I.S.S.N: 1885-6888

ECONOMIC HISTORY

WORKING PAPER SERIES

A VAR ANALYSIS OF THE TRANSPORTATION REVOLUTION IN EUROPE

Marta Felis-Rota (UAM)

Working Paper 01/2014

UNIVERSIDAD AUTONOMA DE MADRID

DEPARTAMENTO DE ANÁLISIS ECONÓMICO: TEORÍA ECONÓMICA E HISTORIA ECONÓMICA

To be presented at the ASSA Meetings 2015, 3-5 January, Boston

A VAR Analysis of the Transportation Revolution in Europe

Marta Felis-Rota[1]

(First version: December 2014)

Abstract

During the second half of the 19th century transportation costs decreased sharply. Among the most notable technological advances that lead to the transportation revolution we find the arrival of the railways. This paper provides a quantitative analysis of the expansion of the railways at the time of the so-called First Globalization in European countries through a vector autoregressive analysis. Total mileage of the railways has been obtained through GIS software for every European country, via a long process of digitalization of historical atlases. Then the vector autoregressive analysis and the impulse-response functions show the interaction between railways and GDP. I find interactions going in both directions of the VAR, and that the persistence of the effects varies from country to country. Thanks to this method, we can compare differentiated patterns of development associated to idiosyncratic transportation revolutions in Europe. (139 words)

Keywords VAR, railways, growth, comparative economic history

JEL Classification C3 N13 R4

[1] Department of Economic Analysis: Economic Theory and Economic History, Universidad Autónoma de Madrid, 28049, Cantoblanco, Madrid, Spain. E-mail: marta.felis@uam.es.

1 Introduction

The European integration process started long ago. We could go back to the Roman Empire and undeniably think of the political annexation of neighboring territories as a European integration process. On the transport and communications' sphere, a plain glance at the map of roads in Europe in the Roman Empire is enough to realize how economic integration must have been taking place at that time (see Sitwell, 1981 for Roman roads in Europe). But my aim here is not going so much further back in time. I look back into the second half of the 19th century. That was the time of the first globalization, when transportation costs decreased sharply thanks to the arrival of the railways, and the steam-propelled iron-hulled ships.

O'Rourke and Williamson (1999) describe how important the fall of transportation costs was for economic integration, associated to price convergence. We see an unprecedented increase in the movements of factors of production worldwide in this period. We can observe an increase in the share of intercontinental trade, especially in the case of Europe (Federico and Tena, 2013).

Understanding that all this movement of capital and workers could have never taken place without the corresponding transportation and communications revolution is of vital importance. Not only railways and steam ships would enter the scene, but also improved postal services and very notably the telegraph facilitated the transmission of information as well as transportation on goods, services, and people. In this paper, I focus on the relevance of the arrival of the railway into the different European countries in terms of associated economic growth.

Although this is an old question, the quantitative approach is new. The most well-known literature on economic development and expansion of the railways refers to the United States (Fogel 1964; and more recently Herrendorf, Schmitz and Teixeira 2009; and Donaldson and Hornbeck 2013), there is a recent renovated interest in the economic impact of the railways in other less-developed regions of the world, like Donaldson (2010) for the Indian case and Banerjee, Duflo, and Quian (2012) for the Chinese case. This renovated interest in the study of the impact of the railways is justified by the use of new empirical methodologies, such as the use of geographical information systems. The usefulness of these new tools for the study of historical geography is explained in Martí-Henneberg (2011).

There have been some previous studies on the impact of the railways to economic development of Europe. For instance, Herranz-Locan (2006) would provide an account of the economic impact of the railways in nineteenth-century Spain. But so far, most of the attention has been focused on the synergies between railway and urbanization. Gregory and Martí-Henneberg (2010) perform an analysis of the population distribution of England and Wales after the arrival of the railways; Mojica and Martí-Henneberg (2011) undertake a similar approach for France, Spain, and Portugal.

This paper provides a quantitative analysis of the expansion of the railways in European countries through impulse-response functions. Total mileage of the railways has been obtained through GIS software for every European country, via a long process of digitalization of historical atlases. Then a vector autoregressions analysis shows the interaction between railways and GDP and the persistence of the effect country by country. Thanks to this method, we can compare differentiated patterns of development associated to idiosyncratic 19th century transportation revolutions in Europe.

2 Data

2.1 Railways

The data on railways are obtained as part of the HGIS-e project (Historical Geographical Information Systems – Europe). This project gathers information from historical railway maps since their introduction as a new means of transportation in the 19th century. It introduces every single new railway line and station in a digital platform, where each point is associated to a geodesic point, similar to that of a GPS system of coordinates. The result is a series of historical digital maps of railway lines in Europe. Additionally, the GIS software is able to calculate the total kilometrical length of the railway network built in every European country at any given point in history. And this is the information I am using in this paper to study the co-evolution of railways and wealth.

As a result, I have data for every 10 years from 1850 to 1920 for all the main European countries. I take 1850 as the initial year because the railway lines opened in the 1830s and 1840s are still rare, and looking for significant correlations with GDP could be quite erratic. Starting in 1850 seems much more reasonable, when most of the countries had already built their first railway lines and some had a significant network. At this respect, the 1830s meant the introduction of the railway as a new means of transportation beyond the strict limits of the mine or factory in Britain. The 1840s saw the diffusion of the first industrial revolution to the first comers in the continent. And, finally, the 1850s brought the more generalized expansion of the railway. The progresses made in the 1910s close the so-called "long 19th century". So I take 1920 as the final year of the sample. As for the set of countries, you can find a list in the appendix.

2.2 GDP

There are just a limited number of GDP historical databases available which are unified for a large set of countries. I use Maddison (2010) not only because it is the more widely used, but also because it is designed to make long time series comparisons. Bolt and van Zanden (2013) produced the first global update to Maddison's last database. However, the focus of their update was on pre-1820 data. On the other hand, Prados de la Escosura (2000) introduced a methodological improvement, which makes it more suitable for cross-country comparisons at current prices. However, his are benchmark estimations, and we are particularly interested in following a time series approach in this paper.

3 Discussion of Methodology: VARs and IRFs

The vector autoregression or VAR is an econometric technique that emerged as an alternative to multivariable models where the order of causality is known. This alternative was first proposed by Sims (1980). It is based in analyzing a set of variables (or vector) as a function of their own lags. Each variable is explained as a function of their own lags, and the lags of the other variables. So it has the advantage of avoiding simultaneity associated problems. Because this technique is eluding the imposing of restrictions, it has been classified as “atheoretical macroeconomics”. Why might it be useful for us?

It is many times true --and very especially in the Social Sciences--, that it is not clear which variable causes which. The statistical study of the history of the railways has typically imposed the causality on one direction by necessity more than believe, that of trains pushing the economy further and not the other way around. Think of the Social Savings literature initiated by Fogel (1964), for instance. But it is also obvious that higher income levels would motivate and facilitate the construction of the railway in the first place. So this seems a clear case where causality can go both directions. Not only that, we should possibly not even aim at achieving the identification of an isolated direction of causality. Given that the synergies may be going in both directions, it seems more appropriate to analyze the co-movement of the two variables (railway development and income) together. What if we analyzed the two variables without imposing theoretical restrictions on the direction of causality? It actually seems adequate.

