbf{A}^{dd} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^s \\ \boldsymbol{x}^d \end{pmatrix} + \begin{pmatrix} \boldsymbol{f}^s \\ \boldsymbol{f}^d \end{pmatrix}, \tag{D.1}$$

where superscripts s and d denote supply and demand constrained industries, respectively. The matrix block $\mathbf{A}^{sd}$ indicates the input recipes of demand constrained customers with respect to output constrained suppliers and analogously for the other blocks. In this framework vectors $\boldsymbol{f}^s$ and $\boldsymbol{x}^d$ are endogenous and $\boldsymbol{f}^d$ and $\boldsymbol{x}^s$ exogenous. Rearranging Eq. (D.1) such that all endogenous variables appear on the left-hand side and all exogenous variables on the right-hand side, yields

$$\begin{pmatrix} \boldsymbol{f}^s \\ \boldsymbol{x}^d \end{pmatrix} = \begin{pmatrix} \mathbb{I} & \mathbf{A}^{sd} \\ \mathbf{0} & \mathbb{I} - \mathbf{A}^{dd} \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{I} - \mathbf{A}^{ss} & \mathbf{0} \\ \mathbf{A}^{ds} & \mathbb{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}^s \\ \boldsymbol{f}^d \end{pmatrix} \tag{D.2}$$

(for details see Miller & Blair (2009), 621-633).

To apply the MEEM in the pandemic context, we categorize industries into a demand and a supply constrained group based on which shock is larger. If the supply shock of an industry exceeds its demand shock in absolute terms, we treat this industry as supply constrained and vice versa. The shock magnitudes for supply and demand for each industry are given as

$$x_i^{\mathrm{SS}} = x_{i,0} - x_i^{\max} = -\epsilon_i^S x_{i,0}, \tag{D.3}$$

$$f_i^{\mathrm{DS}} = c_{i,0} - f_i^{\max} = -\epsilon_i^D f_{i,0}, \tag{D.4}$$

where d_i^{SS} denotes the total *supply shock* and f_i^{DS} the *demand shock*. If $x_i^{\mathrm{SS}} > f_i^{\mathrm{DS}}$, we consider this industry as supply constrained and we will use its gross output values on the right hand side of Eq. (D.2). Otherwise we treat it as demand constrained.

Following this approach, we apply the MEEM to the IO data of Germany, Italy and Spain by calibrating it to the estimated pandemic supply and demand shocks. As shown in Fig. D.1, the MEEM does not yield a feasible solution for any of the three countries. Violations of feasibility conditions are most frequent for Spain, which faces the largest shocks to supply and demand. The model does not compute any negative final consumption values for Germany, but still allocates final consumption values to industries which are larger than $f_i^{\max}$. Note that the results obtained by the MEEM are out of the solution space delimited by the exogenous constraints on output and consumption. Thus, these results are difficult to interpret in this context and we thus

refrain from comparing them more systematically with the results shown in the main text.

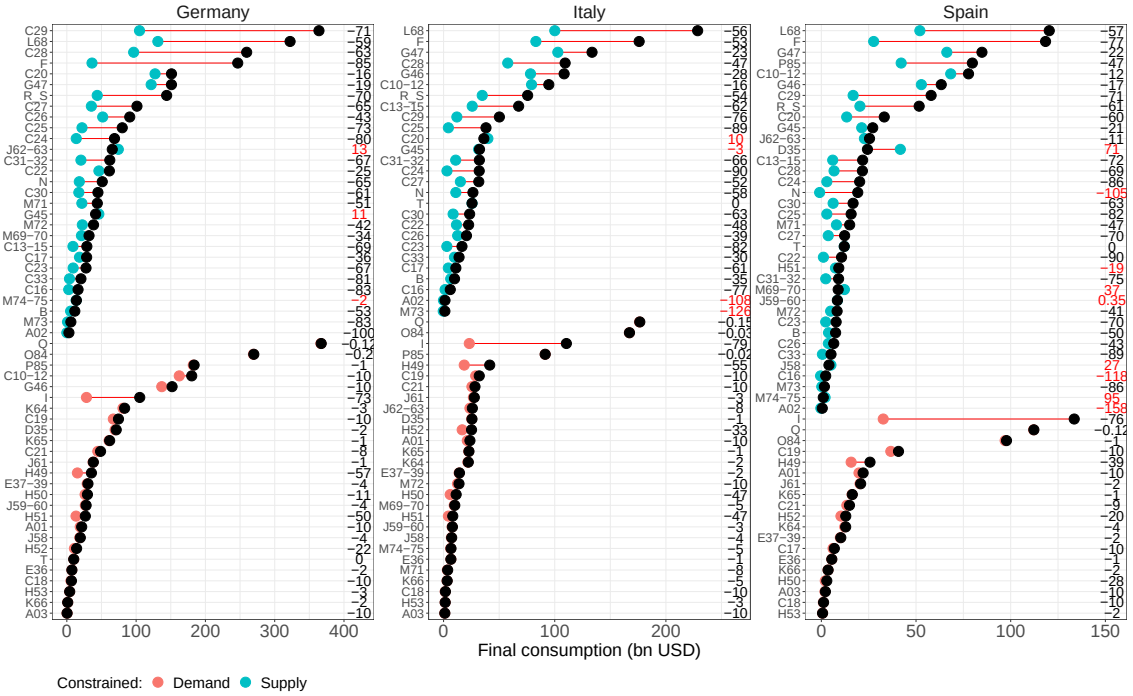

Figure D.1: **Final consumption values from mixed exogenous/endogenous IO modeling with simultaneous supply and demand shocks.** Black circles indicate initial pre-shock values, red circles demand-constrained industries and blue circles supply-constrained industries. Note that demand-constrained consumption values are exogenously determined by the first-order shocks, while in other cases the change in output depends on the higher order effects on the whole economy. Red lines indicate the overall change in production for each sector. Numbers to the right of both panels indicate the value change in percentages and are colored red if the result is infeasible.

