5-2019

Essays on Convex Weighting for Global Vector Autoregressive Models

Garrison Leach
gleach@smu.edu

ESSAYS ON CONVEX WEIGHTING FOR
GLOBAL VECTOR AUTOREGRESSIVE MODELS

Approved by:

Dr. Thomas Fomby
Professor of Economics

Dr. Klaus Desmet
Professor of Economics

Dr. Thomas Osang
Associate Professor of Economics

Dr. Alexander Chudik
Economic Policy Advisor and Senior Economist
Federal Reserve Bank of Dallas

ESSAYS ON CONVEX WEIGHTING FOR GLOBAL VECTOR AUTOREGRESSIVE MODELS

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Economics

by

Garrison Leach

B.S., Finance & Economics, The University of Alabama
M.A., Applied Economics, The University of Alabama

May 18, 2019

Leach, Garrison

B.S., Finance & Economics, The University of Alabama
M.A., Applied Economics, The University of Alabama

Essay on Convex Weighting for

Global Vector Autoregressive Models

Advisor: Dr. Thomas Fomby

Doctor of Philosophy conferred May 18, 2019

Dissertation completed April 29, 2019

This dissertation focuses on studying the impact that weighting schemes can have on forecasting performance and dynamic analysis in global vector autoregressive (GVAR) models. The first chapter discusses an existing gap in the literature regarding weighting scheme choice and develops a simple, yet powerful method for defining richer spatial linkages in a way that doesn't sacrifice economic context. The new technique called convex weighting, extends the set of available options for defining spatial linkages in models that handle the curse of dimensionality via compression and offers a justifiable approach to alleviating uncertainty. The second and third chapters apply the newly developed convex weighting method to regional and international level models to show that improvements in forecasting performance are possible and that inferences drawn from dynamic analysis can be highly sensitive to the underlying weighting scheme.

Although it has been pointed out as an important issue, the GVAR literature has minimally focused on the issue of weighting schemes and has instead liberally applied the most popular schemes which are forms of distance for regional applications and most prominently import-export share weightings for international applications. While in certain circumstances, there might be sufficient theoretical justification for a given linkage mechanism, a greater issue is the lack of attention payed to robustness across schemes. Even with great care given to robustness, existing approaches are limiting in that they still ultimately require the specification of single linkage mechanisms. The first chapter addresses the single-specification limitation by showing how, with special attention to standardization, an arbitrary number of weighting schemes can be combined convexly in an elementwise fashion. This concept of convex mixing opens up the possibility of optimization which is undertaken in chapters two and three.

Chapter two presents a regional application of the Texas housing market and shows that by optimizing the parameters of the convex weighting, forecasting performance can be improved beyond that of pure non-convex weighted alternative models. The application also shows that impulse response functions can be sensitive to the underlying weighting scheme and illustrates how the use of an optimized convex weighting can allow researchers to be more confident in their findings. Beyond the technical contributions are several core substantive findings. First, it is shown that housing markets in Dallas and Houston respond in a highly similar way to all national level shocks while Austin is effected roughly half as much and San Antonio is largely unaffected. Second, it is shown that the housing markets of Houston and San Antonio experience a significant short run spillover effect equal to roughly 50% the magnitude of a Dallas housing price shock while the Austin market experiences no spillover at all. Overall, chapter two suggests that models utilizing convex weights can perform well on small universe type questions for which distance weights are popular.

The third chapter further extends the convex weighting method by combining it with the existing concept of mixing on variables and models the global economy using one of the most popular global macro datasets. The central focus of the third chapter is on how conclusions, or inference, drawn from impulse response functions can differ across models with different weightings. To accomplish this, an algorithm is developed for categorizing impulse response function significance and is applied exhaustively to show that impulse response inferences between the optimized convex weighted model and a purely trade weighted baseline differ significantly in no less than 30% of cases. Lastly, it is shown that the convex weighted model yields a more coherent picture of the global economy than the trade weighted baseline.

The primary concern of this dissertation is in the area of foreign parameter weighting for GVAR models. A flexible method to address an existing gap in the literature is developed and is applied at both the regional and international levels over which it's performance is shown to be favorable. Researchers looking to utilize the GVAR model will no longer, by the methods proposed in this dissertation, need to concern themselves with the exact specification of the linkage mechanism and will no longer have to sacrifice the economic context that existing alternatives required. This dissertation provides a meaningful contribution to the GVAR literature in the area of weighting scheme choice and lays a plentiful foundation for future applications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Thomas Fomby, for his continued support and guidance. I would also like to thank the members of my committee and Dr. Alexander Chudik for providing valuable commentary and suggestions. I am grateful to Dr. Robert Kalescky and Dr. Li Tan for their assistance in high performance computing and to Mr. Tobin Cobb for providing the financial support that made the timely completion of this research possible. Lastly, I am thankful to my family and friends for their encouragement and patience throughout the duration of this project.

CHAPTER 1

Methodological Overview of Convex Weights for GVAR Models

1.1 Introduction

Macroeconomic policy analysis and forecasting require careful consideration of interdependencies that exist across space. In better efforts to understand such linkages, questions such as how a certain phenomenon in one location might impact other locations have become a popular topic within the fields of research in economics. Understanding both spatial and temporal dynamics requires special attention as modeling many locations in a consistent and cohesive manner can lead to econometric difficulties. Specifically, in handling many locations at once, the number of parameters can quickly exceed the number of observations thus forcing econometricians to make decisions over model sparsity.

One particular method for handling this complication is through compression. Spatial econometricians have been utilizing compression for many years, dating all the way back to Anselin et al. (1980) which first describes how variables across space can be aggregated through weighted averaging techniques. The Global VAR (GVAR) model, which was developed by Pesaran et al. (2004) and further expanded by Dees et al. (2007), takes this technique for spatial regression models and applies it to the foreign variables of each spatial unit in a VAR model, thus resulting in a rich spatio-temporal model that yields the same properties of a traditional VAR model such as impulse responses and forecast error variance decompositions. The GVAR model has gained popularity in recent years and is particularly attractive not only due to the wide range of questions that it is well positioned to address, but also due to it's clever and simple handling of the parameterization problem.

There is, however, a common issue that researchers face in utilizing GVAR models. To make use of the weighted averaging compression technique and to subsequently specify a GVAR model, a weighting scheme must first be decided upon. Because of the vast array of international level questions that can easily be formulated, GVAR models have largely been applied to answer how spillovers and transmission effects occur between countries. Careful review of the literature covered in the survey paper Chudik and Pesaran (2016) reveals that a substantial majority of papers utilizing GVAR models to answer international macro questions (over 85%) have relied upon import and export data for the construction of trade based weights with the remaining portion of papers relying on various financial or distance type weighting schemes. While the use of trade weights may be

perfectly valid in many settings, this chapter contributes to the relatively small amount of research that has focused on the development of alternative weighting schemes.[1].

Among the papers that specifically address weighting schemes is the recent work by Martin and Cuaresma (2017), which in building on concepts in Eickmeier and Ng (2015), brings to light that the choice of weighting schemes in GVAR models has historically been "done in an *ad hoc* fashion" and provides evidence that trade weights may not always be best. Martin and Cuaresma (2017) tests the forecasting performance of nine weighting schemes including several forms of each trade, financial, and distance type weights and finds that, depending upon the horizon and variables of interest, financial and distance weights could perform better than trade weights. Interestingly, it was also found that mixing models of differing weights after estimation, in the sense of an ensemble, resulted in relatively worse performance than non-mixed models. Perhaps because of this unexpected finding, it was recommended that further research be conducted on alternative approaches to mixing weights. This chapter directly addresses the weight mixing question through the development of a new convex weighting technique which extends the set of available options for defining spatial linkages in models that handle the curse of dimensionality via compression and offers a justifiable approach to alleviating uncertainty.

This chapter also extends the work of several other papers that have focused on weighting schemes. Specifically, Feldkircher and Huber (2016), like Martin and Cuaresma (2017), tests nine weighting schemes of the same categories while incorporating weight mixing on variables (i.e., economic variables with trade weights and financial variables with financial weights) and also tests various combinations of models through an ensemble-type procedure but focuses mainly on substantive impulse results rather than forecasting. This chapter introduces "pre-averaging" as an alternative to "post-averaging" (i.e., ensemble-type procedures) with the subsequent chapters presenting real world applications at differing levels and focusing on forecasting and impulse response sensitivity. Eickmeier and Ng (2015) specifically studies impulse sensitivity to weightings but focuses mainly on mixing on variables. Their findings show that mixing on variables improves performance and ultimately suggest that "a GVAR based on more sophisticated and carefully chosen weighting schemes can characterize the data better." Chapters two and three follow this trend by showing that additional complexity improves, rather than hinders, performance in GVAR models. The idea for convex weight mixing comes from Sun et al. (2013) who first proposes the combination of weights before estimation as "the sum of trade flows and foreign exposure positions." However, the main focus of their work is not methodological in nature and thus the emphasis is not on weighting scheme sensitivity or performance. Also, the construction of their combined weights requires that

[1] Roughly 15% of the literature, as identified by the author, appears to specifically address the choice of weighting schemes and it's subsequent impact.

the underlying units of the weighting schemes be the same[2]. The method proposed in this chapter eliminates this requirement and provides a clarified and straightforward way to combine weighting schemes. Lastly, the survey paper Chudik and Pesaran (2016) in review of the existing GVAR literature to date states that "selection of optimal weights could be an important issue." To this point, this dissertation lays out a foundation for how to arrive at an optimum without sacrificing economic context[3].

The remainder of this chapter is organized as follows: Section 1.2 formally presents the GVAR model and the generalized impulse response functions that are used in dynamic analysis. Section 1.3 shows how to compute weighting matrices and formally presents convex weights. Lastly, Section 1.4 provides some context and prefaces the following chapters.

1.2 The GVAR Approach

The GVAR approach was originally developed by Pesaran et al. (2004) to provide a global modeling framework capable of generating forecasts while allowing for interdependencies between a set of regions or countries. The framework has since been furthered by works such as Dees et al. (2007) who advanced the model in a number of directions and by Chudik and Pesaran (2013) who contributed an advance with the inclusion of a dominant unit within the GVAR. This chapter advances the framework by providing a relatively straightforward way to construct linkages between spatial units for which any realistic linkage may be less obvious or obtainable due to data availability or some other reason. This is particularly the case when investigating a universe that is smaller than the global scale, such as an intranational or intrastate scale. Critically, analysis of such smaller universes has typically relied on spatial methods, covered extensively in Anselin (1988), that don't focus on dynamics and rely heavily on the distance class of weighting schemes which may in some cases seem unreasonable[4]. For example, Vansteenkiste (2007) uses pure distance weights to link states together in an investigation of regional housing prices. In the case of states linked via pure distance though, California and New York are minimally related simply due to being far away from each other; a consequence which may or may not be reasonable. Similarly, contiguity weighting schemes by construction result in far fewer linkages and thus might for the same reasons be unbelievable. Hence, a more tractable weighting scheme beyond a single choice

[2] For clarity, it should be noted that other papers, namely Chudik and Fratzscher (2011) and Bussière et al. (2011), have utilized multiple weighting schemes without this requirement. However, these papers still rely on "ad hoc" specifications of coefficients.

[3] Gross (2018) also develops a method for arriving at an optimal GVAR weighting by estimating the weights endogenously as part of the system. This endogenous approach is novel and may be well suited for certain applications but in those where it might be of interest to understand how the cross-sectional units are linked, another method is necessary since a consequence of the fully endogenous estimation is that any economic context is lost.

[4] Elhorst et al. (2018) provides a thorough overview of the intersection of spatial econometric methods and the GVAR model.

within the distance class is presented here. First, the GVAR model is presented and second, details regarding the weighting scheme are presented.