The use of VARs in Economic History has been fairly limited. Probably the closest example, saving the distance due to my data limitations, is that of Nicolini (2007), who analyzed the co-evolution of the economy and the population in pre-industrial England using a VAR analysis. Other studies using similar or close time series methodologies are those of Ermisch (2008), and Crafts and Mills (2009), all in historical economic demography.

Regarding the interpretation of the results of the VAR, it typically comes via three alternative and non-excluding methodologies. These are the study of the impulse-response functions (IRFs), the fixed-effects vector decomposition (FEVD), and the dynamic multiplier functions (DMFs).

The IRFs depict how the endogenous variables react over time to a shock happening to another endogenous variable; unlike the DMFs, which shows the reaction of one endogenous variable to a change happening to an exogenous variable. Because I am specifically interested in the response of one endogenous variable to a change in the other one, I focus my empirical analysis interpretation on the IRFs. DMFs are dismissed in this paper because the aim is not analyzing the effect of exogenous shocks, but studying the interaction between both endogenous variables instead.

Finally, the third alternative would be the FEVD, which is a specific case of an instrumental variables estimator, using a Hausman and Taylor estimator (Hausman and Taylor, 1981), which would control for country-specific unobservables. It serves the purpose of estimating a fixed-effects model (i.e. it estimates a time-invariant variable). The FEVD was first proposed by Plümper and Troeger in 2007. This estimator, which can be performed after a VAR analysis, calculates which fraction of the error variance of each variable can be attributed to a given innovation. Unfortunately, the FEVD proved more than controversial. There was a vivid discussion on the virtues and faults of this estimator in the scientific journal *Political Analysis* in 2011. Both William Greene (2011) and Breusch *et al.* (2011) coincide in categorically dis-recommending this method. Breusch *et al.* (2011) proved that the standard errors estimated using the FEVD are simply incorrect. Finally, the journal acknowledged their concern about the several articles which had already been published using this method (Breusch et al. 2011). So, eventually, the FEVD seems a totally non-recommendable strategy. So I will not enter the task of estimating the ins and outs of the idiosyncratic component of every country; more specifically, the fraction of the country variance explained by each innovation, and I will stick to the more traditional IRFs.

4 Analysis

4.1 Stationarity

The first step to run a VAR analysis is determining whether the variables under study are stationary or not. In order to test for unit roots (non-stationarity), I run Augmented Dickey-Fuller tests for every country i. The original Dickey-Fuller model to test is as follows:

$$Kmrail_{t,i} = \alpha_i + \beta_i t + \gamma_i Kmrail_{t-1,i} + u_{t,i}$$

$$GDP_{t,i} = \delta_i + \zeta_i t + \eta_i GDP_{t-1,i} + v_{t,i}$$

, where α_i and δ_i are constant terms, $\beta_i t$ and $\zeta_i t$ are time trends, $u_{t,i}$ and $v_{t,i}$ are error terms, and finally γ_i and η_i are the unit root coefficients to be estimated in the test (unit root means these coefficients are equal to one).

The augmented version of the test avoids the problems associated to serial correlation, and for this reason it runs the test on the variables' increments:

$$\Delta Kmrail_{t,i} = \alpha_i + \beta_i t + \gamma_i Kmrail_{t-1,i} + \delta_{1i}\Delta Kmrail_{t-1,i} + \cdots + \delta_{si}\Delta Kmrail_{t-s,i} + u_{t,i}$$

$$\Delta GDP_{t,i} = \zeta_i + \eta_i t + \varphi_i GDP_{t-1,i} + \lambda_{1i}\Delta GDP_{t-1,i} + \cdots + \lambda_{si}\Delta GDP_{t-s,i} + v_{t,i}$$

, where s is the number of incremental lags. Now the null hypothesis (unit root) is that coefficients γ_i and φ_i be zero.

When I run the Augmented Dickey-Fuller tests without lags, the result is that the null hypothesis of non-stationarity (unit root) cannot be rejected. This is true for the two variables (GDP and kilometers of railway line), and for every country i. But this result is probably due to the existence of serial correlation, which is captured by the coefficient of the plain lag in the absence of incremental lags.[2]

When we incorporate a time trend and the maximum number of lags in the ADF tests for kilometers of railway line, non-stationarity cannot be rejected (only in the

[2] The same happens when we do not include the time trend. There is a bias towards being too optimistic in diagnosing stationarity.

case of Britain when we incorporate a time trend plus the maximum number of lags). When we run the same tests for the GDP variable, the unit root persists even after including a time trend and one incremental lag. For most cases, we run short of degrees of freedom when we want to include a time trend plus two incremental lags.

When we take the natural logarithm of the kilometers of railway, it shows much greater stationarity, although it totally varies depending on the cases. GDP remains strongly non-stationary even after taking logs, including a time trend and the incremental lags. So we can conclude that these series are strongly non-stationary. Because the natural logarithm smoothes the non-stationarity, I will continue working with the logs of the variables.

Given the non-stationarity of the two variables, the next step is testing whether they are cointegrated or not. In order to test for a potential cointegration relationship, I run the Johansen's cointegration tests (Johansen, 1995). These tests reveal that kilometers of railway and GDP are not cointegrated, neither in levels nor in first differences. Therefore, we can rule out the possibility of running vector error-correction model (VEC) with these variables, since the VEC models are used in the case the series are cointegrated.

Once we have:

1. Ruled out the suitability of running a VAR in levels due to the lack of evidence of stationarity.
2. Ruled out the suitability of running a VEC due to the lack of cointegration between the two variables, both in levels and in differences.

We need looking for alternative explanations. It turns out that most of the countries return stationary series when the ADF test of stationary is run with a lag. This is, omitting the last observation, which corresponds to 1920, after the First World War (WWI). This might be due to the fact that the Great War is imposing a structural break on the database.

The database does not span far beyond 1920, except for Britain, France, Italy, Germany and Spain, for which, exceptionally, I also have the railway length in kilometers for 1930. This is enough to run a simple Chow test for structural breaks. I run a Chow test with the unbalanced panel in order to take advantage of all available information.[3] Nevertheless, I do not use the limitedly available data referring to the interwar period for the VAR later, which will be limited to the pre-WWI period (precisely because I find a structural break at the WWI).

The Chow test reveals the presence of a WWI structural break at a 95% confidence level. At the same time, there is enough evidence to show that stationarity dramatically increases when we stick to pre-WWI data only. So it seems that the solution is removing the observations corresponding to the last time-cut and perform the VAR with pre-WWI data only. For this purpose I create the series $lnKmrail_i$ and $lnGDP_i$, which stand for the log of the variables prior to the Great War. Finally, the VAR vector is

$$[lnKmrail_{ti} \quad lnGDP_{ti}]$$

and the model is:

$$lnKmrail_{ti} = \alpha_{0i} + \alpha_{1i} lnKmrail_{t-1,i} + \alpha_{2i} lnGDP_{t-1,i} + u_{ti}$$

$$lnGDP_{ti} = \beta_{0i} + \beta_{1i} lnKmrail_{t-1,i} + \beta_{2i} lnGDP_{t-1,i} + v_{ti}$$

to be estimated simultaneously. Where time *t* runs from 1850 to 1910, decade by decade; and *i* stands for country *i* for all countries *i*=1, 2, …, 13 (listed in appendix B). The alphas and betas are the parameters to be estimated, and u_{ti} and v_{ti} are the error terms.