The MEEM can yield infeasible economic allocations and it depends on the context whether negative final consumption values are meaningful or not (Miller & Blair 2009, p. 628). To see why the MEEM can give negative final consumption values, let us consider the output-constrained part of Eq. (D.2), which can be rewritten as

$$\boldsymbol{f}^s = [\mathbb{I} - \mathbf{A}^{ss}]\boldsymbol{x}^s - \mathbf{A}^{sd}[\mathbb{I} - \mathbf{A}^{dd}]^{-1}(\mathbf{A}^{ds}\boldsymbol{x}^s + \boldsymbol{f}^d). \tag{D.5}$$

Note that the vector $(\mathbf{A}^{ds}\boldsymbol{x}^s + \boldsymbol{f}^d)$ is always non-negative by definition, as is every element of the matrix $\mathbf{A}^{sd}[\mathbb{I} - \mathbf{A}^{dd}]^{-1}$. (This is clear from invoking the Hawkins-Simon conditions). Thus, for $\boldsymbol{f}^s \geq \mathbf{0}$, $[\mathbb{I} - \mathbf{A}^{ss}]\boldsymbol{x}^s$ must be non-negative and larger than the part to the right of the minus in Eq. (D.5). However, this term can be negative for supply shocks that are sufficiently heterogeneous. For example, consider the case of an economy with only two supply-constrained industries without self-loops, i and j. Any supply shocks which lead to $x_i^{\max} < A_{ij}^{ss} x_j^{\max}$ yield negative final consumption

values for industry i. This demonstrates that the MEEM framework is unlikely to yield plausible solutions in the current context.

The MEEM can also lead to final consumption values that lie above any given pre-specified upper limit of consumption $f_i^{\max}$. Let us again consider an example of two industries, i and j, where i is supply constrained and j is demand constrained. If industry i only supplies industry j and final consumers and industry j only supplies to final consumers, it can be verified that $f_i^s > f_i^{\max}$ if $x_i^{\mathrm{SS}} - f_i^{\mathrm{DS}} < A_{ij}^{sd} f_j^{\mathrm{DS}}$. Thus, in case industry i is only slightly supply constrained (supply shocks are only slightly larger than demand shocks) and industry j faces comparatively serious demand constraints, the MEEM would compute larger final consumption values for industry i than are possible.

Note that gross output values of demand constrained industries always lie within the feasible range $x_i^d \in [0, x_i^{\max}]$ if only adverse supply shocks to the economy are allowed ($\epsilon_i^S \geq 0$). By noting that $\boldsymbol{x}^d = [\mathbb{I} - \mathbf{A}^{dd}]^{-1}(\boldsymbol{A}^{ds}\boldsymbol{x}^s + \boldsymbol{f}^d)$, it can easily be verified that $x_i^d \geq 0$. The condition that $\epsilon_i^S x_i^d < \epsilon_i^D f_i^d$ ensures that output of demand constrained industries cannot exceed the maximum output value $x_i^{\max}$.

The supply and demand shocks do not need to be large for the MEEM to generate infeasible solutions. To show this we scale down the size of the supply and demand shocks by varying a parameter $\alpha \in [0, 1]$ to obtain new maximum output and consumption values, according to

$$x_i^{\max} = (1 - \alpha\epsilon_i^S)x_{i,0}, \tag{D.6}$$

$$f_i^{\max} = (1 - \alpha\epsilon_i^D)f_{i,0}. \tag{D.7}$$

Since we scale demand and supply shocks by the same proportion this does not change which industries are supply or demand constrained.

Fig. D.2 shows the MEEM results for varying directs shocks. If $\alpha = 0$, there is no direct shock, resulting in a feasible market allocation since in this case the MEEM simply recovers the pre-shock economy. This is indicated by the green colors at the very left of all three panels. But even for very small $\alpha > 0$ we obtain infeasible solutions for all three countries as shown by the gray colors. If we increase shock sizes further, it becomes more likely that the model computes negative final consumptions values as can be seen from the transition of red colors into gray when following the x-axis from left to right.

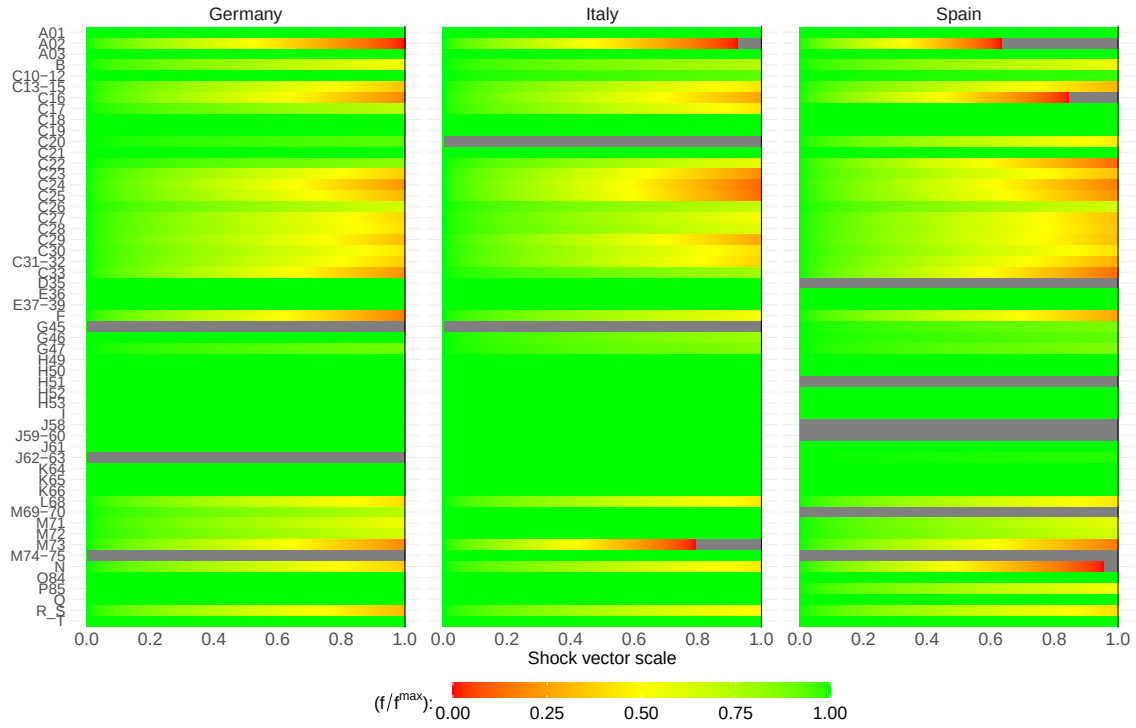

Figure D.2: **Infeasible results for final consumption for the MEEM model as a function of shock size.** The parameter α that scales the supply and demand shocks varies along the x-axis. The color codes indicate the ratio $f/f^{\max}$ where f is the MEEM result of final consumption. Note that this ratio is always equal to one (green color) for industries which are demand constrained. Infeasible values, $f \notin [0, f^{\max}]$, are indicated in gray.