When dealing with smaller universes, as does the application in chapter 2, it is of interest to understand how units respond not only to cross-sectional shocks but also to common factor shocks from an above universe. In the case of a city, national level variables, such as the interest rate or the oil price, are shared among all cross-sectional units and represent variation from a higher universe. To allow for such higher level interactions, a GVAR specification that includes a dominant unit containing all common factors or global variables is presented. Because the model in chapter 2 uses a small number of cities as cross-sectional units, the dominant unit model, as is discussed in Chudik and Pesaran (2013), is presented without feedback effects because a small number of cities in a single state could reasonably be assumed not to substantially impact national (interest rates) or truly global (oil price) variables.

1.2.1 A Generic Dominant Unit Global VAR Model

This section formally presents the GVAR model to be used in the applications that follows.[5] The application in chapter 2 mirrors the model that is presented in this chapter and the application in chapter 3 mirrors the model in Dees et al. (2007) which excludes the dominant unit. Beginning with the dominant unit, denote it's variables by the $m_\omega \times 1$ vector of observables ω_t and consider the following VAR(p_ω) specification

$$\omega_t = \mu_t + \mu_1 t + \phi_1 \omega_{t-1} + \ldots + \phi_{p_\omega} \omega_{t-p_\omega} + \eta_t \tag{1.1}$$

where p_ω is the lag order of the vector of observables ω and is selected by information criterion[6]. Such an inclusion . In the presence of I(1) variables, eq. (1.1) can be rewritten in error correction representation under case IV (unrestricted intercept, restricted trend) as

$$\Delta\omega_t = c - \alpha_\omega \beta'_\omega [\omega_{t-1} - \kappa(t-1)] + \sum_{j=1}^{p_\omega - 1} \Gamma_j \Delta\omega_{t-j} + \eta_t \tag{1.2}$$

where α_ω and β_ω are $m_\omega \times r_\omega$ vectors, and r_ω denotes the number of cointegrating relations. Denote the $r_\omega \times 1$ vector of error correction terms by $\xi_{\omega,t-1} = \beta'_\omega[\omega_{t-1} - \kappa(t-1)]$, and it's estimate by

[5] The Appendix of the User Guide for the GVAR Toolbox 2.0 by L. Vanessa Smith & Alessandro Galesi provides an excellent reference for understanding GVAR models. Readers are highly encouraged to explore the GVAR Toolbox as it represents a powerful contribution to the GVAR community.

[6] For clarification, eq. (1.1) can be augmented by lagged granular cross-sectional averages of non-dominant units to account for possible unobserved common factors. See Chudik and Pesaran (2013) for details.

$\hat{\xi} = \hat{\beta}'_{\omega}[\omega_{t-1} - \hat{\kappa}(t-1)]$. The remaining parameters of eq. (1.2) are estimated by OLS applied to

$$\Delta\omega_t = c + \delta\hat{\xi}_{\omega,t-1} + \sum_{j=1}^{p_\omega - 1} \Gamma_j \Delta\omega_{t-j} + \eta_t \tag{1.3}$$

where $\hat{\xi}_{\omega,t-1}$ is taken as given from the first stage estimation. Following estimation, eq. (1.3) can be re-expressed as eq. (1.1) which is then used in the second stage for establishing the dynamic properties of the global model.

In addition to the dominant unit model that captures global variables, individual spatial units are modeled under the following VARX*(p_i, q_i) structure for the i^{th} spatial unit

$$\begin{aligned} x_{it} = a_{i0} + a_{i1}t + \phi_{i1}x_{i,t-1} + \ldots + \phi_{ip_i}x_{i,t-p_i} + \Lambda_{i0}x^*_{it} + \Lambda_{i1}x^*_{i,t-1} + \ldots \\ + \Lambda_{iq_i}x^*_{i,t-q_i} + \Psi_{i0}\omega_t + \Psi_{i1}\omega_{t-1} + \ldots + \Psi_{iq_i}\omega_{t-q_i} + u_{it}, \end{aligned} \tag{1.4}$$

for $i = 0,1,\ldots,N$. where x_{it} is a k-dimensional column vector of domestic variables for cross-sectional unit i in period t, a_{i0} is a vector of constants, $a_{i1}t$ is a linear trend, x^*_{it} are k^*-dimensional column vectors of weighted foreign variables (assumed weakly exogenous), ω_{it} is the m_ω-dimensional vector of global variables, and u_{it} is a k-dimensional vector of serially uncorrelated error terms. ϕ_{it}, Λ_{it}, and Ψ_{it} are corresponding coefficient matrices[7]. Foreign variables x^*_{it} in the GVAR model are highly similar to spatially lagged variables in a spatial regression model. That is to say, assuming $k^* = k_i = k$ for all i, they are constructed as weighted averages of other spatial units

$$x^*_{it} = \sum_{j=0}^{N} w_{ij}x_{jt}, \quad w_{ii} = 0 \tag{1.5}$$

with w_{ij}, $j = 0,1,\ldots,N$ being a set of weights such that $\sum_{j=0}^{N} w_{ij} = 1$[8].

Defining in terms of $z_{it} = (x_{it}\ x^*_{it})'$, a vector that stacks domestic and foreign variables, and assuming the lag orders on domestic and foreign (global) variables $p_i = q_i$ are equal for expositional purpose, we have

$$\begin{aligned} G_{i0}z_{it} = a_{i0} + a_{i1}t + G_{i1}z_{i,t-1} + \ldots + G_{ip_i}z_{i,t-p_i} \\ + \Psi_{i0}\omega_t + \Psi_{i1}\omega_{t-1} + \ldots + \Psi_{iq_i}\omega_{t-q_i} + u_{it}, \end{aligned} \tag{1.6}$$

with $G_{i0} = (I_{k_i}, -\Lambda_{i0})$ and $G_{ij} = (\phi_{ij},\Lambda_{ij})$ for $j = 1,\ldots,p_i$. Using the identity $z_{it} = W_i x_t$ where W_i are link matrices containing bilateral exposures between the spatial units at hand and x_t is a $K \times 1$ vector including all non-global endogenous variables of the system, eq. (1.6) can be written

[7] In practice, the macroeconomic variables in the VARX* models typically have unit roots and it is possible that they have cointegrating relationships among themselves. Due to this, it is plausible to estimate equation eq. (1.4) in error correction form (VECMX*). See Section 33.3 in Pesaran (2015).

[8] A complete description on weighting matrix construction is provided in Section 1.3

as

$$G_{i0}W_i x_t = a_{i0} + a_{i1}t + G_{i1}W_i x_{t-1} + \ldots + G_{ip_i}W_i x_{t-p_i} + \Psi_{i0}\omega_t + \Psi_{i1}\omega_{t-1} + \ldots + \Psi_{iq_i}\omega_{t-q_i} + u_{it}.$$

These individual models are then stacked to yield the model for x_t given by

$$G_0 x_t = a_0 + a_1 t + G_1 x_{t-1} + \ldots + G_p x_{t-p} + \Psi_0\omega_t + \Psi_1\omega_{t-1} + \ldots + \Psi_q\omega_{t-q} + u_t \tag{1.7}$$

where both the contemporaneous and lagged values of ω_t now appear on the right hand side of eq. (1.7) with $p = max(p_i)$ and $q = max(q_i)$ and

$$G_0 = \begin{pmatrix} G_{00}W_0 \\ G_{10}W_1 \\ \vdots \\ G_{N0}W_N \end{pmatrix}, \; G_j = \begin{pmatrix} G_{0j}W_0 \\ G_{1j}W_1 \\ \vdots \\ G_{Nj}W_N \end{pmatrix}, \; j = 1, \ldots, p,$$

$$a_0 = \begin{pmatrix} a_{00} \\ a_{10} \\ \vdots \\ a_{N0} \end{pmatrix}, \; a_1 = \begin{pmatrix} a_{01} \\ a_{11} \\ \vdots \\ a_{N1} \end{pmatrix}, \; u_t = \begin{pmatrix} u_{0t} \\ u_{1t} \\ \vdots \\ u_{Nt} \end{pmatrix}.$$

The η_t and u_t error terms are assumed to be uncorrelated.

Defining the $(k + m_\omega) \times 1$ vector $y_t = (x_t', \omega_t')'$, and eq. (1.7) for $p = \tilde{p} = \tilde{q} = q$ can be written as

$$H_0 y_t = h_0 + h_1 t + H_1 y_{t-1} + \ldots + H_p y_{t-p} + \zeta_t \tag{1.8}$$

where

$$H_0 = \begin{bmatrix} G_0 & -\Psi_0 \\ 0_{m_\omega \times k} & I_{m_\omega} \end{bmatrix}, h_0 = \begin{bmatrix} a_0 \\ \mu_0 \end{bmatrix}, h_1 = \begin{bmatrix} a_1 \\ \mu_1 \end{bmatrix},$$
$$H_j = \begin{bmatrix} G_j & -\Psi_j \\ \Lambda_j W_j & \phi_j \end{bmatrix}, j = 1, \ldots, p, \zeta_t = \begin{bmatrix} u_t \\ \eta_t \end{bmatrix},$$

or

$$y_t = c_0 + c_1 t + C_1 y_{t-1} + \ldots + C_p y_{t-p} + e_t \tag{1.9}$$

with

$$c_j = H_0^{-1}h_j,\ j = 0,1;\ C_j = H_0^{-1}H_j,\ j = 1,\ldots,p,\ e_t = H_0^{-1}\zeta_t.$$

To specify and estimate the spatial unit-specific models given by eq. (1.4), standard procedure suggested in Pesaran et al. (2004) is followed whereby x_t^* and ω_t are combined and treated jointly as weakly exogenous. The models are then estimated using the Pesaran et al. (2000) I(1) modified version of Johansen (1991) reduced rank regression techniques now for VECMX* models where Johansens trace and maximal eigenvalue statistics are used to determine the rank order of each spatial unit VARX* model. Lag orders in chapter 2 are determined by SBC with an assumed maximum lag order $p = q = 12$ while both AIC and SBC are tested in chapter 3 with maximum order matching that of Dees et al. (2007).

1.2.2 Impulse Responses Analysis with GVARs

To analyze shocks in a way that combats ordering complications, which without theoretical guidance is a large problem particularly in GVAR models due to having to decide over the spatial dimension, the Generalized Impulse Response Function (GIRF) is used. The GIRF approach which was proposed in Koop et al. (1996) and developed for vector error-correcting models in Pesaran et al. (2000) reports how shocks to one variable affect the other variables of the system, on impact and over time, regardless of the source of the change, but taking into account the possibility that the error terms of the GVAR are contemporaneously correlated. Formally,

$$\begin{aligned}\mathcal{GIRF}_{\zeta_j}(h) &= \mathrm{E}(y_{t+h}|\zeta_{jt} = \sqrt{\sigma_{jj}},\mathcal{I}_{t-1}) - \mathrm{E}(y_{t+h}|\mathcal{I}_{t-1}) \\ &= \frac{\mathrm{R}_h H_0^{-1}\Sigma e_j}{\sqrt{e_j'\Sigma e_j}}\end{aligned} \tag{1.10}$$

for $j = 1,2,\ldots,k+m_\omega$ and $h = 0,1,2,\ldots$ where $\sigma_{jj} = \mathrm{E}(\zeta_{jt}^2)$ is the size of the shock which is set to one standard deviation of ζ_{jt}, $\mathcal{I} = \{y_t, y_{t-1}, \ldots\}$ is the information set consisting of all available information at time t, Σ is the sample covariance matrix of the error term in the global model, e_j is a $(k+m_\omega)\times 1$ selection vector of weighted non-zero values only for elements associated with the variable to be shocked for the cross sectional units involved, and the matrices R_h are obainted recursively as

$$\mathrm{R}_h = \sum_{j=1}^{p} C_j \mathrm{R}_{h-j} \text{ with } \mathrm{R}_0 = \mathrm{I}_{k+\omega} \text{ and } R_j = 0 \text{ for } j < 0.$$

Since consideration of the exact nature of shocks lies beyond the scope of the applications in this dissertation, the GIRF is a well positioned alternative to the more traditional orthogonolized impulse

responses analysis originally proposed by Sims (1980). Even without theoretically motivated *a priori* beliefs over the ordering, the GIRF approach can provide useful information about how variables respond across space. In the results of the following applications, figures display bootstrap estimates of the GIRFs and their associated 90% confidence bounds[9].