4.2 Vector Autoregressions (VAR)

There is not much room for lag-order selection given the sample size. I am performing a VAR analysis with one lag, which is equivalent to 10 years. This means that it gathers the effect of the impact of one variable on the other for a

[3] The Chow F(2,176) statistic is 2.7011/.5904=4.575 > F(2,176) from the tables at 5% level =2.9957. Therefore, we can confirm a structural break at the WWI. This test is based on the sum of squares of the errors, and controls for the number of observations and the number of parameters estimated in total. See Chow (1960).

period of 10 years. Table 1 compiles the results for all European countries in the sample (13 countries).

When I look at how the lagged value of the log of GDP has had an impact on the log of Km. of rail, I find a significant effect for Belgium, Denmark, Sweden and Switzerland. The interpretation is that it was important to have an initial rise in the level of incomes for building the railways in those countries. For the rest, I do not find a statistically significant effect for the importance of having a previous income before building the railways. Either both things grew and developed hand-in-hand thanks to private investor's initiative (like in the case of Britain) or there was an investment initiative or facilitation through regulation on behalf of the State (France, Spain), which did not require a pre-existing income rise.

For those countries were income levels in the previous decade mattered, the coefficients can be directly interpreted as elasticities, since I run the regressions in logs. Then, from table 1 it follows that a 1% increase in the income levels is associated with around a 0.5% increase in the railway network length in the next decade, except for Sweden, where the associated percentual increase in railway length is as high as 2.1%. Think that all in this sub-sample except Sweden are geographically compact and relatively small size countries. Sweden might have experienced a stronger impact of the unification of markets thanks to the arrival of the railways due to its geography. The arrival of the railway in Sweden was very important because it longitudinally linked North and South. More puzzling is the case of Belgium, which renders a small but negative significant coefficient of magnitude 0.2%. This is hard to interpret because it means that a rise in the log of income is associated to a small fall in the log of km. of rail in the next decade. Think that Belgium is the first country to adopt the Industrial Revolution in continental Europe. This small though negative relationship, which is exclusively happening in the Belgium case, might be due to the fact that the Belgians became richer earlier and maybe started to adopt other means of transportation earlier, like improved roads, which is more typical of 20^{th}-century transportation transformations.

4.3 Impulse-Response Functions (IRF)

The impulse-response functions show the impact that a one-time one-standard deviation shock in one variable has on itself and on the other variable. These are graphically represented in appendix C, country by country. These are shown in levels instead of logarithms in order to have a more intuitive interpretation.

There are two sets of reactions we are interested in. On the one hand, what is the reaction of the variable kilometers of railway to a one standard deviation change in GDP in the pre-WWI period? Most of the countries do not show immediate responses. This is, simply the fact of being richer does not bring your country to building the railways. The exception to the norm is Norway, where we can observe a quick rise in the expansion of the railways after one period only, and then a significant permanent rising effect in the long run. For Norway, being richer brought the development of transportation, in particular, the railways.

The other question, which has been much more debated in the Economic History literature, is whether the construction of the railways brought economic growth associated to it. Is this new methodology shedding some light on this question? The impulse-response functions vary from country to country. We can differentiate two groups of countries. On the one hand, there is a group of countries for which the rise of the railways per se is not leading to a strong reaction in GDP in the impulse-response functions (left-bottom graph in the figures). Its long term effects could be larger. Germany, Finland, and especially Switzerland show a long-run upward tendency with high persistence of the impact of railways on future GDP. Unfortunately, we observe wide 95% confidence intervals for the distant future, which means that the precise long-term impact is difficult to predict with exactitude. One the other hand, we have another set of countries (Spain, France, Portugal and Britain), for which a one standard-deviation initial shock in the kilometers of railway induces a quick reaction of GDP in the next decade. The GDP impulse-response function reacts positively to a railways' positive shock for all cases. This effect lasts for about 5 periods (50 years), although it depends on the country. The shortestlasting effect takes place in Portugal. Its IRF reaction starts fading away after 1 decade. And the longest-lasting effect takes place in France, where the effect of a positive shock starts falling after 3 decades and then dies out very slowly.

5 Conclusions

Traditionally, Economic History has paid attention to the impact that the arrival of the railways had on future economic growth. In this paper, I look at the synergies between the development of the railways and income from a new perspective. I analyze the effect that railways had on future income, and simultaneously that of income in turn on railways. I find statistically significant effects in both directions in a VAR analysis.

On the one hand, the arrival of the railways during the period of the First Globalization impacted economic growth in the years to follow. Countries revealing a short-run economic impact have a long-lasting effect of about 50 years. Others have a smaller short-term impact and then react after several decades. The particularities of these relationships vary from one country to the other.

On the other hand, having a rise in income would also affect the construction of the railways in some countries. This question has been seldom addressed in the literature, and never before from a quantitative perspective. For those countries with a significant effect, a 1% increase in the income levels is associated with around a 0.5% increase in the railway network length in the next 10 years. This impact seems to vary with the geography and morphology of the country, and probably with the stage of development in which the country found itself at (although this remains a conjecture).

Acknowledgements: I would like to thank funding from the Spanish Ministry of Science and Innovation (grants ref. ECO2013-46980, and CSO2010-16389).

Appendix A

Table 1 – Pre-First World War VAR for Log Km. railway line and Log GDP for European countries, data from 1850 to 1910

	Germany	Belgium	Denmark	Spain	Finland	France	Netherl.	Italy	Norway	Portugal	Sweden	Switzerl.	Britain
Ln Km Rail													
L.Ln KmRail	0.864***	0.679***	0.295***	0.352***	.571***	0.622***	0.888***	0.625***	0.411***	0.250**	-0.0244	0.135***	0.624***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.002)	(0.012)	(0.829)	(0.003)	(0.000)
L.Ln GDP	-0.342	-0.215***	0.489***	0.273	-0.194	-0.158	-0.708	-0.163	0.680	0.723	2.132***	0.553***	-0.00289
	(0.102)	(0.002)	(0.000)	(0.349)	(0.273)	(0.227)	(0.440)	(0.774)	(0.213)	(0.196)	(0.000)	(0.001)	(0.978)
Ln GDP													
L.Ln KmRail	0.0727	-0.0530***	-0.0461***	0.0331	.150***	0.0388	0.0629	-0.128***	-0.0802***	0.0374	-0.0394*	0.0112	0.0418
	(0.199)	(0.000)	(0.003)	(0.314)	(0.000)	(0.570)	(0.258)	(0.001)	(0.000)	(0.462)	(0.088)	(0.290)	(0.624)
L.Ln GDP	0.973***	0.999***	1.247***	0.653**	.586***	0.773***	0.891***	1.972***	1.268***	0.705**	1.224***	1.010***	0.863***
	(0.000)	(0.000)	(0.000)	(0.014)	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.014)	(0.000)	(0.000)	(0.000)
N decades	6	6	5	6	4	6	6	6	5	5	5	6	6
AIC	-31.55	-61.74	-36.35	-15.81	-	-28.63	-14.68	-21.53	-28.27	-9.076	-23.76	-23.49	-40.17
BIC	-32.80	-62.99	-38.70	-17.06	-	-29.88	-15.93	-22.78	-30.61	-11.42	-26.10	-24.74	-41.42