D.3 Details on network density

In Chapter 4.4.3.2 we removed links randomly and repeated this procedure multiple times for any desired network density value. While random edge removal is one possible approach, other procedures could be followed too. A natural alternative to random edge deletion is to first delete small links. While the high aggregation of the data results in almost complete graphs, link sizes are highly heterogeneous, implying the existence of many very small links (McNerney et al. 2013, Cerina et al. 2015). It could be argued that many of these links are rather an artifact of data aggregation instead of encoding a fixed production recipe. We therefore repeat the procedure of Chapter 4.4.3.2 but eliminate smaller before larger links to achieve a given level of network density. Doing this results in Fig. D.3 which indicates qualitatively similar results as Fig. 4.3 of the main text.

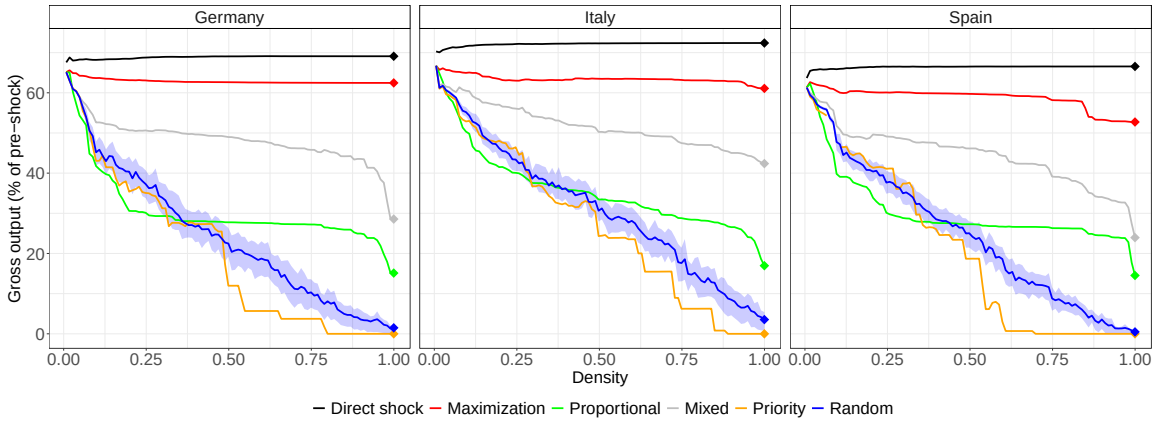

Figure D.3: **Economic impact as a function of network density.** The figure is the same as Fig. 4.3 in the main text, except that network density is changed by eliminating links based on their size instead of random deletion.

The elimination of existing IO links changes key properties of the underlying economic system. As discussed in Chapter 4.4.3.2, removing intermediate consumption values will reduce aggregate output. Figs. D.4(a) and D.4(b) visualize the relationship between output and network density following the random link and "smallest-first" removal approach, respectively. Similarly, the ratio intermediate consumption over aggregate output will be reduced as a consequence. A density value equal to zero means that there is no intermediate consumption left and firms only use primary factors as inputs in production. Fig. D.4 also shows that the average output multiplier decreases when making the network sparser, although not necessarily monotonously. Overall, the economic indicators change fairly linearly with respect to network density if links are randomly removed, whereas these relationships are highly nonlinear if smaller edges are eliminated before larger ones.

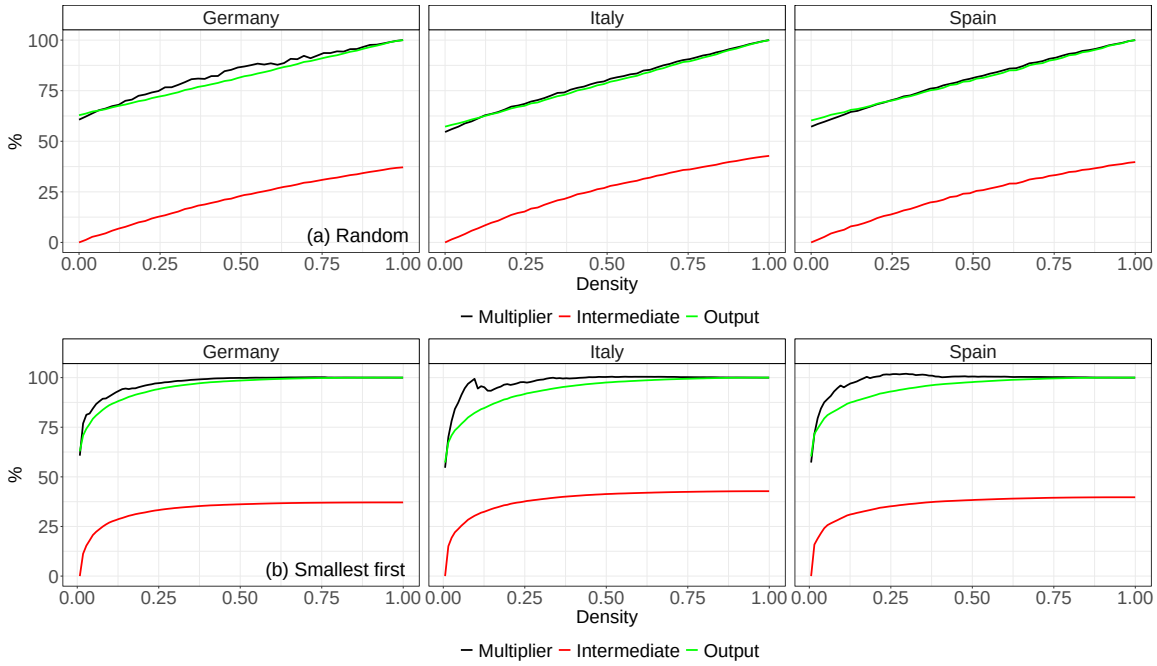

Figure D.4: **Key economic measures as a function of network density.** (a) The network density is changed by eliminating links randomly (corresponding to Fig. 4.3). (b) The network density is changed by eliminating smaller before larger links (corresponding to Fig. D.3). *Multiplier* refers to the (unweighted) average multiplier of the economy, $\sum_{ij} L_{ij}/N$, after rebalancing as percentage of the initial economy. *Intermediate* denotes the share of intermediate consumption in total output, $\sum_{ij} Z_{ij} / \sum_i x_i$ after rebalancing the economy. *Output* is total output after rebalancing divided by initial total output.