1.3 Convex Weights

One of the core factors behind the GVAR framework is how it approaches and handles the curse of dimensionality by compressing foreign variables through weighted averaging techniques via eq. (1.5). However, this only reframes the underlying issue into a selection problem. To execute the compression, a weighting scheme must be constructed that will ultimately define the transmission between the cross-sectional units of the model. Traditionally, when conducting analysis at the international level with countries as cross-sectional units, researchers have used bilateral import-export balances to construct trade weights. A range of other weighting schemes have also been used though, including but not limited to portfolio investment, foreign direct investment, banking claims, trade costs, and geodesic distance. Martin and Cuaresma (2017) study comprehensively in the international setting the forecasting performances of each of these weighting schemes and find generally mixed results. Their study also tested model averaging techniques, of which various ensemble models were found to perform surprisingly poorly.

An interesting problem presents itself when the scope of the analysis is less than international, however, specifically when the above listed weights may not be available or even exist. In these cases, such as in the analysis of states within a country or cities within a state, a different linkage mechanism must be chosen. At this sub-international level, many researches have chosen weighting schemes of the distance class[10]. Distance weights have been very popular in spatial econometrics since the development of the spatial regression in Anselin et al. (1980) and have been used to analyze causal effects of variation in one location on variation in another. The core schemes of the distance class are pure distance, threshold distance, contiguity, or a combination thereof where the same principle generally governs how much weight is given to units; the further away a unit is, the less weight is attributed to it. Interestingly, Martin and Cuaresma (2017) report that some of the best performing forecasting models are those utilizing either distance weights or financial linkages and suggest that further research to assess optimal approaches to combining data on weights for GVAR models should be undertaken. To that end, this section begins to address their suggestion in presenting a way to combine information from multiple weighting schemes prior to model averaging.

The weighting scheme procedure proposed in this section is tractable in that it provides a way

[9] All bootstraps throughout this dissertation are conducted with 1000 replications.

[10] The term 'class' is used informally in a computer programming sense because the distance class contains different templates for creating objects, which in this case are weighting matrices.

to construct bilateral pairs from non-bilateral data by exploiting the original purpose of foreign variable compression. This could be particularly useful when pure distance weights might seem unsatisfactory, as could be the case for certain financial variables whose relation across space is digital, or in the case where other bilateral measures simply aren't available. The final weighting scheme proposed here is essentially an elementwise convex combination of weights that undergo a series of standardizations. In the application in chapter 2, geodesic distance and output are used at the metro level and thus for simplicity, the formal presentation of convex weights is restricted to those variables. Note however, that it is trivial to expand on the number of matrices included in the convex combination using the framework presented below.

To begin, define the arc-distance between two spatial units i and j as follows

$$d_{ij} = arc\ distance.$$

Then, so as to conform to Tobler's first law which implies a *distance decay* effect, a continuous parameterized function of distance itself is applied

$$h_{ij} = f(d_{ij},\gamma),$$

with $\partial h_{ij}/\partial d_{ij} < 0$ and γ as a choice parameter that controls the strength of distance decay. In other words, the function ensures that less weight is attributed as distance grows. Throughout the applications in the chapters that follow, the inverse function $h_{ij} = 1/d_{ij}^{\alpha}$ is used with $\alpha = -1$. In practice, α is typically set to a value of -1 to represent standard inverse weights or to a value of -2 to represent gravity weights. Conventionally, the diagonal elements of the spatial weights are set to zero and are not computed (i.e., plugging in a value of $d_{ii} = 0$ would yield division by zero for inverse distance weights).

The inverse distances h_{ij} are row standardized to yield relative weights for each bilateral pair

$$h_{ij(s)} = h_{ij} / \sum_{j} h_{ij}.$$

The standardized weights are then arranged into a weighting matrix as follows

$$\mathbf{H} = \begin{bmatrix} h_{11(s)} & \dots & h_{1n(s)} \\ \vdots & \ddots & \vdots \\ h_{n1(s)} & \dots & h_{nn(s)} \end{bmatrix} = (h_{ij(s)}) \in \mathbb{R}^{\mathrm{N}\times\mathrm{N}}.$$

The H matrix is then transposed so as to conform to the GVAR toolbox which requires column standardized weights rather than row standardized. Define the column standardized bilateral matrix as:

$$B = \mathbf{H}'$$

Up to this point, nothing beyond the standard way for calculating distance weights has been presented. Importantly though, any bilaterally structured data can be used in place or in addition to distance so long as it's B matrix is hollow and column standardized. Following this however, we depart by incorporating non-bilateral vectors, denoted as "global vectors", since their measure is the same across all spatial units. In what follows, we allow for c bilateral matrices and k global vectors to be incorporated. First define the k raw valued global vectors as follows

$$\begin{aligned} G_1^r &= [g_1^1 \;\; g_2^1 \;\; \dots \;\; g_n^1]' \\ &\vdots \\ G_k^r &= [g_1^k \;\; g_2^k \;\; \dots \;\; g_n^k]' \end{aligned}$$

and standardize each in the usual way

$$g_{j(s)}^i = g_j^i / \sum_i g_j^i \;\; \text{for all} \;\; \begin{matrix} i = 1, \dots, k \\ j = 1, \dots, n \end{matrix}$$

to get

$$\begin{aligned} G_1^s &= [g_{1(s)}^1 \;\; g_{2(s)}^1 \;\; \dots \;\; g_{n(s)}^1]' \\ &\vdots \\ G_k^s &= [g_{1(s)}^k \;\; g_{2(s)}^k \;\; \dots \;\; g_{n(s)}^k]' \end{aligned}$$

Since the purpose of the weighting scheme in the GVAR model is to construct foreign variables that account for activity occurring outside of a given domestic unit, the value in each domestic unit can be dropped so that the global vectors now only contain foreign contributions of the otherwise complete total. To do this efficiently, $N \times N$ global matrices are formed for each global vector by duplicating the vector N number of times

$$\begin{aligned} G_1^m &= [G_1^s \;\; G_1^s \;\; \dots \;\; G_1^s] \\ &\vdots \\ G_k^m &= [G_k^s \;\; G_k^s \;\; \dots \;\; G_k^s]. \end{aligned}$$

To drop the domestic contribution, the diagonal elements of the G^m matrices are set to zero and together with the other non-zero elements, global hollow matrices can be denoted as

$$G_k^{m(h)} = \begin{bmatrix} 0 & g_{1,2}^{k,m} & \cdots & g_{1,n}^{k,m} \\ g_{2,1}^{k,m} & 0 & \ddots & \vdots \\ \vdots & \ddots & 0 & g_{n-1,n}^{k,m} \\ g_{n,1}^{k,m} & \cdots & g_{n,n-1}^{k,m} & 0 \end{bmatrix}.$$

Each global hollow matrix is then column standardized in usual way to yield finalized global matrices of which the k^{th} can be written as

$$G_k^{m(f)} = \begin{bmatrix} g_{(k)11}^{m(f)} & \cdots & g_{(k)1n}^{m(f)} \\ \vdots & \ddots & \vdots \\ g_{(k)n1}^{m(f)} & \cdots & g_{(k)nn}^{m(f)} \end{bmatrix}$$

with

$$g_{(k)ij}^{m(f)} = \frac{g_{ij}^{k,m}}{\sum_{j=1}^{n} g_{ij}^{k,m}}$$

.

With bilateral and global matrices that are both standardized, it is now straightforward to construct convex combinations. Begin by defining the bilateral weights θ_ρ and global weights α_δ such that the following conditions are met:

$$\theta_\rho \in (0,1) \text{ for each } \rho = 1,\ldots,c$$
$$\alpha_\delta \in (0,1) \text{ for each } \delta = 1,\ldots,k$$
$$\sum_{\rho=1}^{c} \theta_\rho + \sum_{\delta=1}^{k} \alpha_\delta = 1$$

The share of weight given to bilaterally defined matrices is represented by θ_ρ while α_δ represents the shares given to global type weights. The convex combination weighting matrix is then constructed by computing elementwise convex combinations of the bilateral B_ρ matrices and global $G_\delta^{m(f)}$ matrices in the following way:

$$\mathbf{W}_{ij} = \sum_{\rho=1}^{c} \theta_\rho B_{\rho,ij} + \sum_{\delta=1}^{k} \alpha_\delta G_{\delta,ij}^{m(f)} \text{ , } for \begin{array}{l} i = 1,\ldots,n \\ j = 1,\ldots,n \end{array} \tag{1.11}$$

Furthermore, since the bilateral matrices B and the global matrices G are of equal dimension, they

can be combined into a single representative unit so that the final form of the convex weights is as follows:

$$\mathbf{W}_{ij} = \sum_{\rho=1}^{|M|} \theta_\rho M_{\rho,ij} \ , for \ \begin{matrix} i = 1,\ldots,n \\ j = 1,\ldots,n \end{matrix} \tag{1.12}$$

where M is a weighting matrix, θ are the convexity parameters, and $|M|$ represents the cardinality of the set of weighting matrices. As is mentioned earlier, the application in chapter 2 uses one bilateral matrix (distance) and one global vector matrix (output) and thus the weightings are notated in the spirit of Equation (1.11). The application in chapter 3 utilizes more matrices and thus the results are notated in accordance with Equation (1.12). Without loss of generality, the convex weightings presented above can be applied in a time varying way similarly to how trade weights were applied in Dees et al. (2007).

1.4 Choice Parameters

Choice parameters up to the econometrician are always typically a matter of interest. Chosen improperly, the results of the model could quickly become unbelievable. Chosen in an ad hoc fashion, the model might lose it's proper footing in the real and instead move too far towards the abstract, thus losing interpretability. In the case of the standard GVAR model, there are several choice parameters that the econometrician must consider. First and foremost are the variables that go into the model, of which such decisions are beyond the scope of this work. Secondly, are decisions over the technical workings of the model such as the rank of the individual VARX* models or the lag orders for which well developed techniques exist to aid in the decisions over these factors. The primary concern of this dissertation is in the area of foreign parameter weighting, for which there isn't yet any particularly well established techniques to help govern our decision making. Decisions over the choice of weighting schemes have historically been made, as Martin and Cuaresma (2017) state, in an "ad hoc" fashion particularly with the substantial prevalence of trade weights. Critically, the issue lies not with the use of trade weights specifically but rather with the requirement of exact specification. The fact that the econometrician must choose a weighting scheme means that there will always be room for discussion over his method of choosing. Thus, the aim of the applications that follow is to relax the requirement of exact specification through an optimization procedure over the convex weighting scheme developed in this chapter. By replacing choice with optimization, the magnitude of any uncertainty over whether the weighting scheme was adequate should be minimized. To properly put any further discussion of weighting optimality into context, Chapter 2 develops a small scale GVAR model for illustrative purposes and Chapter 3 develops a full scale model with the aim of showing a real world empirical example of inference sensitivity to weighting.