Notes: L. indicates lagged variable. Constant omitted. P-values in parentheses *p<0.1, **p<0.05, ***p<0.001

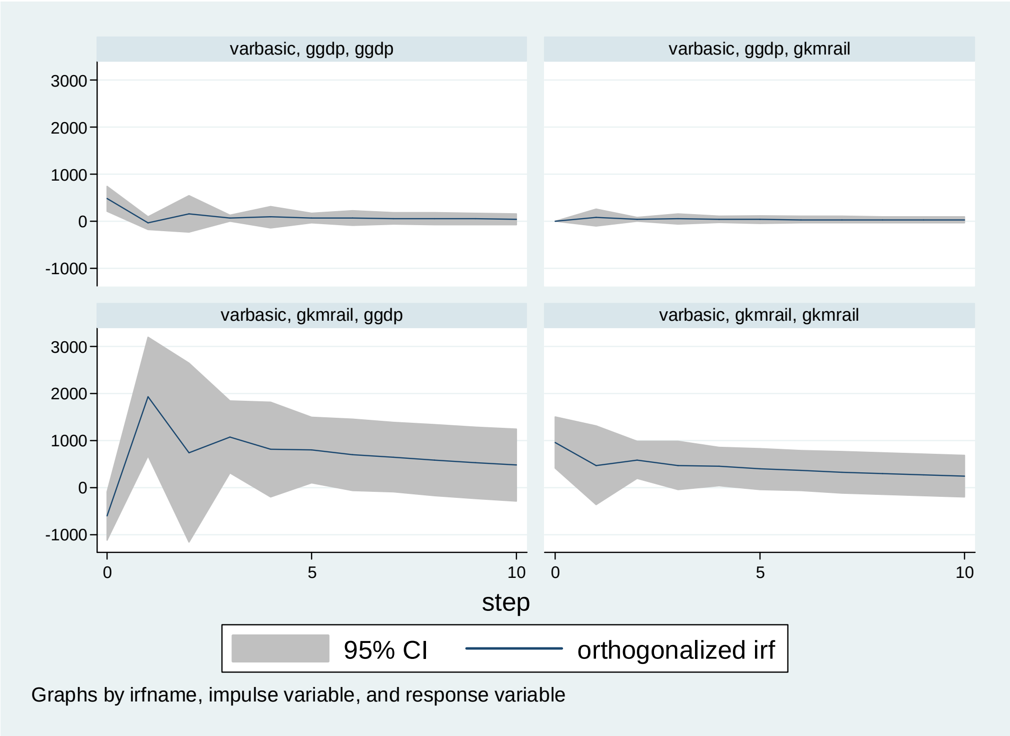
Appendix B

Set of countries:

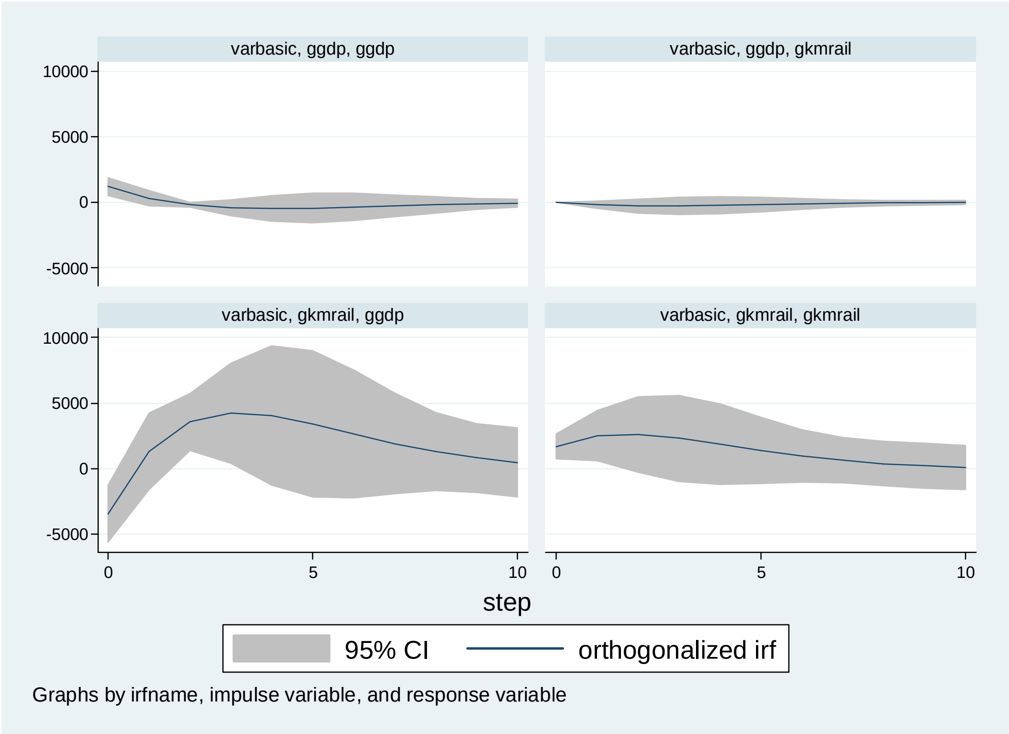
Belgium, Britain, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland.

Appendix C

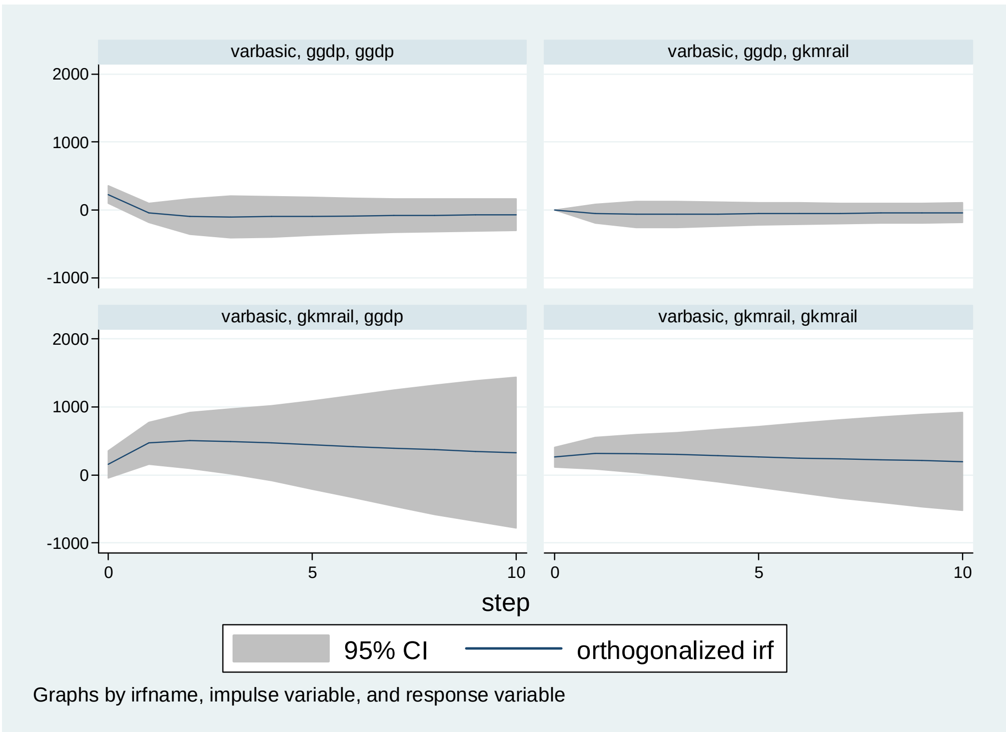
Impulse-Response Functions by country:



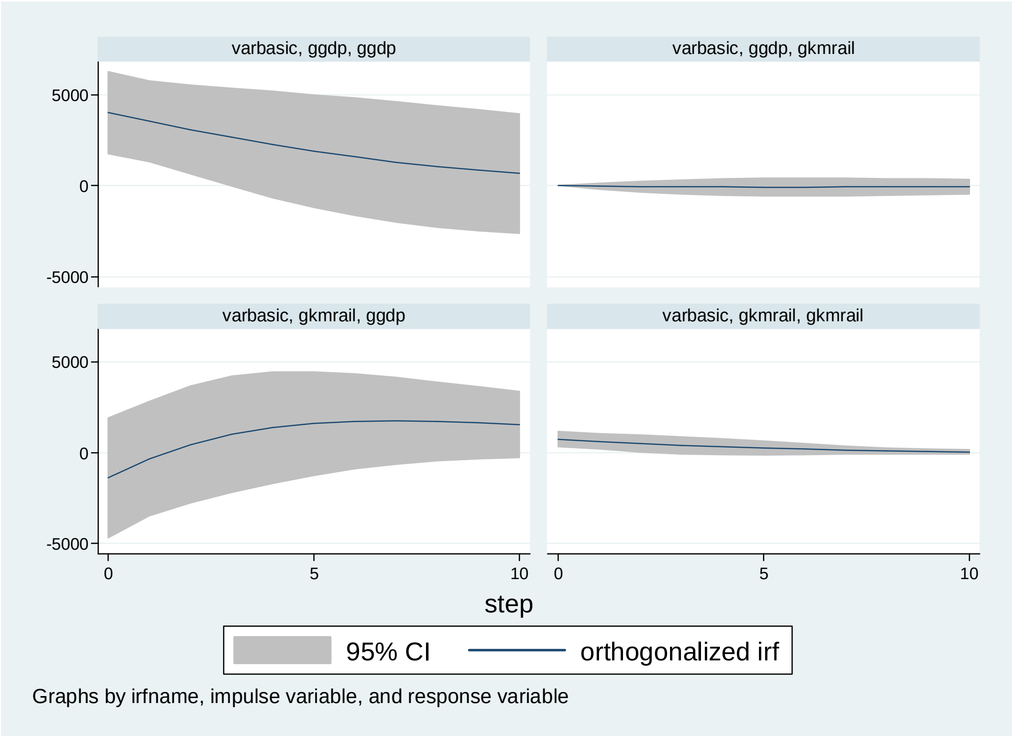
Spain



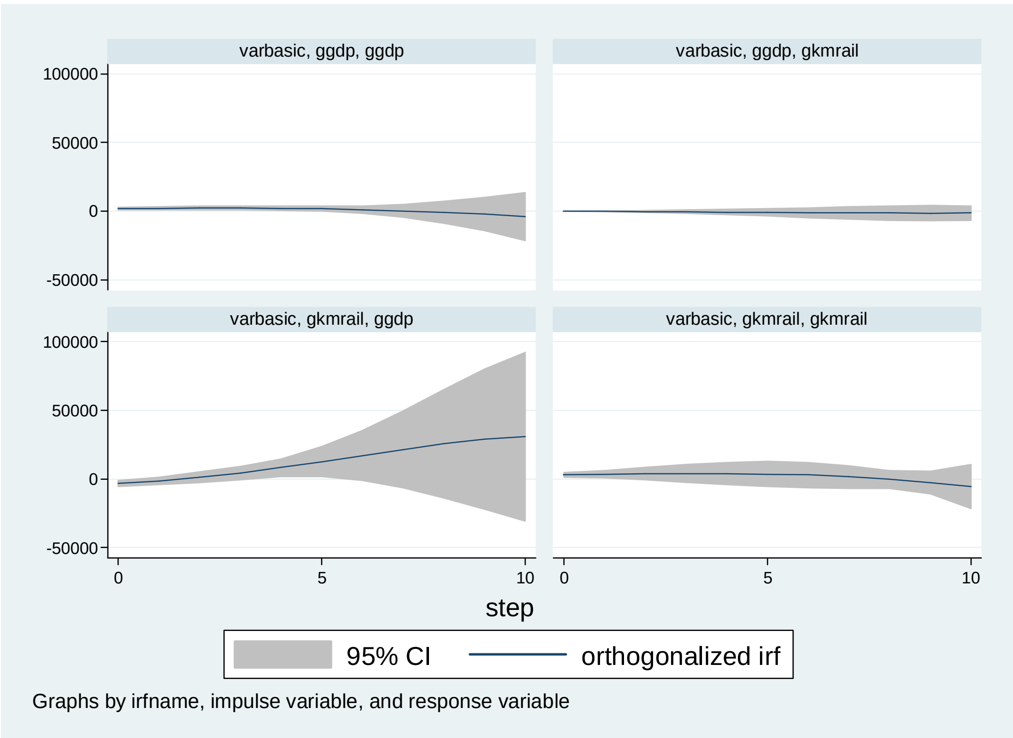
France



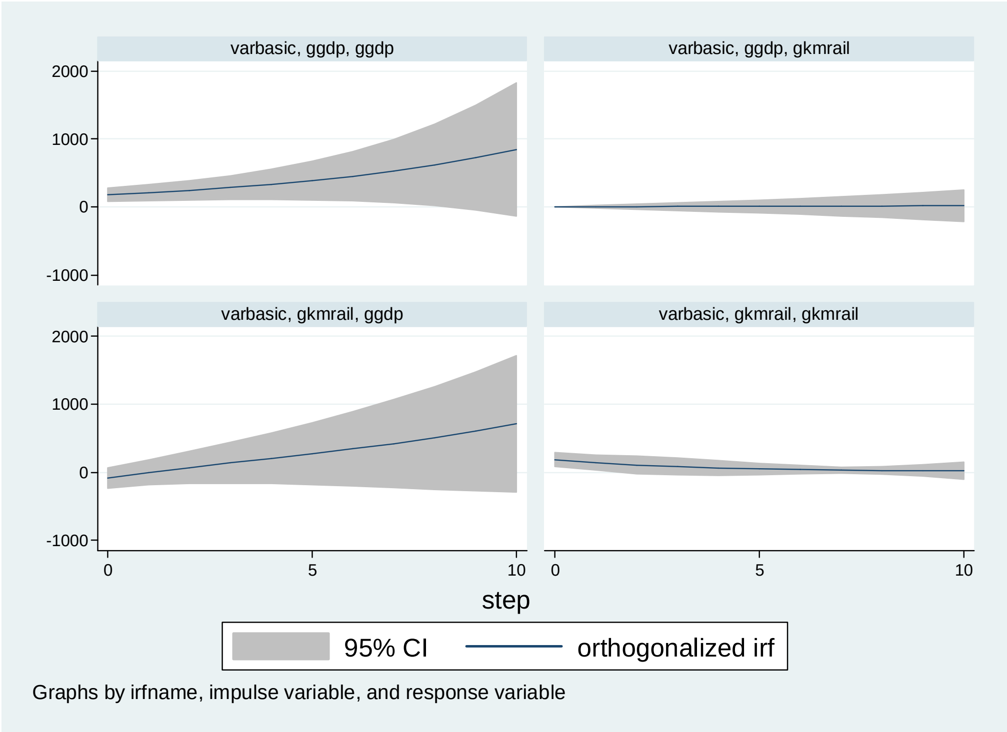
Portugal



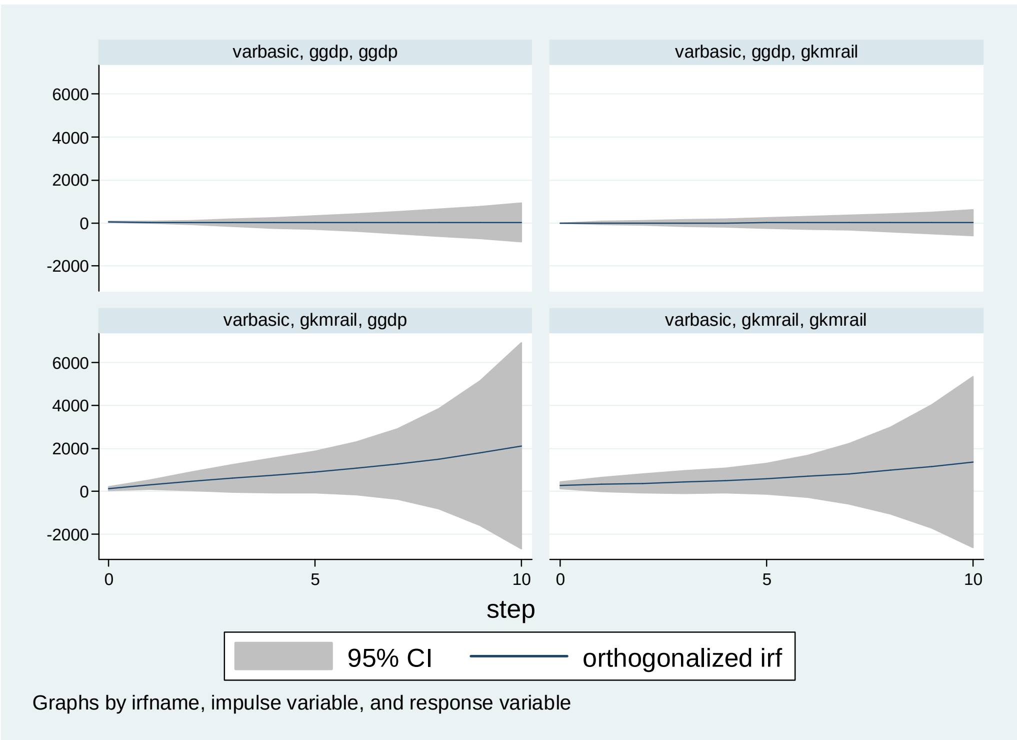
Britain



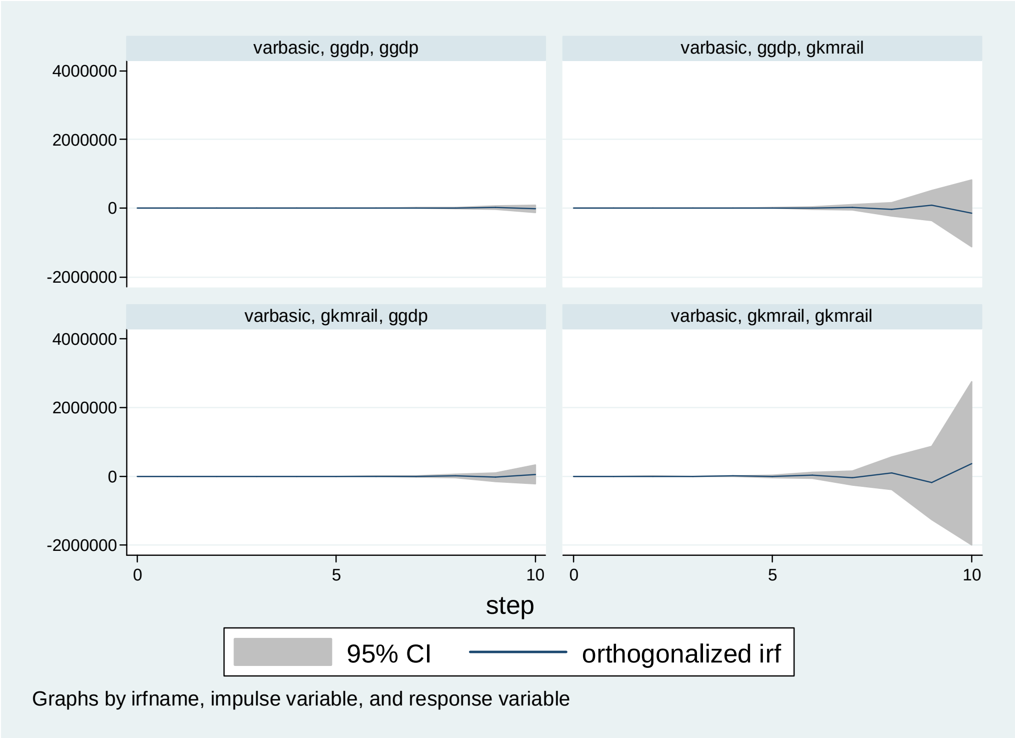
Germany



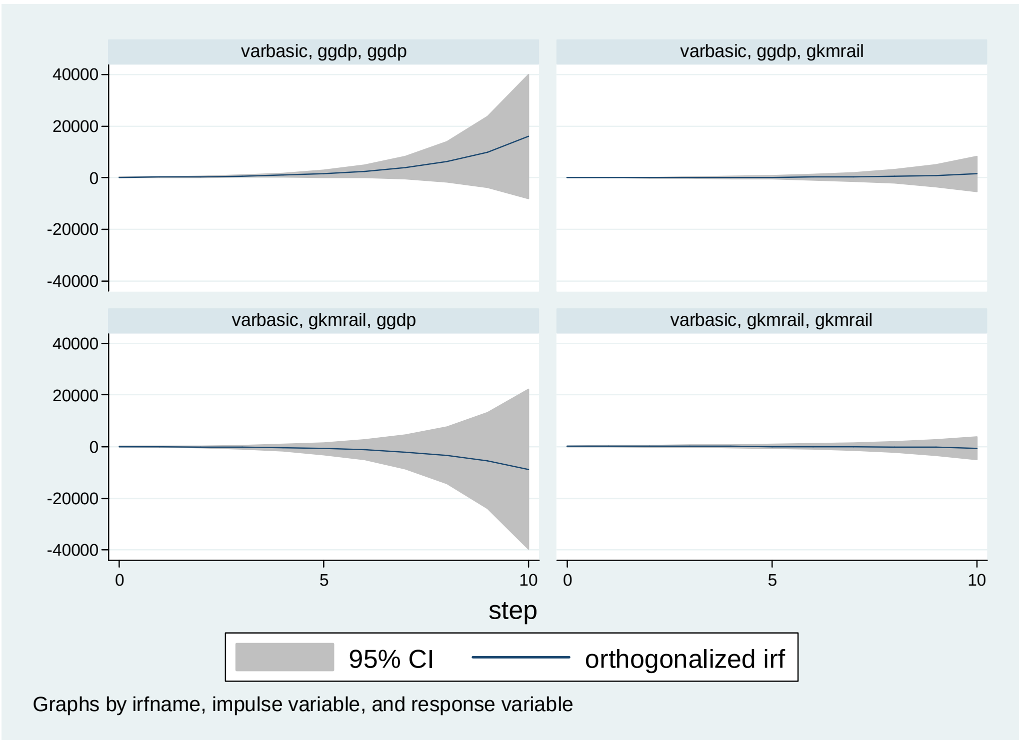
Switzerland



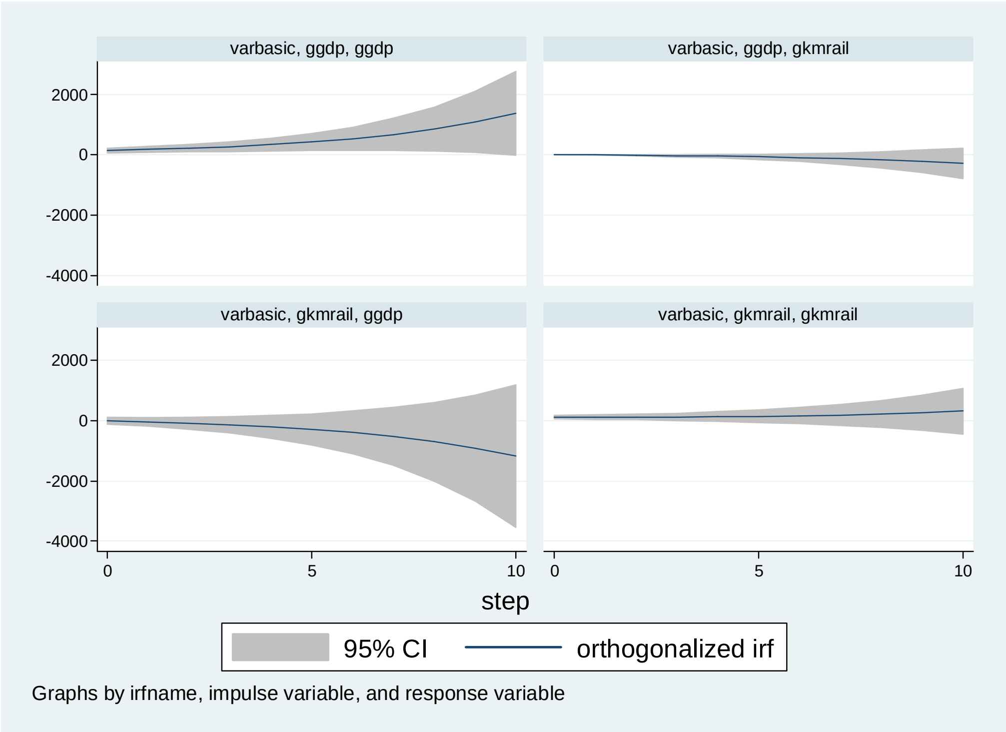
Finland



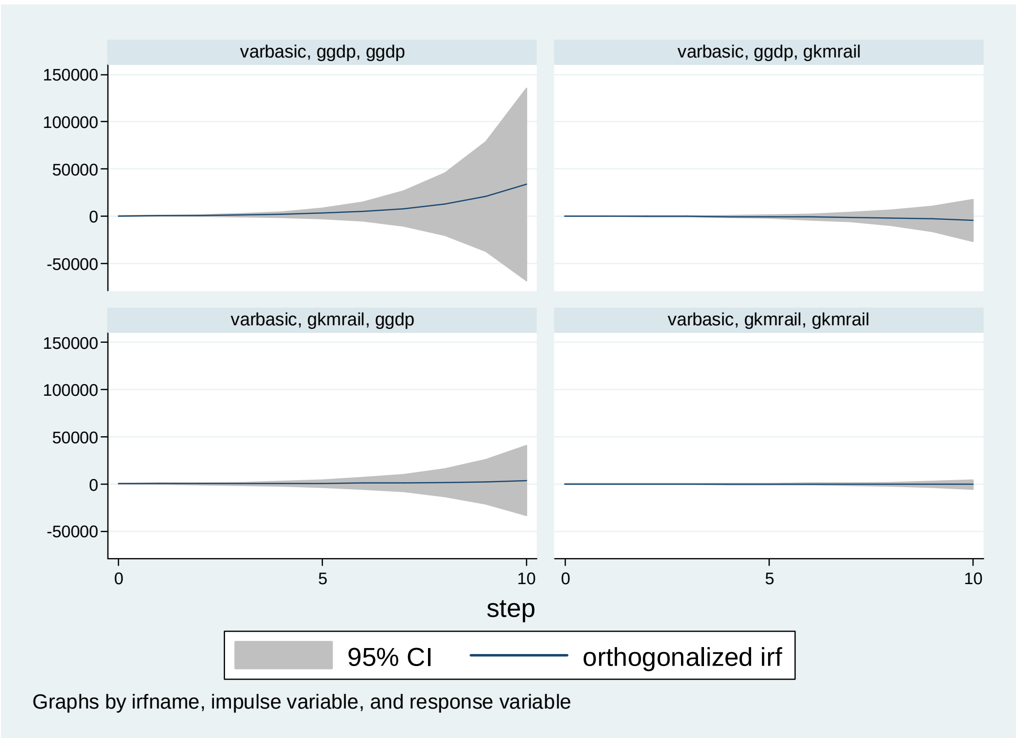
Sweden



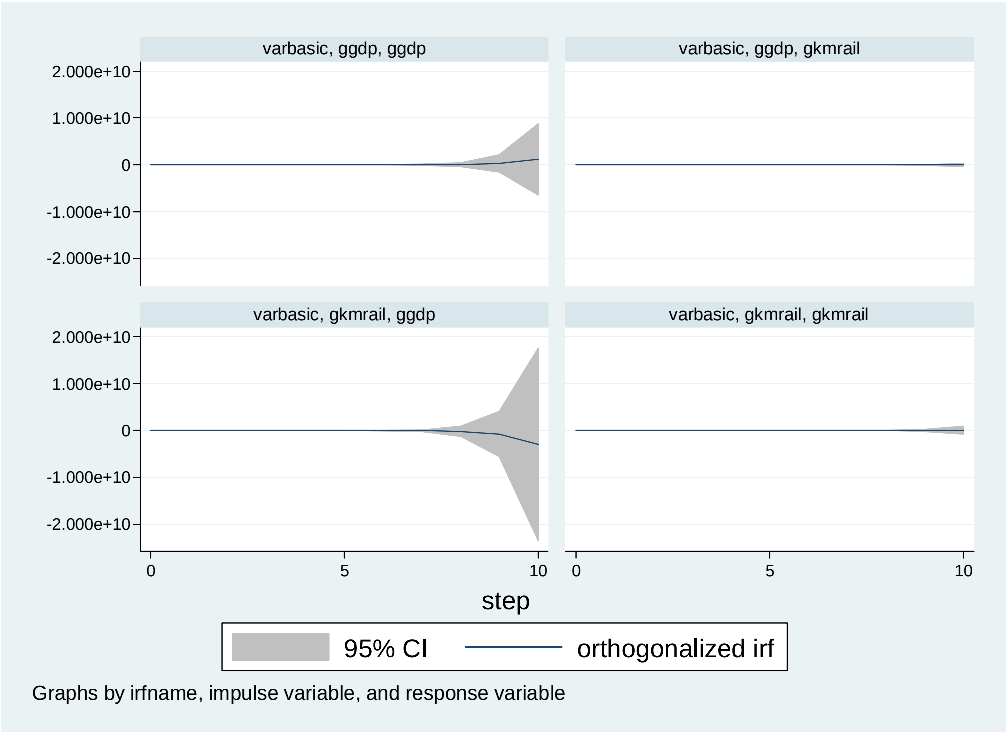
Denmark



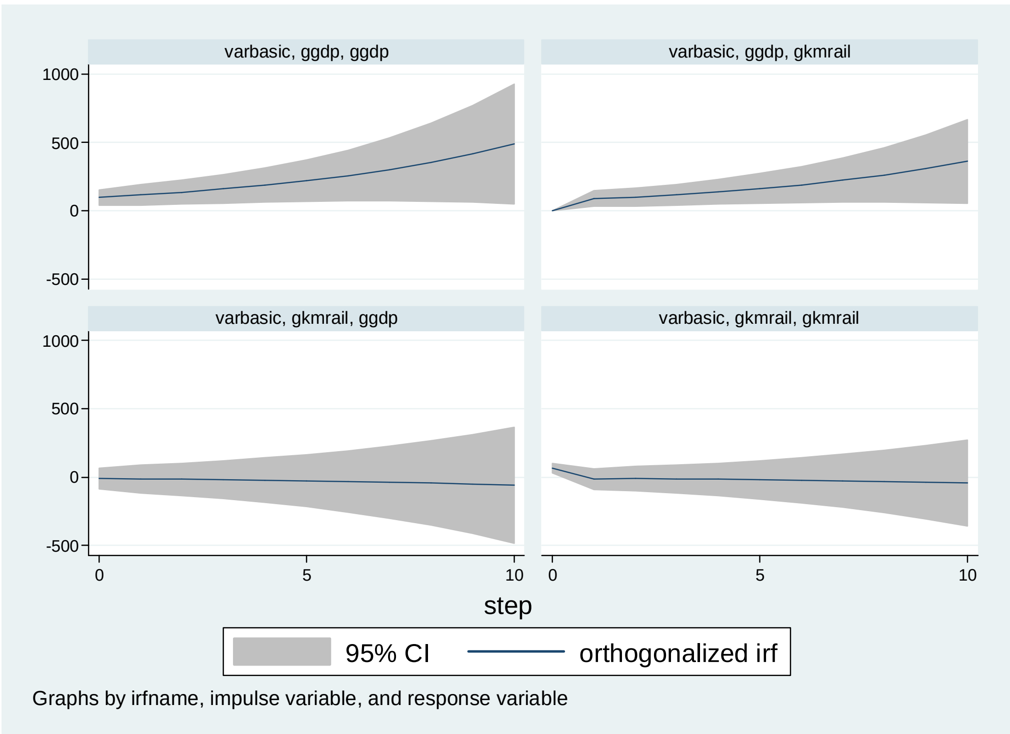
Belgium



Netherlands



Italy



Norway

References

Banerjee A., Duflo E., and Qian N. (2012) On the road: access to transportation infrastructure and economic growth in China. NBER Working Paper, no. 17897.

Breusch, T., Ward M. B., Nguyen H. T. M., and Kompas T. (2011) FEVD: Just IV or just mistaken? *Political Analysis*, 19: 165-169.

Bolt, J. and van Zanden, J. L. (2013) The First Update of the Maddison Project; Re-Estimating Growth Before 1820. Maddison Project Working Paper 4.