D.4 Details on shock magnitude effects

In the main text we have shown the effect of scaling either supply or demand shocks on aggregate output. We also explored shock amplification effects when scaling both, supply and demand shocks, simultaneously. Fig. D.5 shows the effect of different shock magnitudes on aggregate output and final consumption values. Results are fairly similar for aggregate values of output and consumption and are also in qualitative agreement with the results presented in Fig. 4.2.

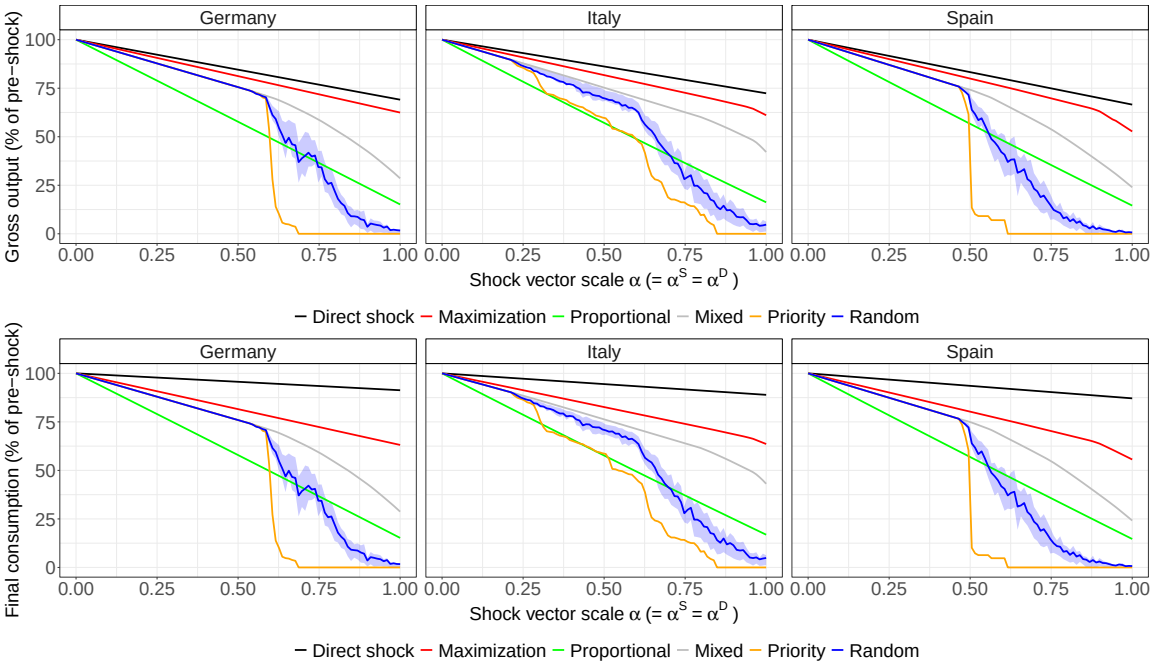

Figure D.5: **Economic impact as a function of shock magnitude.** Aggregate gross output and final consumption levels as a function of scaling demand and supply shocks equally between zero and one ($\alpha = \alpha^S = \alpha^D \in [0, 1]$).

Bibliography

Acemoglu, D., Akcigit, U. & Kerr, W. R. (2016), 'Innovation network', *Proceedings of the National Academy of Sciences* **113**(41), 11483–11488.

Acemoglu, D., Carvalho, V. M., Ozdaglar, A. & Tahbaz-Salehi, A. (2012), 'The network origins of aggregate fluctuations', *Econometrica* **80**(5), 1977–2016.

Aghion, P. & Howitt, P. (1990), A model of growth through creative destruction, Technical Report 3223, National Bureau of Economic Research.

Ahmadpoor, M. & Jones, B. F. (2017), 'The dual frontier: Patented inventions and prior scientific advance', *Science* **357**(6351), 583–587.

Alstott, J., Triulzi, G., Yan, B. & Luo, J. (2017), 'Mapping technology space by normalizing patent networks', *Scientometrics* **110**(1), 443–479.

Andersen, A. L., Hansen, E. T., Johannesen, N. & Sheridan, A. (2020), 'Consumer reponses to the COVID-19 crisis: Evidence from bank account transaction data', *SSRN 3609814* .

Ando, A., Fisher, F. M. & Simon, H. A. (1963), *Essays on the structure of social science models*, MIT Press.

Anselin, L. (2013), *Spatial econometrics: methods and models*, Vol. 4, Springer Science & Business Media.

Antràs, P., Chor, D., Fally, T. & Hillberry, R. (2012), 'Measuring the upstreamness of production and trade flows', *American Economic Review* **102**(3), 412–16.

Armstrong, J. S. (1985), *Long-Range Forecasting: From Crystal Ball to Computer*, Vol. 2, Wiley.

Arthur, W. B. (2009), *The nature of technology: What it is and how it evolves*, Simon and Schuster.

Arto, I., Andreoni, V. & Rueda Cantuche, J. M. (2015), 'Global impacts of the automotive supply chain disruption following the Japanese earthquake of 2011', *Economic Systems Research* **27**(3), 306–323.

Avelino, A. F. & Dall'erba, S. (2019), 'Comparing the economic impact of natural disasters generated by different input–output models: An application to the 2007 Chehalis river flood (WA)', *Risk Analysis* **39**(1), 85–104.

Ayres, R. U. (1969), *Technological forecasting and long-range planning*, McGraw-Hill Book Company.

Bagley, R. J., Farmer, J. D. & Fontana, W. (1992), 'Evolution of a metabolism', *Artificial life II* **10**, 141–158.