CHAPTER 2

Regional Application on the Texas Housing Market

2.1 Introduction

This chapter presents the first application of convex weightings in a GVAR model. Of particular interest in this chapter is forecasting performance and the sensitivity of impulse response functions to differing weights. In favor of simplicity, the scope of this chapter is local rather than international so that the popular usage of purely distance weights can be clearly contrasted with the convex alternative. Specifically, using a metro level model of the Texas housing market, the core issue to be revealed is that the outcomes, both forecasting and dynamic analysis, can vary meaningfully across models with different weights. Due to this finding, an important issue is brought to light which is the fact that, to date, only a very small portion of the literature has presented findings with specific focus on robustness to weights.

Regarding forecasting performance, this chapter shows that models utilizing convex weights can outperform those utilizing pure weights[1]. In consideration of the generally understood trade-off between complexity and forecasting accuracy, this finding is somewhat surprising[2]. However, the forecasting performance improvement is not completely unexpected taking into account the substantial parameter variability across differing weighting schemes shown in Gross (2018) and the forecasting improvements from utilizing the mixing on variables concept developed by Eickmeier and Ng (2015) that are shown in Martin and Cuaresma (2017).

As for the sensitivity of impulse response functions to weighting scheme choice, this chapter lays the foundation for the more in depth analysis of chapter 3 by showing that the conclusions reached from analyzing impulse responses can differ greatly across models utilizing different weighting schemes. This finding is extremely relevant in its illustration that the conclusions argued by a researcher might drastically change if only he were to estimate under a different weighting scheme. The demonstration of significant impulse response sensitivity in the simple but relatable context of this chapter in combination with the prevalence of the distance class at the regional level and trade weights at the international level leaves something to be desired among the existing literature[3].

[1] Concerning robustness, the GVAR models are compared to more rudimentary methods and are found to outperform in all cases except for at the shortest horizon. See Table 2.9 for details.

[2] Green and Armstrong (2015) reviews evidence from 32 papers on the accuracy of forecasts from complex vs. simple methods and finds that complexity fails to improve accuracy in all but 16 of 97 comparisons.

[3] The meaning of significance in terms of impulse response sensitivity is illustrated by the case for which a researcher might conclude that a response is significant under one weighting scheme but insignificant under another.

A major concern is that any research not having presented robustness to weighting might be in question. Of course, such questions don't apply to cases where the linkage mechanism had theoretical justification but for those cases where the weighting scheme may have been "determined in an ad hoc fashion," to quote Martin and Cuaresma (2017), the question may be more than relevant. Furthermore, this chapter shows that beyond the simulation exercises in Gross (2018), weighting scheme sensitivity is more than just a theoretical issue.

The remaining sections of this chapter are organized as follows: Section 2.2 presents the data and the model as well as a brief discussion on it's place among the existing regional GVAR literature after which subsections detailing the convex weightings and the standard diagnostics follow. Section 2.3 discusses forecasting performance and the notion of an optimal weighting scheme. Section 2.4 covers impulse response sensitivity to weighting and provides context for the substantive spillover and transmission effects results. Section 2.5 concludes.

2.2 Regional Application (1990-2017)

The small scale GVAR to be used for the illustration of weighting optimality covers the four largest Metropolitan Statistical Areas (MSA) in Texas shown in Table 2.1. With the current coverage, the present GVAR model accounts for just over 80% of the state's total output and around 75% of the state's population.

Table 2.1: MSA's in the GVAR Model

Dallas, TX	San Antonio, TX
Houston, TX	Austin, TX

The models are estimated over the period 1990(1)2017(10). The use of monthly data is in itself an improvement over the existing literature that most commonly use quarterly data[4]. In order to capture more fully the possible effects of the global economy on localities, the US GDP (y_t), the US Fed Funds rate (r_t) to represent the nominal short-term rate, the oil price (oil_t), the S&P 500 closing price (q_t), and inflation (π_t) are included as global variables. Other variables included are local total monthly wages across all industries (w_{it}), local total employment across all industries (e_{it}), and local housing prices (hp_{it}) as well as the foreign housing counterpart ($hp^{\star}_{it}$). More specifically,

$$
\begin{aligned}
&y_t = \ln(GDP_t/CPI_t), && p_t = \ln(CPI_t), && q_t = \ln(EQ_t/CPI_t),\\
&\rho^S_t = .08\bar{3} \times \ln(1 + R^S_t/100), && hp_{it} = \ln(HP_{it}/CPI_t), && w_{it} = \ln(W_{it}/CPI_t),\\
&e_{it} = \ln(E_{it}), && oil_t = \ln(OIL_t)
\end{aligned}
$$

[4] Vansteenkiste and Hiebert (2011), Jannsen (2010), and Cipollini et al. (2018) all focus on housing prices and use quarterly data.

where GDP_t is the nominal Gross Domestic Product, CPI_t the consumer price index, EQ_t the S&P500 closing price, R_t^S the short-term rate, OIL_t the West Texas Intermediary (WTI) closing price, HP_{it} the nominal housing price, W_{it} the total wages, and E_{it} the total employment for MSA i in period t[5]. All global variables are collected from Federal Reserve Economic Data (FRED), with employment and wage data coming from the Bureau of Labor Statistics (BLS), and the housing price variables from the Real Estate Center of Texas A&M University.

The only MSA specific foreign variable, housing price hp_{it}^*, is constructed using eq. (1.11) with distance and output weights. Vansteenkiste (2007) uses distance weights to describe the linkages of housing prices in the United States at the state level and concludes that the weighting scheme works well, but Martin and Cuaresma (2017) shows that the forecasting performance of distance weights is significantly worse than some other weights. Other weighting schemes such as contiguity have been used by Cipollini et al. (2018) and Choi and Chudik (2017) in small universe type analyses, however, contiguity weighting is fundamentally the same as inverse distance weighting with high decay so it is reasonable to assume that the conclusion of Martin and Cuaresma (2017) might also hold in regard to contiguity. Initially, fixed weights are used based on the average over the three year period 2013-2016. Allowing for time-varying weights is straightforward and is considered in Section 2.3

With the exception of the dominant unit model, all models include the MSA-specific foreign variable hp_{it}^* and the global variables y_t, eq_t, ρ_t^s, oil_t, p_t as weakly exogenous in the sense discussed in Dees et al. (2007). The dominant unit model contains all global variables exclusively without feedbacks.

2.2.1 Weighting Schemes

The baseline convex combination weights using distance and output to construct the MSA-specific foreign variables are given in the 4×4 matrix below:

[5] Quarterly nominal GDP is converted to monthly frequency by fitting a local quadratic polynomial for each observation of the quarterly series, then using this polynomial to fill in all observations of the monthly series. The quadratic polynomial is formed by taking sets of three adjacent points from the quarterly series and fitting a quadratic so that the average of the monthly frequency points matches the quarterly frequency data. This procedure is carried out in Eviews.

Table 2.2: Convex Combination Weights

MSA	Dallas	Austin	Houston	San Antonio
Dallas	0.000	0.340	0.467	0.316
Austin	0.293	0.000	0.296	0.356
Houston	0.482	0.347	0.000	0.328
San Antonio	0.224	0.313	0.237	0.000

Note: Convexity parameters: $\theta_1 = \alpha_1 = 0.5$. Columns sum to unity.

In this case, the convex weights are computed as follows: $w_{ij} = \theta_1(d_{ij}) + \alpha_1(GDP_{ij})$ where i and j are locations in the ij location pair, θ_1 and α_1 are convexity parameters, d_{ij} is the ij^{th} element from the column standardized inverse arc distance hollow B matrix, and GDP_{ij} is the ij^{th} element from the column standardized nominal output hollow G matrix.

First considering Dallas, in contrast to purely distance weights where Houston would receive a much smaller share simply due to being further away, under the convex weighting scheme it receives the largest share of the cities included, as it arguably should, since it is a much larger and more influential city. Regarding Austin, notice that rather than San Antonio receiving greater than 50% and Dallas and Houston receiving less than 30% as would be the case under pure distance, all cities receive roughly uniform weight. Dallas receives the largest share for Houston, and the weights for San Antonio are roughly the same as Austin. The weighting behavior in Table 2.2 is a result of the chosen convexity parameters $\theta_1 = 0.5$ and $\alpha_1 = 0.5$. The complete range of θ and α values is considered in Section 2.3

2.2.2 Unit Root Tests

Pesaran et al. (2004) is followed with the assumption that the variables included in the model are integrated of order one ($I(1)$) so that we can distinguish between short-run and long-run relations and interpret the latter as cointegrating. To test this assumption, ADF tests are conducted on each individual series in levels, first difference, and second differences. Following Dees et al. (2007), Table 2.3 reports unit root t-statistics based on weighted symmetric (WS) estimation of ADF type regressions introduced by Park and Fuller (1995). The lag length employed in the WS unit root tests is selected by the SBC based on standard ADF regressions[6].

[6] In results not reported, additional unit root tests using sequential-t and MAIC described in Ng and Perron (2001) to select lag orders are conducted to which no meaningful differences are found.

Table 2.3: Weighted Symmetric ADF Unit Root Test Statistics - Domestic/Foreign Variables

Variable	Critical Value	Dallas	Austin	Houston	San Antonio
e	-3.24	-2.73	-2.41	-2.79	-2.07
Δe	-2.55	-2.53	-2.79	-3.91	-3.96
w	-3.24	-1.76	-1.62	-1.99	-1.57
Δw	-2.55	-5.36	-4.62	-5.69	-8.18
hp	-3.24	-1.09	-1.69	-2.07	-3.32
Δhp	-2.55	-14.15	-13.60	-5.61	-8.48
$hp^{\star}$	-3.24	-2.44	-2.19	-1.96	-2.01
$\Delta hp^{\star}$	-2.55	-14.55	-14.87	-12.16	-11.91

Note: WS statistics for all levels variables are based on regressions including a linear trend. Lag orders are based on SBC Order Selection.

Total employment, total wages, local housing prices, and foreign housing are generally $I(1)$ across all MSAs. Total employment in Dallas and local house prices in San Antonio are borderline $I(0)/I(1)$ but are very close. Broadly speaking however, the test results for domestic and foreign variables support the unit root hypothesis.

Next regarding global variables, Table 2.4 reports *WS* statistics for consumer prices, the nominal short-term rate, real output, the oil price, and equity prices which all are $I(1)$. These results are also in support of the unit root hypothesis.

Table 2.4: Weighted Symmetric ADF Unit Root Test Statistics - Global Variables

Variable	Critical Value	Statistic
p	-3.24	-0.07
Δp	-2.55	-11.87
r	-2.55	-1.21
Δr	-2.55	-6.59
y	-3.24	-1.62
Δy	-2.55	-10.42
oil	-3.24	-2.23
Δoil	-2.55	-10.71
q	-3.24	-1.64
Δq	-2.55	-11.51

Note: WS statistics for all levels variables are based on regressions including a linear trend, except for the nominal short rate variable. Lag orders are based on SBC Order Selection.

2.2.3 Specification and Estimation of the MSA-Specific Models

The current version of the model imposes the same specification across the MSA-specific models, however this need not be the case. For each model, local employment (e), local total wages (w), and local housing prices (hp) are included as endogenous variables. Foreign housing prices ($hp^{\star}$), the national consumer price index (p), the national nominal short-term rate (r), the national real output (y), the oil price (oil), and real equity prices (q) are included as weakly exogenous. The inclusion of the global variables allows for each model to be more fully integrated in the national economy and hence to take a more satisfactory account of the second round effects in the national economic system as a whole. Hence it is crucial that the weak exogeneity of these variables be tested, as is done in Section 2.2.4

After having specified the variables to be included in the individual models, the corresponding cointegrating VAR models are estimated and the rank of their cointegrating space is determined. In the models that are considered in this chapter, the order of the individual MSA VARX*(p_i,q_i) models is determined by SBC where p_i is the lag order on domestic variables and q_i is the lag order on foreign variables[7]. Note that p_{max} and q_{max} are equal to 12 in accordance with the monthly frequency of the data. We then proceed with the cointegration analysis, where the MSA-specific models are estimated subject to reduced rank restrictions[8].