Chow, G. C. (1960) Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* 28 (3): 591–605.

Crafts, N., and Mills, T. C. (2009) From Malthus to Solow: How did the Malthusian economy really evolve? *Journal of Macroeconomics*, 31(1): 68-93.

Donaldson D. (2010) Railroads of the Raj: Estimating the Impact of Transport Infrastructure. NBER Working Paper No. 16487, http://www.nber.org/papers/w16487.

Donaldson, D., and Hornbeck, R. (2013) Railroads and American Economic Growth: A "Market Access" Approach, NBER working paper 19213, July 2013.

Ermisch, J. (2008) An Economic History of Bastardy in England and Wales, in Rosenbloom, J. L. (ed.) *Quantitative Economic History: The Good of Counting*, Routledge (Oxon and New York):8-33.

Federico G. and Tena A. (2013) Bilateral Trade Base from MCI: ECO2011-25713 research project: "Globalization and growth in the peripheral countries. Europe and Latin America between 1820 y 1940".

Fogel W.R. (1964) *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins University Press.

Greene, W. H. (2011) Fixed effects vector decomposition: A magical solution to the problem of time invariant variables in fixed effects models? *Political Analysis*, 19:135–46.

Gregory I. N., and Martí-Henneberg J. (2010) The Railroads, Urbanization, and Local Demography in England and Wales, 1825-1911. *Social Science History* 34 (2): 199-228.

Hausman J. A. and Taylor W. E. (1981) Panel Data and Unobservable Individual Effects, *Econometrica*, 49 (6): 1377-1398.

Herranz-Lorcán A. (2006) Railroad impact in backward economies: Spain, 1850–1913. *Journal of Economic History*, 66: 853–881.

Herrendorf B., Schmitz J.A., and Teixeira A. (2009) Transportation and Development: Insights from the U.S., 1840-1860. Federal Reserve Bank of Minneapolis Research Department Staff Report 425, http://minneapolisfed.org/research/sr/sr425.pdf

Johansen, S (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Maddison, A. (2010) Historical Statistics of the Word Economy: 0-2008 AD, University of Groningen, http://www.ggdc.net/maddison/oriindex.htm (last accessed 28/11/2014).

Marti-Henneberg J. (2011) Geographical information systems and the study of history. *Journal of Interdisciplinary History*, 42 (1): 1–13.

Mojica L. and Marti-Henneberg J. (2011) Railways and population distribution: France, Spain and Portugal, 1870–2000. *Journal of Interdisciplinary History*, 42 (1): 15–28.

Nicolini, E. A. (2007) Was Malthus right? A VAR analysis of economic and demographic interactions in pre-industrial England. *European Review of Economic History*, 11(1): 99-121.

Plümper T., and Troeger V. E. (2007) Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects. *Political Analysis*, 15 (2): 124-139.

Prados de la Escosura, L. (2000) International Comparisons of Real Product, 1820-1990: An Alternative Data Set, *Explorations in Economic History* 37(1):1-41.

Sims, C. A. (1980) Macroeconomics and Reality, *Econometrica* 48 (1):1-48.