Bak, P., Chen, K., Scheinkman, J. & Woodford, M. (1993), 'Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics', *Ricerche economiche* **47**(1), 3–30.

Balint, T., Lamperti, F., Mandel, A., Napoletano, M., Roventini, A. & Sapio, A. (2017), 'Complexity and the economics of climate change: a survey and a look forward', *Ecological Economics* **138**, 252–265.

Baqaee, D. R. & Farhi, E. (2020), Nonlinear production networks with an application to the Covid-19 crisis, Technical Report 27281, National Bureau of Economic Research.

Barrot, J.-N., Grassi, B. & Sauvagnat, J. (2020), 'Sectoral effects of social distancing', *Covid Economics 3, 10 April 2020: 85-102* .

Barrot, J.-N. & Sauvagnat, J. (2016), 'Input specificity and the propagation of idiosyncratic shocks in production networks', *The Quarterly Journal of Economics* **131**(3), 1543–1592.

Basurto, A., Dawid, H., Harting, P., Hepp, J. & Kohlweyer, D. (2020), Economic and epidemic implications of virus containment policies: insights from agent-based simulations, Technical Report 05, Bielefeld Working Papers in Economics and Management.

Bates, D. & Maechler, M. (2019), *Matrix: Sparse and Dense Matrix Classes and Methods*. R package version 1.2-18.
URL: *https://CRAN.R-project.org/package=Matrix*

Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B. & Stiglitz, J. E. (2007), 'Credit chains and bankruptcy propagation in production networks', *Journal of Economic Dynamics and Control* **31**(6), 2061–2084.

Bonadio, B., Huo, Z., Levchenko, A. A. & Pandalai-Nayar, N. (2020), Global supply chains in the pandemic, Technical Report 27224, National Bureau of Economic Research.

Borsos, A. & Mero, B. (2020), Shock propagation in the banking system with real economy feedback, Technical Report Working papers 6, Magyar Nemzeti Bank (Central Bank of Hungary).

Borsos, A. & Stancsics, M. (2020), Unfolding the hidden structure of the hungarian multi-layer firm network, Technical Report Occasional papers 139, Magyar Nemzeti Bank (Central Bank of Hungary).

Breschi, S., Lissoni, F. & Malerba, F. (2003), 'Knowledge-relatedness in firm technological diversification', *Research Policy* **32**(1), 69–87.

Brinca, P., Duarte, J. B. & Faria e Castro, M. (2020), Measuring labor supply and demand shocks during COVID-19, Technical report, Federal Reserve Bank of St. Louis, St. Louis, Missouri.

Carvalho, V. M., Hansen, S., Ortiz, Á., Garcia, J. R., Rodrigo, T., Rodriguez Mora, S. & Ruiz de Aguirre, P. (2020), Tracking the COVID-19 crisis with high-resolution transaction data, Technical Report DP14642, CEPR.

Carvalho, V. M. & Tahbaz-Salehi, A. (2019), 'Production networks: A primer', *Annual Review of Economics* **11**, 635–663.

Cerina, F., Zhu, Z., Chessa, A. & Riccaboni, M. (2015), 'World input-output network', *PloS one* **10**(7), e0134025.

Chen, H., Qian, W. & Wen, Q. (2021), 'The impact of the COVID-19 pandemic on consumption: Learning from high-frequency transaction data', *American Economic Association Papers and Proceedings* **111**, 307–11.

Chetty, R., Friedman, J. N., Hendren, N., Stepner, M. et al. (2020), How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data, Technical report, National Bureau of Economic Research.

Clements, M. P. & Hendry, D. F. (2001), 'Forecasting with difference-stationary and trend-stationary models', *The Econometrics Journal* **4**(1), 1–19.

Congressional Budget Office (2006), 'Potential influenza pandemic: Possible macroeconomic effects and policy issues'. `https://www.cbo.gov/sites/default/files/109th-congress-2005-2006/reports/12-08-birdflu.pdf`.

Cowan, R. & Jonard, N. (2003), 'The dynamics of collective invention', *Journal of Economic Behavior & Organization* **52**(4), 513–532.

De Mesnard, L. (2009), 'Is the Ghosh model interesting?', *Journal of Regional Science* **49**(2), 361–372.

del Rio-Chanona, R. M., Mealy, P., Pichler, A., Lafond, F. & Farmer, J. D. (2020), 'Supply and demand shocks in the covid-19 pandemic: An industry and occupation perspective', *Oxford Review of Economic Policy* **36**(1), 94–137.

Delli Gatti, D. & Grugni, E. (2021), Breaking bad: Supply chain disruptions in a streamlined agent based model, Technical Report 9029, CESifo.

Delli Gatti, D. & Reissl, S. (2020), ABC: An agent based exploration of the macroeconomic effects of Covid-19, Technical Report 8763, CESifo.

Demir, B., Javorcik, B., Michalski, T. K., Ors, E. et al. (2020), Financial constraints and propagation of shocks in production networks, Technical Report 23, University of Nottingham, Globalisation, Productivity and Technology Programme.

Diem, C., Borsos, A., Reisch, T., Kertész, J. & Thurner, S. (2021), 'Quantifying firm-level economic systemic risk from nation-wide supply networks', *arXiv preprint arXiv:2104.07260* .

Dietzenbacher, E. & Miller, R. E. (2015), 'Reflections on the inoperability input–output model', *Economic Systems Research* **27**(4), 478–486.

Dingel, J. & Neiman, B. (2020), 'How many jobs can be done at home?', *Covid Economics* **1**, 16–24.

Dosi, G. (1982), 'Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change', *Research Policy* **11**(3), 147–162.

Duchin, F. (2005), 'A world trade model based on comparative advantage with m regions, n goods, and k factors', *Economic Systems Research* **17**(2), 141–162.

Duchin, F. & Levine, S. H. (2012), 'The rectangular sector-by-technology model: not every economy produces every product and some products may rely on several technologies simultaneously', *Journal of Economic Structures* **1**(1), 1–11.

Elhorst, J. P. (2014), *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, Vol. 479, Springer.

EPO & USPTO (2017), Guide to the CPC (Cooperative Patent Classification), Technical report.

Ertur, C. & Koch, W. (2007), 'Growth, technological interdependence and spatial externalities: theory and evidence', *Journal of Applied Econometrics* **22**(6), 1033–1062.