The orders of the VARX* models and the number of cointegration relationships are reported in Table 2.5. For most MSAs, a VARX*(4,1) specifications seemed to be satisfactory. For San Antonio however, a VARX*(2,1) was favored by SBC. Regarding the number of cointegrating relationships, 2 are found for all MSAs except for Austin for which 1 is found. The cointegration analysis follows Dees et al. (2007) with results based on the trace statistics (at the 95% critical value level).

[7] SBC is used throughout this application due to the findings of Koehler and Murphree (1988) who compare AIC and SBC and conclude that SBC is superior for forecasting applications.

[8] The rank of the cointegrating space for each MSA model is computed using Johansens trace and maximal eigenvalue statistics as set out in Pesaran et al. (2000) for models with weakly exogenous I(1) regressors, under case IV.

Table 2.5: VARX* order and number of cointegration relationships in the MSA-specific models

	VARX*(p_i,q_i)		# Cointegrating relationships
Country	p_i	q_i	
Dallas	4	1	2
Austin	4	1	1
Houston	4	1	2
San Antonio	2	1	2

Note: The rank of the cointegrating space for each MSA was computed using Johansen's trace and maximal eigenvalue statistics as set out in Pesaran et al. (2000) for models with weakly exogenous $I(1)$ regressors, in the case where unrestricted constants and restricted trend coefficients are included in the individual country error correction models.

2.2.4 Testing Weak Exogeneity

The main assumption underlying the estimation strategy is that the foreign aggregate variables $\mathbf{x}^{\star}$ are weakly exogenous with respect to the long-run parameters of the conditional models. This section presents a formal test of this assumption and the corresponding results of testing the foreign star variables and the global variables.

Table 2.6: F-statistics for testing weak exogeneity of MSA-specific foreign and global variables

MSA		Foreign and Global Variables					
		hp^*	p	r	y	oil	q
Dallas	F(2,295)	0.36	3.29*	0.28	3.87*	0.05	1.01
Austin	F(1,296)	1.63	0.08	0.01	0.03	2.18	0.11
Houston	F(2,295)	0.86	2.47	0.55	3.15*	1.03	0.75
San Antonio	F(2,295)	2.65	0.19	1.59	0.10	1.17	0.15

Note: * denotes statistical significance at the 5% level.

The weak exogeneity test conducted here is the same as is used in Dees et al. (2007) who describe it's origination from Johansen (1992) and Harbo et al. (1998). A test is carried out of the joint significance of the estimated error correction terms in auxiliary equations for the MSA-specific foreign variables, $\mathbf{x}^*_{it}$. Specifically for each lth element of the $\mathbf{x}^*_{it}$ the following regression is carried

out:

$$\Delta x^*_{it,l} = \mu_{it} + \sum_{j=1}^{r_i} \gamma_{ij,l} ECM^j_{i,t-1} + \sum_{k=1}^{s_i} \varphi'_{ik,l} \Delta \mathbf{x}_{i,t-k} + \sum_{m=1}^{n_i} \vartheta'_{im,l} \Delta \tilde{\mathbf{x}}^*_{i,t-m} + \epsilon_{it,l}$$

where $ECM^j_{i,t-1}$, j = 1,2,...,r_i are the estimated error correction terms corresponding to the r_i cointegrating relations found for the ith MSA model and $\Delta \tilde{\mathbf{x}}^*_{it} = (\Delta \mathbf{x}'^*_{it}, \Delta p_t, \Delta r_t, \Delta y_t, \Delta oil_t, \Delta q_t)'$. The test for weak exogeneity is an F-test of the joint hypothesis that $\gamma_{ij,l} = 0$, j = 1,2,...,r_i in the above regression. The lag orders s_i and n_i, need not be the same as the orders p_i and q_i of the underlying MSA-specific VARX* models. In the test that is carried out, the lag order was chosen by SBC and under these specifications, 3 out of 24 cases are found to be significant at the 5% significance level. The test results for this specification are reported in Table 2.6

For the set of MSAs, as can be seen from this table, the weak exogeneity assumptions are rejected only for inflation, in Dallas, and output in both Dallas and Houston. Since the vast majority of the weak exogeneity tests are rejected, the tests suggest that the foreign and global variables can be considered as weakly exogenous and that the key assumptions that underlie the GVAR modelling are not generally violated.

2.3 Forecasting Sensitivity

In the context of a GVAR model, just as any specification change can alter the results of a model, so too should the choice of weighting scheme since it alters the underlying data behind the specification. This section addresses the impact that weighting matrices can have on the underlying data, shows how convex weightings computed via Equation (1.11) can lead to improvements, and outlines a procedural approach for how to think about weighting optimization through an example.

2.3.1 Foreign Variable Sensitivity

In the preceding analysis, baseline convexity parameters $\theta_1 = \alpha_1 = 0.5$ were used simply on the grounds that they give equal weight to both distance and output. Moving beyond this, it is helpful to first illustrate what the convexity parameters actually control. Hence, the foreign housing variables (hp^*) is computed across the range $\theta_1 \in$ [0,1] in increments of 0.1 under both fixed and time varying settings. Although eq. (1.11) can be computed under either setting, it is commonly found in many applications that fixed and time varying weights typically correlate strongly and thus the utilization of the relatively more simple fixed weights is justified. Another justification for fixed weights is their necessity in the generation of forecasts. Unless time-varying weights are endogenously determined, usually weights from the last available observation are used as fixed to generate forecasts. This inconsistency is unappealing and hence it is more satisfying to show that

fixed weights are entirely sufficient. Before discussing convex weighting optimality, this section first provides evidence of the relationship between the various weighting parameter choices in the simple case of the two parameter application. Table 2.7 provides correlation coefficients of the MSA specific foreign housing price variables hp^* across the θ_1 range in the fixed setting while Table 2.8 provides correlation coefficients between fixed and time varying settings at each θ_1 value.

Table 2.7: Correlation coefficients of MSA-specific foreign housing price variables using different values of θ_1

Housing Price - Levels - 2016 Fixed Weights								
MSA / Convexity	0.0 - 0.5	0.2 - 0.5	0.4 - 0.5	0.6 - 0.5	0.8 - 0.5	1.0 - 0.5	0.0 - 1.0	0.1 - 0.9
Dallas	0.9994	0.9998	1.0000	1.0000	0.9995	0.9982	0.9956	0.9974
Austin	0.9966	0.9987	0.9998	0.9998	0.9985	0.9955	0.9843	0.9900
Houston	0.9983	0.9992	0.9999	0.9999	0.9983	0.9933	0.9849	0.9911
San Antonio	0.9974	0.9990	0.9999	0.9999	0.9985	0.9950	0.9854	0.9908
Housing Price - 1st diff. - 2016 Fixed Weights								
MSA / Convexity	0.0 - 0.5	0.2 - 0.5	0.4 - 0.5	0.6 - 0.5	0.8 - 0.5	1.0 - 0.5	0.0 - 1.0	0.1 - 0.9
Dallas	0.9662	0.9855	0.9981	0.9977	0.9759	0.9240	0.7941	0.8668
Austin	0.9456	0.9770	0.9970	0.9966	0.9668	0.9037	0.7153	0.8105
Houston	0.9730	0.9878	0.9983	0.9978	0.9741	0.9088	0.7880	0.8675
San Antonio	0.9310	0.9693	0.9958	0.9950	0.9494	0.8522	0.6024	0.7318

Note: Recall that the convexity weights are computed via Equation (1.11). Each column presents the correlation of the foreign housing price variable under the respective θ_1 value to the baseline case of $\theta_1 = 0.5$ (e.g., the third column '$0.4 - 0.5$' presents the correlation between the foreign housing price variables under the two cases $\theta_1 = 0.4$, $\alpha_1 = 0.6$ and $\theta_1 = \alpha_1 = 0.5$). In all cases θ_1 is calibrated such that $\theta_1 + \alpha_1 = 1$. The last two columns report the correlations between the cases $\theta_1 = 0$ to $\theta_1 = 1$ and $\theta_1 = 0.1$ to $\theta_1 = 0.9$ respectively.

As is evidenced by the results in Table 2.7, across all θ_1 specifications, there is little difference in the foreign variable values to the baseline in either levels or in first differences. However, focusing more on first differences as the foreign housing price variables are likely to be I(1), the last two columns show that there is at least a moderate difference between pure GDP weights ($\theta_1 = 0$) and pure distance weights ($\theta_1 = 1$) with the correlation coefficients ranging from 0.60 to 0.79. This is evidence, although not statistically formal, that unsurprisingly a difference exists between weighted foreign aggregates computed under purely distance and purely GDP weighting schemes. Hence, we might expect the results to differ if not meaningfully for impulses, perhaps to a more noticeable degree for forecasting. Furthermore, the non-unity correlation suggests that some optimal mix exists which captures the best blend between distance and GDP.

Lastly regarding robustness to fixed versus time varying weighting schemes, Table 2.8 shows the

correlations between each temporal strategy for the same θ_1 specification.

Table 2.8: Correlation coefficients of MSA-specific foreign housing price variables using fixed and time varying weights while varying θ_1

Housing Price - Levels - Fixed to Time Varying Correlation						
MSA/θ_1	0.0	0.2	0.4	0.6	0.8	1.0
Dallas	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000
Austin	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000
Houston	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
San Antonio	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000

Housing Price - 1st diff. - Fixed to Time Varying Correlation						
MSA/θ_1	0.0	0.2	0.4	0.6	0.8	1.0
Dallas	0.9985	0.9990	0.9995	0.9998	1.0000	1.0000
Austin	0.9989	0.9993	0.9996	0.9998	1.0000	1.0000
Houston	0.9995	0.9997	0.9999	0.9999	1.0000	1.0000
San Antonio	0.9985	0.9990	0.9995	0.9998	1.0000	1.0000

Note: Each column represents the correlation between fixed weights and time varying weights at each respective θ_1 value. The first panel reports levels and the second reports first differences.

Simply by definition, as θ_1 approaches unity the correlation coefficients approach unity. Across all values of θ_1 however, the coefficients are extremely close to unity and thus it is unlikely that the use of time varying weights would meaningfully impact either the impulse responses or the forecasts. Thus, the exclusive use of fixed weights is justified.

2.3.2 Convex Optimality

To follow the discussion in section section 1.4 on parameter choice, the definition of optimality is also of critical importance. Informational evaluation methods such as AIC and SBC are based on the trade-off between the goodness of fit and model simplicity and while these methods are certainly helpful, they are not the only methods. Alternatively, model predictability as was argued by Milton Friedman in his 1953 work on Positive Economics [9], is paramount to success of any theory. It was argued that without any accurate predictions, a theory is nothing more than useless. To this end and in a similar manner to Martin and Cuaresma (2017), while informational methods are later entertained, forecasting accuracy is focused on as the chief evaluation method.

[9] Chapter seven of Friedman (1953) presents a discussion on the benefits of focusing on positive rather than normative economics and emphasizes that in order to move closer to the positive, one must focus on the forecasting performance of a model.

Model complexity is an issue not only related to information criterion but also one that is central to the ability of a model to forecast well. As more and more degrees of freedom are used up, the noisiness of coefficient estimates increases and as a result, so too does the forecasting performance degrade. This issue over complexity is particularly of interest here even prior to any discussion over weighting schemes. Since the GVAR in itself is a relatively more complicated model with more parameters than some of it's simpler predecessors, it should first be illustrated that a GVAR model is even necessary in forecasting applications. In order to assess forecasting ability, a period of roughly 25 years starting in January of 1990 and ending in October of 2014 is used for training, leaving 36 months (2014M11-2017M11) available for out-of-sample forecasting evaluation. Following Martin and Cuaresma (2017), out-of-sample forecasting performance is assessed using the root mean squared forecasting errors (RMSE) relative to the RMSE that would be obtained using a random walk prediction.

Housing prices are forecast for each MSA using a series of different models and are evaluated across seven time intervals ranging from one month ahead ($h = 1$) to three years ahead ($h = 36$). The models tested include two 'naive' VAR models, both with lag orders optimally selected via SBC, of which the first includes only the housing prices of the four MSAs and the second includes the global variables as well. Also included in the testing are SBC optimal univariate ARIMA models for each location. As for GVAR models, three are initially tested with one utilizing convex weights. The two non-convex weighted GVARs are those weighted by pure distance and pure GDP. The convexly weighted GVAR model is the baseline model that mixes equally between distance and GDP. Each GVAR model is exactly specified in accordance with what is described in Section 1.2 To best evaluate the overall performance, irrespective of location, RMSEs can be averaged across locations at each horizon and then compared to the random walk forecast RMSE. These results are presented in Table 2.9 and illustrate the general advantage of utilizing GVAR models in particular but also GVAR models with mixed weights. In this sense of overall performance, GVAR models perform better than their non GVAR counterparts in all horizons except for the one step ahead forecast, for which the ARIMA models performed best. Across the other six horizons, 100% of the winning models are from the GVAR with mixed weights. Another noteworthy finding is the increased performance over the random walk as the horizon lengthens. In the one step ahead forecast, averaged RMSEs from the best performing model are only 23% lower than the random walk RMSE while at three year horizon, averaged RMSEs from the best model are 32% better.

Table 2.9: Predictive Accuracy Results - Model Comparison

Model (SBC Lag)	Average Universal Performance - Housing Prices							
	h=1	h=6	h=12	h=18	h=24	h=30	h=36	avg
VAR(2) - Only hp	2.089	1.915	1.790	1.642	1.527	1.419	1.362	1.678
VAR(3) - All Variables	1.872	1.752	1.631	1.493	1.383	1.284	1.228	1.520
SBC Optimal Univariate ARIMA	**0.769**	0.841	0.772	0.748	0.714	0.684	0.668	0.742
GVAR : Pure Distance	0.890	0.864	0.739	0.683	0.639	0.591	0.579	0.712
GVAR : Pure GDP	0.780	0.813	0.748	0.732	0.689	0.642	0.625	0.718
GVAR : 0.5 Dist + 0.5 GDP	0.833	**0.805**	**0.692**	**0.658**	**0.624**	**0.579**	**0.571**	**0.680**

Note: Bold figures indicate minima. Average universal performance is calculated as the average RMSE across the four included MSAs relative to the Random Walk RMSE.

The results in Table 2.9 provide strong evidence for convex weights defined by Equation (1.11) and directly imply that an optimal weighting must exist since, for a given horizon, the performance of the mix is better than the performance at either convexity bound[10]. To find the best performing mix, a grid search procedure is conducted with 21 GVAR models estimated at each mix across the θ_1 range in intervals of 0.05. The results from this grid search are presented below in Table 2.10.

[10] In results not reported, the null hypothesis of the Diebold and Mariano (2002) test of *equal predictive accuracy* to the baseline is rejected for every forecast within Table 2.9 which further illustrates that the convexly weighted GVAR resulted in better forecasting performance.

Table 2.10: Convex Weights Grid Search - RMSE Results
Detailed Tabular Representation

Model (sbc lag)	Average Universal Performance - Housing Prices							
	h=1	h=6	h=12	h=18	h=24	h=30	h=36	avg
GVAR : 0.00 Dist + 1.00 GDP	**0.780**	0.813	0.748	0.732	0.689	0.642	0.625	0.718
GVAR : 0.05 Dist + 0.95 GDP	0.781	0.822	0.756	0.737	0.692	0.644	0.627	0.723
GVAR : 0.10 Dist + 0.90 GDP	0.783	0.828	0.761	0.740	0.694	0.646	0.628	0.726
GVAR : 0.15 Dist + 0.85 GDP	0.786	0.836	0.767	0.743	0.696	0.648	0.630	0.729
GVAR : 0.20 Dist + 0.80 GDP	0.788	0.844	0.774	0.747	0.699	0.650	0.632	0.733
GVAR : 0.25 Dist + 0.75 GDP	0.791	0.853	0.781	0.751	0.702	0.653	0.634	0.738
GVAR : 0.30 Dist + 0.70 GDP	0.795	0.862	0.788	0.756	0.705	0.655	0.636	0.742
GVAR : 0.35 Dist + 0.65 GDP	0.798	0.872	0.796	0.760	0.708	0.657	0.638	0.747
GVAR : 0.40 Dist + 0.60 GDP	0.802	0.882	0.804	0.764	0.711	0.660	0.640	0.752
GVAR : 0.45 Dist + 0.55 GDP	0.806	0.892	0.812	0.768	0.714	0.661	0.641	0.756
GVAR : 0.50 Dist + 0.50 GDP	0.833	**0.805**	**0.692**	**0.658**	**0.624**	**0.579**	**0.571**	**0.680**
GVAR : 0.55 Dist + 0.45 GDP	0.839	0.812	0.698	0.662	0.626	0.581	0.573	0.685
GVAR : 0.60 Dist + 0.40 GDP	0.845	0.818	0.702	0.664	0.628	0.582	0.574	0.688
GVAR : 0.65 Dist + 0.35 GDP	0.851	0.824	0.707	0.666	0.629	0.583	0.574	0.691
GVAR : 0.70 Dist + 0.30 GDP	0.857	0.830	0.711	0.669	0.631	0.584	0.575	0.694
GVAR : 0.75 Dist + 0.25 GDP	0.863	0.836	0.716	0.671	0.632	0.585	0.576	0.697
GVAR : 0.80 Dist + 0.20 GDP	0.868	0.841	0.721	0.674	0.634	0.586	0.576	0.700
GVAR : 0.85 Dist + 0.15 GDP	0.874	0.847	0.725	0.676	0.635	0.587	0.577	0.703
GVAR : 0.90 Dist + 0.10 GDP	0.879	0.853	0.730	0.679	0.637	0.589	0.578	0.706
GVAR : 0.95 Dist + 0.05 GDP	0.884	0.859	0.735	0.681	0.638	0.590	0.579	0.709
GVAR : 1.00 Dist + 0.00 GDP	0.890	0.864	0.739	0.683	0.639	0.591	0.579	0.712

Note: Bold figures indicate minima. Average universal performance is calculated as the average RMSE across the four included MSAs relative to the Random Walk RMSE. Lag order for all models is selected via SBC.

Interestingly, the midpoint mix $\theta_1 = \alpha_1 = 0.5$ resulted in the best performance at every horizon except for the the shortest. To better illustrate the results in Table 2.10 and to further exemplify any patterns, Figure 2.1 presents a graphical representation of the all horizon average relative RMSE. While the convexity mix $\theta_1 = \alpha_1 = 0.5$ coincidentally resulted in the lowest relative RMSE, the trend towards better performance as θ_1 declines is consistent across the entire range but with a performance break at $\theta_1 = 0.45$.

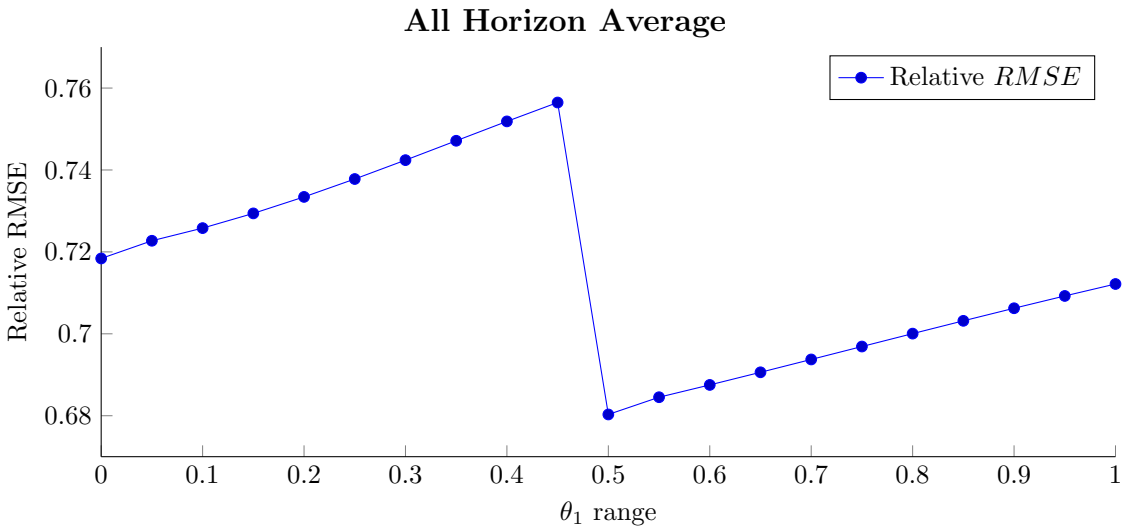

Figure 2.1: Convex Weights Grid Search - RMSE Results
Select Graphical Representation

Upon inspection of the lag orders of each GVAR model in the grid search, it is found that $p_{i|\theta_1<0.5} \neq p_{i|\theta_1\geq 0.5}$. Specifically in this case, the lag order for domestic variables in the San Antonio model is where the change occurs (i.e., $p_{san} = 4$ when $\theta_1 \geq 0.5$ and $p_{san} = 2$ when $\theta_1 < 0.5$). Hence, it can be said that a 'lag break' occurs at the critical point $\theta_1 = 0.5$.

Table 2.11: Convex Weights Grid Search - Log Likelihood and Information Criterion

Model (sbc lag)	Average Universal Performance		
	logLik	Akaike	Schwartz Bayesian
GVAR : 0.00 Dist + 1.00 GDP	2797.771	2753.521	2672.632
GVAR : 0.10 Dist + 0.90 GDP	2797.878	2753.628	2672.739
GVAR : 0.20 Dist + 0.80 GDP	2797.900	2753.650	2672.762
GVAR : 0.30 Dist + 0.70 GDP	2797.791	2753.541	2672.652
GVAR : 0.40 Dist + 0.60 GDP	2797.564	2753.314	2672.425
GVAR : 0.50 Dist + 0.50 GDP	**2811.753**	**2763.003**	**2673.888**
GVAR : 0.60 Dist + 0.40 GDP	2811.564	2762.814	2673.699
GVAR : 0.70 Dist + 0.30 GDP	2811.278	2762.528	2673.413
GVAR : 0.80 Dist + 0.20 GDP	2810.955	2762.205	2673.091
GVAR : 0.90 Dist + 0.10 GDP	2810.604	2761.854	2672.739
GVAR : 1.00 Dist + 0.00 GDP	2810.232	2761.482	2672.368

Note: Bold figures indicate maxima. Average universal performance is calculated as the average log likelihood or information value across the four included MSAs. Maximum values are considered best for AIC/SBC due to the Dees et al. (2007) formulation. Increments of 0.1 are reported for brevity.