Evans, G. W. & Honkapohja, S. (2012), *Learning and expectations in macroeconomics*, Princeton University Press.

Fadinger, H. & Schymik, J. (2020), 'The costs and benefits of home office during the Covid-19 pandemic: Evidence from infections and an input-output model for germany', *Covid Economics* **9**, 107–134.

Fana, M., Tolan, S., Perez, S. T., Brancati, M. C. U., Macias, E. F. et al. (2020), The COVID confinement measures and eu labour markets, Technical Report JRC120578, Joint Research Centre (Seville).

Farmer, J. D., Kauffman, S. A. & Packard, N. H. (1986), 'Autocatalytic replication of polymers', *Physica D: Nonlinear Phenomena* **22**(1-3), 50–67.

Farmer, J. D. & Lafond, F. (2016), 'How predictable is technological progress?', *Research Policy* **45**(3), 647–665.

Farmer, J., Hepburn, C., Ives, M., Hale, T., Wetzer, T., Mealy, P., Rafaty, R., Srivastav, S. & Way, R. (2019), 'Sensitive intervention points in the post-carbon transition', *Science* **364**(6436), 132–134.

Fink, T. & Reeves, M. (2019), 'How much can we influence the rate of innovation?', *Science Advances* **5**(1), eaat6107.

Fink, T., Reeves, M., Palma, R. & Farr, R. (2017), 'Serendipity and strategy in rapid innovation', *Nature Communications* **8**(1).

Fleming, L. (2001), 'Recombinant uncertainty in technological search', *Management Science* **47**(1), 117–132.

Freeman, C. & Soete, L. (2012), *The economics of industrial innovation*, Routledge.

Funk, S., Salathé, M. & Jansen, V. A. (2010), 'Modelling the influence of human behaviour on the spread of infectious diseases: a review', *Journal of the Royal Society Interface* **7**(50), 1247–1256.

Galbusera, L. & Giannopoulos, G. (2018), 'On input-output economic models in disaster impact assessment', *International Journal of Disaster Risk Reduction* **30**, 186–198.

Ghosh, A. (1958), 'Input-output approach in an allocation system', *Economica* **25**(97), 58–64.

Griliches, Z. (1990), 'Patent statistics as economic indicators: A survey', *Journal of Economic Literature* **28**(4), 1661.

Grossman, G. M. & Helpman, E. (1991), 'Quality ladders in the theory of growth', *The Review of Economic Studies* **58**(1), 43–61.

Gruver, G. W. (1989), 'On the plausibility of the supply-driven input-output model: A theoretical basis for input-coefficient change', *Journal of Regional Science* **29**(3), 441–450.

Guan, D., Wang, D., Hallegatte, S., Davis, S. J., Huo, J., Li, S., Bai, Y., Lei, T., Xue, Q., Coffman, D. et al. (2020), 'Global supply-chain effects of covid-19 control measures', *Nature Human Behaviour* pp. 1–11.

Haimes, Y. Y. & Jiang, P. (2001), 'Leontief-based model of risk in complex interconnected infrastructures', *Journal of Infrastructure Systems* **7**(1), 1–12.

Hall, B. H., Jaffe, A. B. & Trajtenberg, M. (2001), The NBER patent citation data file: Lessons, insights and methodological tools, Technical report, National Bureau of Economic Research.

Hall, B. H., Jaffe, A. & Trajtenberg, M. (2005), 'Market value and patent citations', *RAND Journal of Economics* pp. 16–38.

Hall, B. H., Mairesse, J. & Mohnen, P. (2010), Measuring the Returns to R&D, *in* 'Handbook of the Economics of Innovation', Vol. 2, Elsevier, pp. 1033–1082.

Hallegatte, S. (2008), 'An adaptive regional input-output model and its application to the assessment of the economic cost of katrina', *Risk Analysis* **28**(3), 779–799.

Hallegatte, S. (2014), 'Modeling the role of inventories and heterogeneity in the assessment of the economic costs of natural disasters', *Risk Analysis* **34**(1), 152–167.

Henriet, F., Hallegatte, S. & Tabourier, L. (2012), 'Firm-network characteristics and economic robustness to natural disasters', *Journal of Economic Dynamics and Control* **36**(1), 150–167.

Higham, N. J. (2002), 'Computing the nearest correlation matrix — a problem from finance', *IMA Journal of Numerical Analysis* **22**(3), 329–343.

Ho, C.-Y., Wang, W. & Yu, J. (2018), 'International knowledge spillover through trade: A time-varying spatial panel data approach', *Economics Letters* **162**, 30–33.

Hötte, K. (2021), 'Demand-pull and technology-push: What drives the direction of technological change? An empirical network-based approach', *arXiv preprint arXiv:2104.04813* .

Hötte, K., Pichler, A. & Lafond, F. (2021), 'The rise of science in low-carbon energy technologies', *Renewable and Sustainable Energy Reviews* **139**, 110654.

Huneeus, F. (2018), Production network dynamics and the propagation of shocks, Technical report, Princeton University.

Inoue, H. & Todo, Y. (2019), 'Firm-level propagation of shocks through supply-chain networks', *Nature Sustainability* **2**(9), 841–847.

Inoue, H. & Todo, Y. (2020), 'The propagation of the economic impact through supply chains: The case of a mega-city lockdown against the spread of COVID-19', *Covid Economics* **2**, 43–59.

Jain, S. & Krishna, S. (2001), 'A model for the emergence of cooperation, interdependence, and structure in evolving networks', *Proceedings of the National Academy of Sciences* **98**(2), 543–547.

Jantsch, E. (1967), *Technological forecasting in perspective*, Vol. 3, Citeseer.

Jolliffe, I. T. (2002), *Principal component analysis*, Vol. 2, Springer.

Jones, C. I. (1995), 'R&D-based models of economic growth', *Journal of Political Economy* **103**(4), 759–784.

Kauffman, S. A. (1993), *The origins of order: Self-organization and selection in evolution*, OUP USA.

Kerschner, C. & Hubacek, K. (2009), 'Erratum to "Assessing the suitability of input-output analysis for enhancing our understanding of potential effects of peak-oil"', *Energy* **34**(10), 1662–1668.

Klimek, P., Poledna, S. & Thurner, S. (2019), 'Quantifying economic resilience from input-output susceptibility to improve predictions of economic growth and recovery', *Nature Communications* **10**(1), 1–9.

Kogan, L., Papanikolaou, D., Seru, A. & Stoffman, N. (2017), 'Technological innovation, resource allocation, and growth', *The Quarterly Journal of Economics* **132**(2), 665–712.

Koks, E. E., Bočkarjova, M., de Moel, H. & Aerts, J. C. (2015), 'Integrated direct and indirect flood risk modeling: development and sensitivity analysis', *Risk Analysis* **35**(5), 882–900.

Koks, E. E. & Thissen, M. (2016), 'A multiregional impact assessment model for disaster analysis', *Economic Systems Research* **28**(4), 429–449.

König, M. D., Battiston, S., Napoletano, M. & Schweitzer, F. (2011), 'Recombinant knowledge and the evolution of innovation networks', *Journal of Economic Behavior & Organization* **79**(3), 145–164.

Koren, M. & Pető, R. (2020), 'Business disruptions from social distancing', *Covid Economics* **2**, 13–31.

Kumar, A., Chakrabarti, A. S., Chakraborti, A. & Nandi, T. (2021), 'Distress propagation on production networks: Coarse-graining and modularity of linkages', *Physica A: Statistical Mechanics and its Applications* **568**.

Lafond, F., Bailey, A. G., Bakker, J. D., Rebois, D., Zadourian, R., McSharry, P. & Farmer, J. D. (2018), 'How well do experience curves predict technological progress? A method for making distributional forecasts', *Technological Forecasting and Social Change* **128**, 104–117.

Lafond, F. & Kim, D. (2019), 'Long-run dynamics of the US patent classification system', *Journal of Evolutionary Economics* **29**, 1–34.

Leontief, W. W. (1936), 'Quantitative input and output relations in the economic systems of the United States', *The Review of Economics and Statistics* pp. 105–125.

Li, J., Crawford-Brown, D., Syddall, M. & Guan, D. (2013), 'Modeling imbalanced economic recovery following a natural disaster using input-output analysis', *Risk Analysis* **33**(10), 1908–1923.

Long, J. B. & Plosser, C. I. (1983), 'Real business cycles', *Journal of Political Economy* **91**(1), 39–69.

Magee, C. L., Basnet, S., Funk, J. L. & Benson, C. L. (2016), 'Quantitative empirical trends in technical performance', *Technological Forecasting and Social Change* **104**, 237–246.

Mandel, A. & Veetil, V. (2020*a*), 'The economic cost of COVID lockdowns: An out-of-equilibrium analysis', *Economics of Disasters and Climate Change* **4**, 431–451.

Mandel, A. & Veetil, V. P. (2020*b*), 'Disequilibrium propagation of quantity constraints: Application to the COVID lockdowns', *SSRN 3631014* .

Marco, A. C., Carley, M., Jackson, S. & Myers, A. (2015), 'The USPTO historical patent data files. two centuries of innovation'.

Martino, J. P. (1993), *Technological forecasting for decision making*, McGraw-Hill, Inc.

Masuda, N. & Lambiotte, R. (2016), *A Guide to Temporal Networks*, World Scientific.

McNerney, J., Farmer, J. D., Redner, S. & Trancik, J. E. (2011), 'Role of design complexity in technology improvement', *Proceedings of the National Academy of Sciences* **108**(22), 9008–9013.

McNerney, J., Fath, B. D. & Silverberg, G. (2013), 'Network structure of inter-industry flows', *Physica A: Statistical Mechanics and its Applications* **392**(24), 6427–6441.

McNerney, J., Savoie, C., Caravelli, F. & Farmer, J. D. (2018), 'How production networks amplify economic growth', *arXiv preprint arXiv:1810.07774* .

Miller, R. E. & Blair, P. D. (2009), *Input-output analysis: foundations and extensions*, Cambridge University Press.

Miller, R. E. & Temurshoev, U. (2017), 'Output upstreamness and input downstreamness of industries/countries in world production', *International Regional Science Review* **40**(5), 443–475.

Mincer, J. A. & Zarnowitz, V. (1969), The evaluation of economic forecasts, *in* 'Economic forecasts and expectations: Analysis of forecasting behavior and performance', NBER, pp. 3–46.

Mokyr, J. (1992), *The lever of riches: Technological creativity and economic progress*, Oxford University Press.

Muellbauer, J. (2020), 'The coronavirus pandemic and U.S. consumption', *VoxEU.org, 11 April* . `https://voxeu.org/article/coronavirus-pandemic-and-us-consumption`.

Nagy, B., Farmer, J. D., Bui, Q. M. & Trancik, J. E. (2013), 'Statistical basis for predicting technological progress', *PloS one* **8**(2).

Napolitano, L., Evangelou, E., Pugliese, E., Zeppini, P. & Room, G. (2018), 'Technology networks: the autocatalytic origins of innovation', *Royal Society Open Science* **5**(6), 172445.

National Research Council (2009), 'Persistent forecasting of disruptive technologies'.

Oosterhaven, J. (1988), 'On the plausibility of the supply-driven input-output model', *Journal of Regional Science* **28**(2), 203–217.

Oosterhaven, J. (2017), 'On the limited usability of the inoperability IO model', *Economic Systems Research* **29**(3), 452–461.

Oosterhaven, J. & Bouwmeester, M. C. (2016), 'A new approach to modeling the impact of disruptive events', *Journal of Regional Science* **56**(4), 583–595.

Ord, K. (1975), 'Estimation methods for models of spatial interaction', *Journal of the American Statistical Association* **70**(349), 120–126.

Pangallo, M. (2020), 'Synchronization of endogenous business cycles', *SSRN 3538999* .

Peydró, J.-L., Jiménez, G., Huremovic, K., Moral-Benito, E. & Vega-Redondo, F. (2020), Production and financial networks in interplay: Crisis evidence from supplier-customer and credit registers, Technical Report Discussion Paper No. DP15277, CEPR.

Pichler, A. & Farmer, J. D. (2021), 'Simultaneous supply and demand constraints in input-output networks: The case of Covid-19 in Germany, Italy, and Spain', *Forthcoming in Economics Systems Research* .