This prompts an inspection of the information criterion of the models within the grid search. Table 2.11 presents the log likelihood, AIC and SBC values for each model and reveals an identical pattern. It is, however, worth noting the rule of thumb outlined in Burnham and Anderson (2004) which states that if the difference in information between the ith model and the best performing model is less than the value of 2, then there is still substantial support for model i. Given that this is indeed the case across the results of the entire grid search, information criteria evaluation methods are hardly helpful and suggest that the choice of weighting has a minimal impact. This is to be expected though as the models are of the exact same specification, only with the underlying data of one foreign variable being minimally altered (recall column seven of Table 2.7). Hence, this is additional evidence that it is more useful to rely upon evaluations of forecasting performance. Nevertheless, the same pattern being found in both forecasting evaluation and information criteria indicates that yet a greater improvement could be made. It should be noted though that the potential for refinement is conditional on a single break, as is the case in this application, or a harmonious group of 'breaks'. To formalize, under the single lag break observed above, the following conditions can provide guidance:

$$\left(\frac{\delta RMSE}{\delta \theta_1}\right) > 0 \implies \text{Use Maximum } \theta_1 \text{ Lag Order}$$
$$\left(\frac{\delta RMSE}{\delta \theta_1}\right) < 0 \implies \text{Use Minimum } \theta_1 \text{ Lag Order}$$

In words, if the slope of the relative RMSE curve on both sides of the lag break is increasing (and since lower relative RMSE is better), then an improvement beyond the mix suggested by the initial grid search could be made by fixing the lag order at the order selected under the maximum value of θ_1 instead of allowing it to be automatically selected via criterion[11]. To evaluate this claim, a second grid search, hereafter referred to as the CX-Refinement, is run with lag orders fixed at the $\theta_1 = 1$ (maximum value / pure distance) order (i.e., $p_i = p_{i|\theta=1}$ and $q_i = q_{i|\theta=1}$. Results of the CX-Refinement grid search are reported below.

Table 2.12: CX-Refinement Grid Search - RMSE Results
Detailed Tabular Representation

Model ($\theta_1 = 1$ lag order)	Average Universal Performance - Housing Prices							
	h=1	h=6	h=12	h=18	h=24	h=30	h=36	avg
GVAR : 0.00 Dist + 1.00 GDP	**0.780**	**0.767**	**0.673**	**0.651**	**0.621**	**0.578**	**0.571**	**0.663**
GVAR : 0.10 Dist + 0.90 GDP	0.789	0.775	0.677	0.653	0.622	0.579	0.572	0.667
GVAR : 0.20 Dist + 0.80 GDP	0.798	0.780	0.678	0.653	0.622	0.578	0.571	0.669
GVAR : 0.30 Dist + 0.70 GDP	0.809	0.787	0.682	0.654	0.622	0.578	0.571	0.672
GVAR : 0.40 Dist + 0.60 GDP	0.821	0.796	0.687	0.657	0.623	0.579	0.571	0.676
GVAR : 0.50 Dist + 0.50 GDP	0.833	0.805	0.692	0.658	0.624	0.579	0.571	0.680
GVAR : 0.60 Dist + 0.40 GDP	0.845	0.818	0.702	0.664	0.628	0.582	0.574	0.688
GVAR : 0.70 Dist + 0.30 GDP	0.857	0.830	0.711	0.669	0.631	0.584	0.575	0.694
GVAR : 0.80 Dist + 0.20 GDP	0.868	0.841	0.721	0.674	0.634	0.586	0.576	0.700
GVAR : 0.90 Dist + 0.10 GDP	0.879	0.853	0.730	0.679	0.637	0.589	0.578	0.706
GVAR : 1.00 Dist + 0.00 GDP	0.890	0.864	0.739	0.683	0.639	0.591	0.579	0.712

Note: Bold figures indicate minima. Average universal performance is calculated as the average RMSE across the four included MSAs relative to the Random Walk RMSE. Lag order for all models is fixed at the $\theta_1 = 1$ sbc selected order.

Table 2.12 shows that improvements are made across all horizons from taking the best performing model from the CX-Refinement grid search ($\theta_1 = 0$) over the best performing model from the CX-Automatic grid search ($\theta_1 = 0.5$) with a magnitude of improvement around 2.5%. Figure 2.2

[11] To simplify the terminology, the initial grid search is referred to as the CX-Automatic grid search since the lag orders and ranks of each model are determined automatically using selection criterion.

illustrates graphically the improvement for the all horizon average relative RMSE from fixing the lag order.

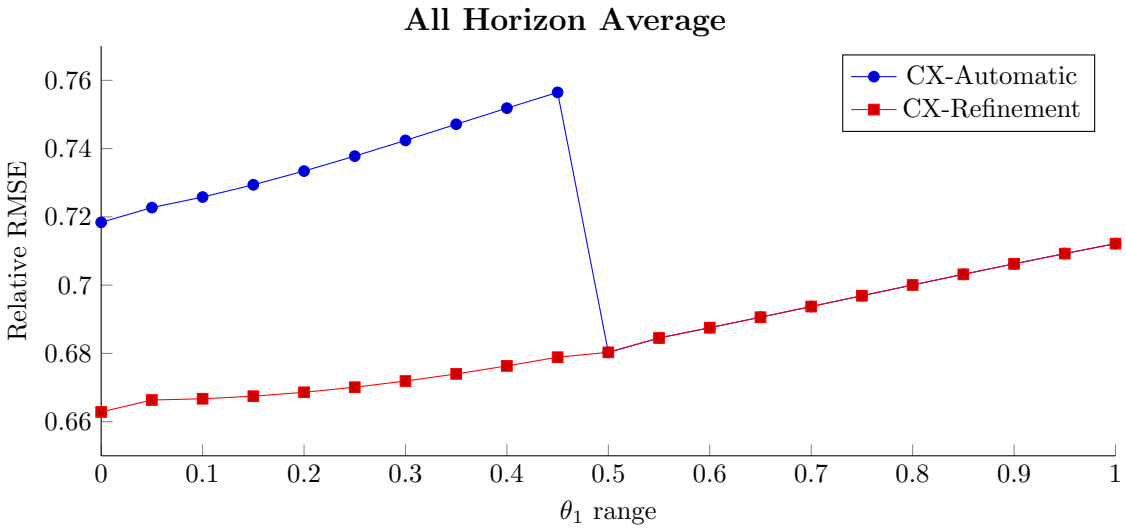

Figure 2.2: CX-Refinement Grid Search - RMSE Results
Graphical Performance Comparison

The specific magnitudes of improvement from the entire procedure are illustrated in Table 2.13. Starting from a commonly used non-convex baseline pure distance weighting scheme, performing the CX-Automatic grid search resulted in identification of a Dist-GDP mixed weights model that performs around 4.5% better. Moving further to the CX-Refinement procedure with fixed lag order, a model performing 2.6% better than the CX-Automatic optimal model was identified. The overall performance from the entire procedure resulted in an increase of nearly 7% over the baseline before convexity or refinements are introduced.

Table 2.13: Overall Performance Increases
Convex Weights and Refinement

Model	Average Horizon	Relative Gain	Total Gain
Pure Distance	0.712		
CX-Automatic	0.680	-4.5%	
CX-Refinement	0.663	-2.6%	-6.9%

Note: CX models with the best performance. Gains are reported as the percentage reduction in relative RMSE from one row to the next.

Up to this point, the discussion has revolved around what was found in the simple illustrative application with only four spatial units. It is worth noting that the guidance for improvement

beyond interior convexity (fixed lag/rank grid search) and thus the CX-Refinement procedure utilized here cannot without loss of generality be expanded to cases for which multiple lag breaks, rank breaks, or any combination of more than a single change occurs or to cases where the signs of the slopes of the respective zones are not the same. It may also become less than feasible to conduct the refinement in the presence of convex mixes containing more than two units, however we can remain hopeful that this particular issue might resolve itself as computing power continues to increase. Presently though, it is likely best to trust the model selection criterion as the largest improvement occurs when moving from the baseline to the initial convex mix under automatic lag and rank selection. The guidance presented here only serves as a starting point for what can be done with regard to weighting matrices and illustrates that improvements are possible. Another last point of interest is on how the optimality is defined. In the case of the model presented here, it is over universal performance where in practical applications one might choose to optimize on a single location or on an average of two or more locations that are of specific interest to the problem at hand. It suffices to say that regardless of definition, convex weight mixing and optimization should at least be entertained as an alternative to exact specification.

2.4 Impulse Response Sensitivity

Beyond forecasting, an extremely common exercise performed in modern applied macroeconomics is impulse response analysis. This type of analysis, as defined by Pesaran (2015) in his recent econometric manual, is used to "measure the time profile of the effect of shocks at a given point in time on the (expected) future value of the variables within a system." The focus of this section is to investigate the sensitivity of impulses to changes in weighting schemes. In addition to this, standard consideration is also given to each shock since the coverage of the model is substantial for the housing markets in Texas. The following set of shocks is considered: (1) a one standard error negative shock to US real GDP; (2) a one standard error negative shock to US equity prices; (3) a one standard error negative shock to oil prices; (4) a one standard error positive shock to the US interest rate; (5) a one standard error positive shock to housing prices in Dallas[12]. A formal summary of the GIRF method which is used throughout this section is presented in Section 1.2.2.

2.4.1 Shock to US GDP

Consider first the GIRFs for a one standard error negative shock to US GDP. This shock is equivalent to a fall of around 0.18% in US GDP per month. As is the main focus of this section, Figure 2.3 reports the response of Austin housing prices using four different weighting schemes.

[12] In the discussion of sensitivity for each shock, a single location is presented for illustrative purposes. The full set of responses for all locations to all shocks can be found in Appendix A.1.

Of these four are two pure (distance and GDP) weighted models and both the CX-Automatic and CX-Refinement optimal models from each respective grid search. As expected, the responses show a highly similar shape across specifications but more critically, by inspection of the distance and GDP panels, it is seen that the confidence bands are indeed different. These top two panels directly illustrate, albeit not in the most extreme way, that the choice of weighting scheme could influence the strength to which an author might argue that a certain effect exists. Specifically, if pure distance weights were used, the top left panel indicates that we might strongly argue that housing prices in Austin are negatively impacted in a significant way. However, if pure GDP weights were used, the top right panel would indicate that the Austin housing market might only marginally be impacted over the long run. Hence, even in this small scale application we see how two moderately different arguments could be formed from what has traditionally been an ad hoc choice. The bottom two panels report the responses of the optimally selected models from each of the CX-Automatic and CX-Refinement grid searches respectively which appear quite similar and are more resemblant to the pure distance weights in terms of significance.

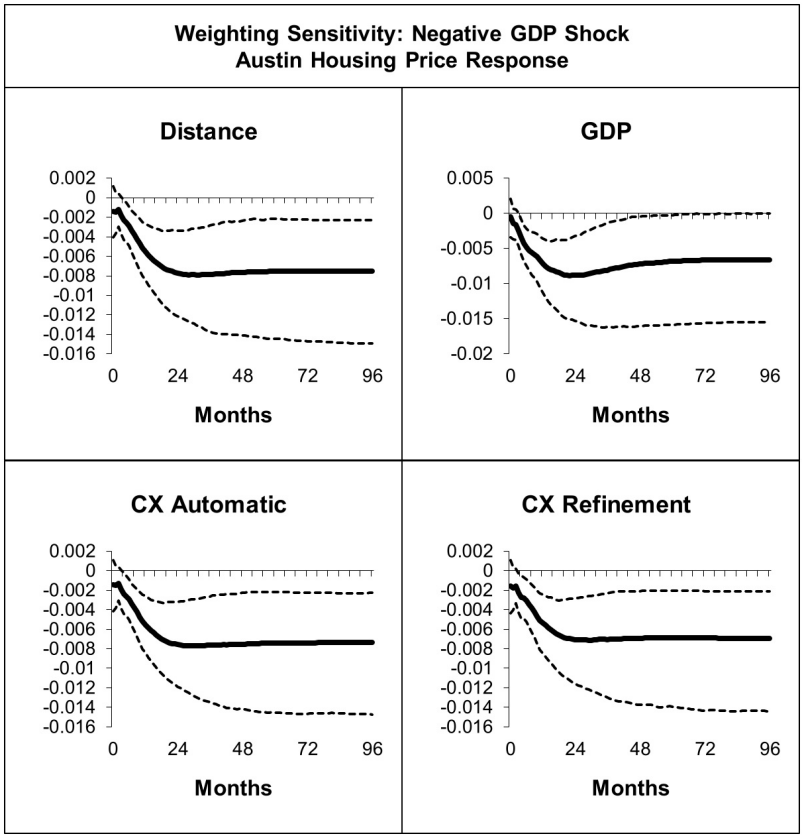

Figure 2.3: GIRFs of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). Austin housing price response under each weighting scheme.