Pichler, A., Lafond, F. & Farmer, J. D. (2020), 'Technological interdependencies predict innovation dynamics', *arXiv preprint arXiv:2003.00580* .
URL: *https://arxiv.org/abs/2003.00580*

Pichler, A., Lafond, F. & Farmer, J. D. (2021), 'Persistence and predictability of patenting', *In preparation* .

Pichler, A., Pangallo, M., del Rio-Chanona, R. M., Lafond, F. & Farmer, J. D. (2020), 'Production networks and epidemic spreading: How to restart the UK economy?', *Covid Economics* **23**, 79–151.
URL: *https://arxiv.org/abs/2005.10585*

Pichler, A., Pangallo, M., del Rio-Chanona, R. M., Lafond, F. & Farmer, J. D. (2021), 'In and out of lockdown: Propagation of supply and demand shocks in a dynamic input-output model', *SSRN 3788494* .
URL: *http://dx.doi.org/10.2139/ssrn.3788494*

Pless, J., Hepburn, C. & Farrell, N. (2020), 'Bringing rigour to energy innovation policy evaluation', *Nature Energy* pp. 1–7.

Poledna, S., Hochrainer-Stigler, S., Miess, M. G., Klimek, P., Schmelzer, S., Sorger, J., Shchekinova, E., Rovenskaya, E., Linnerooth-Bayer, J., Dieckmann, U. et al. (2018), 'When does a disaster become a systemic event? Estimating indirect economic losses from natural disasters', *arXiv preprint arXiv:1801.09740* .

Poledna, S., Miess, M. G. & Hommes, C. H. (2019), 'Economic forecasting with an agent-based model', *SSRN 3484768* .

Popp, D. (2016), 'Economic analysis of scientific publications and implications for energy research and development', *Nature Energy* **1**(4), 1–8.

Porter, A. L., Cunningham, S. W., Banks, J., Roper, A. T., Mason, T. W. & Rossini, F. A. (2011), *Forecasting and management of technology*, Wiley Online Library.

Reissl, S., Caiani, A., Lamperti, F., Guerini, M., Vanni, F., Fagiolo, G., Ferraresi, T., Ghezzi, L., Napoletano, M. & Roventini, A. (2021), Assessing the economic effects of lockdowns in italy: a dynamic input-output approach, Technical Report 2021/03, LEM.

Richiardi, M., Bronka, P. & Collado, D. (2020), 'The economic consequences of COVID-19 lockdown in the UK. An input-output analysis using consensus scenarios', *Working paper* .

Romer, D. (2012), *Advanced Macroeconomics*, McGraw-Hill, New York.

Romer, P. M. (1990), 'Endogenous technological change', *Journal of Political Economy* **98**(5/2), 71–102.

Sampson, M. (1991), 'The effect of parameter uncertainty on forecast variances and confidence intervals for unit root and trend stationary time-series models', *Journal of Applied Econometrics* **6**(1), 67–76.

Santos, J. R. & Haimes, Y. Y. (2004), 'Modeling the demand reduction input-output (I-O) inoperability due to terrorism of interconnected infrastructures', *Risk Analysis* **24**(6), 1437–1451.

Schmookler, J. (1966), *Invention and economic growth*, Harvard University Press.

Schumpeter, J. A. (1939), *Business Cycles*, McGraw-Hill Book Company, New York.

Sharma, D., Bouchaud, J.-P., Gualdi, S., Tarzia, M. & Zamponi, F. (2021), 'V–, U–, L–or W–shaped economic recovery after Covid-19: Insights from an agent based model', *PloS one* **16**(3), e0247823.

Solow, R. M. (1956), 'A contribution to the theory of economic growth', *The Quarterly Journal of Economics* **70**(1), 65–94.

Spray, J. (2017), Reorganise, replace or expand? The role of the supply-chain in first-time exporting, Technical Report Cambridge Working Papers in Economics: 1741, Cambridge-INET.

Steenge, A. E. & Bočkarjova, M. (2007), 'Thinking about imbalances in post-catastrophe economies: an input–output based proposition', *Economic Systems Research* **19**(2), 205–223.

Steinback, S. R. (2004), 'Using ready-made regional input-output models to estimate backward-linkage effects of exogenous output shocks', *Review of Regional Studies* **34**(1), 57–71.

Surana, K., Doblinger, C., Anadon, L. D. & Hultman, N. (2020), 'Effects of technology complexity on the emergence and evolution of wind industry manufacturing locations along global value chains', *Nature Energy* **5**(10), 811–821.

Surico, P., Känzig, D. & Hacioglu, S. (2020), Consumption in the time of Covid-19: Evidence from UK transaction data, Technical Report DP14733, CEPR.

Taalbi, J. (2020), 'Evolution and structure of technological systems-an innovation output network', *Research Policy* **49**(8), 104010.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. & De Vries, G. J. (2015), 'An illustrated user guide to the world input–output database: the case of global automotive production', *Review of International Economics* **23**(3), 575–605.

Tria, F., Loreto, V., Servedio, V. D. P. & Strogatz, S. H. (2014), 'The dynamics of correlated novelties', *Scientific Reports* **4**, 5890.

Triulzi, G., Alstott, J. & Magee, C. L. (2020), 'Estimating technology performance improvement rates by mining patent data', *Technological Forecasting and Social Change* **158**, 120100.

Tsay, R. S. (2010), *Analysis of financial time series*, 3 edn, John Wiley & Sons, Hoboken, New Jersey.

Usher, A. P. (1954), *A History of Mechanical Invention*, Harvard University Press, Cambridge, MA.

Verbeek, M. (2017), *A guide to modern econometrics*, 5 edn, John Wiley & Sons, Hoboken, New Jersey.

Wang, W. & Yu, J. (2015), 'Estimation of spatial panel data models with time varying spatial weights matrices', *Economics Letters* **128**, 95–99.

Wang, W., Zhang, X. & Paap, R. (2019), 'To pool or not to pool: What is a good strategy for parameter estimation and forecasting in panel regressions?', *Journal of Applied Econometrics* **34**(5), 724–745.

Youn, H., Strumsky, D., Bettencourt, L. M. & Lobo, J. (2015), 'Invention as a combinatorial process: evidence from US patents', *Journal of The Royal Society Interface* **12**(106), 1–8.

Yule, G. U. (1926), 'Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series', *Journal of the Royal Statistical Society* **89**(1), 1–63.