As for more traditional considerations, Figure 2.4 reports the responses of each metro to the same national GDP shock using the CX-Automatic grid search optimal model. To briefly summarize the

findings, the transmission to each housing market takes place rather quickly and in a moderately homogeneous way across each metro. Regarding on impact effects, Dallas, Austin, and Houston are all impacted negatively with monthly declines ranging from around -0.15% to -0.30%. These negative effects are only significant in Dallas and Houston however. San Antonio is the only metro to have a positive point estimate response on impact, although the confidence bands clearly show that the effect is not significantly different from zero. The temporal effects over the first year are negative across all metros with Houston and Dallas having the steepest average monthly declines with an impact after twelve months of -0.80% for both locations. Austin is impacted about half as much as Dallas or Houston and the effect on the San Antonio housing market isn't statistically different from zero. Concerning the long run, all metros show a permanent negative impact again with San Antonio as the outlier for statistical insignificance.

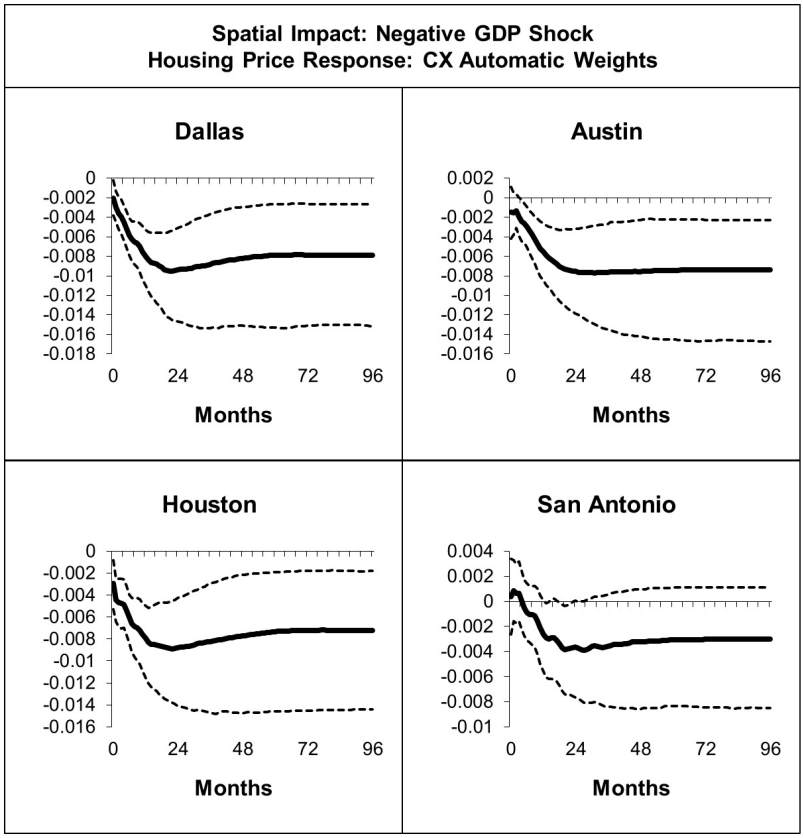

Figure 2.4: GIRFs of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic optimal model housing price responses.

The greater impacts in Dallas and Houston could be driven by a number of factors but it would seem that the effect could simply be related to the size of the metro. Dallas and Houston are substantially larger than Austin and San Antonio and we see that both Dallas and Houston appear to be affected more substantially. Understanding the exact causal impacts of these minimally heterogeneous effects is beyond the scope of this chapter and is a potential path for further research.

2.4.2 Shock to US Equity Prices

Consider next the GIRFs for a one standard error negative shock to US equity prices. This shock is equivalent to a fall of around 2-3% in US real equity prices per month. Figure 2.5 compares the Austin housing market response across different weighting schemes where it can be observed from the top panels that the point estimates exhibit substantial sensitivity.

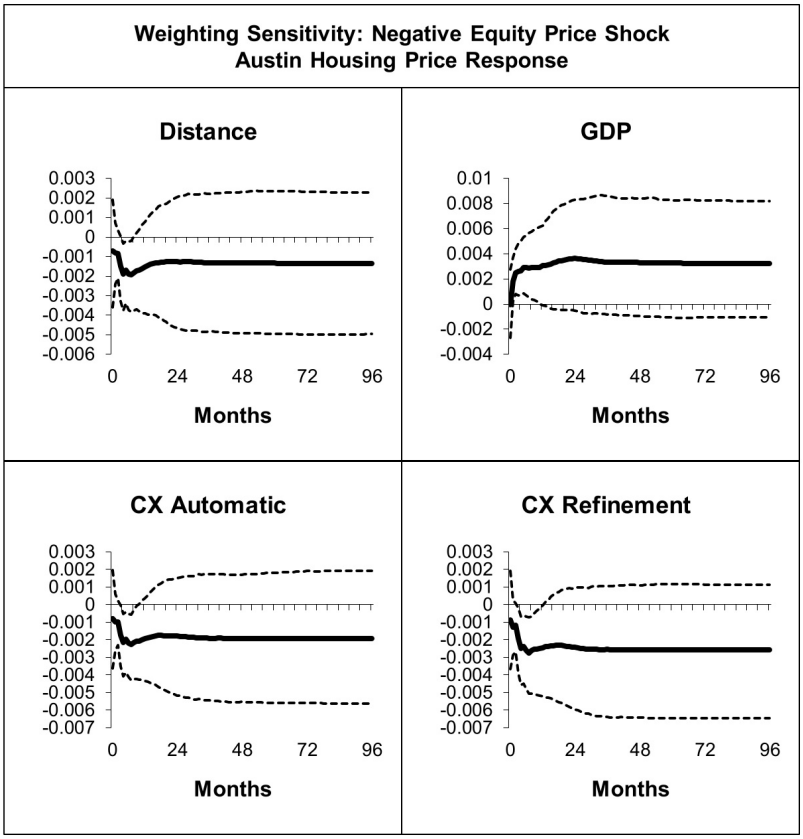

Figure 2.5: GIRFs of a negative (1 s.e.) shock to US Equity Prices (bootstrap mean estimates with 90% bootstrap error bounds). Austin housing price response under each weighting scheme.

The top left panel shows the response from the pure distance weighted model where a negative but insignificant response is observed that persists over the long run. Contrast this with the top right panel that reports the response from the purely GDP weighted model which has a positive and lasting response across the entire horizon. Focusing on the confidence intervals of the top two panels, the conclusion drawn from the distance weighted model would be that housing prices in Austin are not significantly impacted by the equity price shock whereas the conclusion from the GDP weighted model would be that of a borderline positive effect. This case specifically illustrates the danger in estimating and drawing conclusions from a single weighting scheme without checking robustness.

Regarding the transmission to the major housing markets in Texas, Figure 2.6 reports the

responses of each metro from the CX Automatic grid search optimal model.

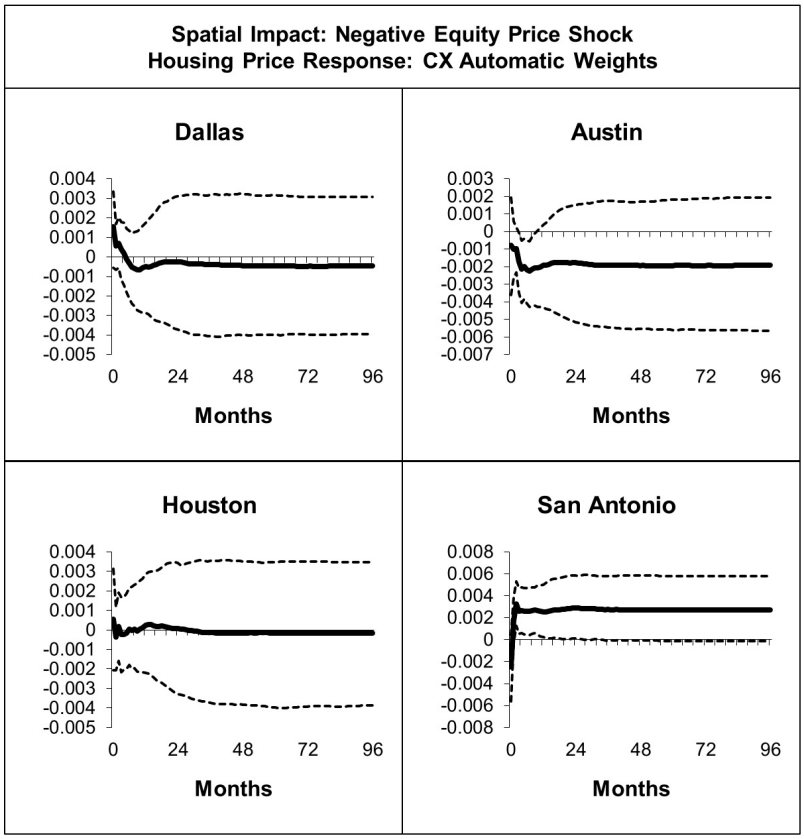

Figure 2.6: GIRFs of a negative (1 s.e.) shock to US Equity (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic optimal model housing price responses.

Responses across all spaces take place rather quickly and in only a moderately heterogeneous way again with San Antonio as the outlier. Regarding on impact effects, Dallas and Houston are impacted positively with Dallas seeing a monthly price jump of 0.16% and Houston to a lesser degree 0.06%. Conversely, Austin and San Antonio are negatively affected on impact with magnitudes of -0.08% and -0.25% respectively. Over the first two years however, Dallas and Houston appear to be more resilient than Austin and San Antonio with the responses being largely insignificant. The temporal effect in Austin and San Antonio do appear significant though but in opposite directions. Austin is impacted negatively by 0.19% on average for the first year while San Antonio is impacted positively by 0.26%. After the first year, effects are clearly insignificant everywhere save for San Antonio where the effect is only borderline. It is worth noting that, in results not reported, under a purely distance weighted model the lower bound for the San Antonio response lies below zero and thus a researcher might alter his argument by stating that the long run response is insignificant across the board whereas utilization of the CX-Automatic optimal model makes such a statement less plausible since the San Antonio bounds are borderline.

The heterogeneous impacts could be driven by a number of factors but it would seem that the

resiliency of a metro to stock prices movements could be due to the same reason listed for the GDP shock. Dallas and Houston are substantially larger than Austin and San Antonio and we see that neither Dallas or Houston appears to be affected. Another contributing factor could be the diversification of labor forces or even more simply, differential abilities across metros in permitting and construction lead times. Again, the exact causal impacts of these heterogeneous effects is beyond the scope of this work and is a potential path for further research.

2.4.3 Shock to Oil Prices

The third shock of interest, a positive one standard error shock to the oil price, represents a monthly increase of around 5-6%.

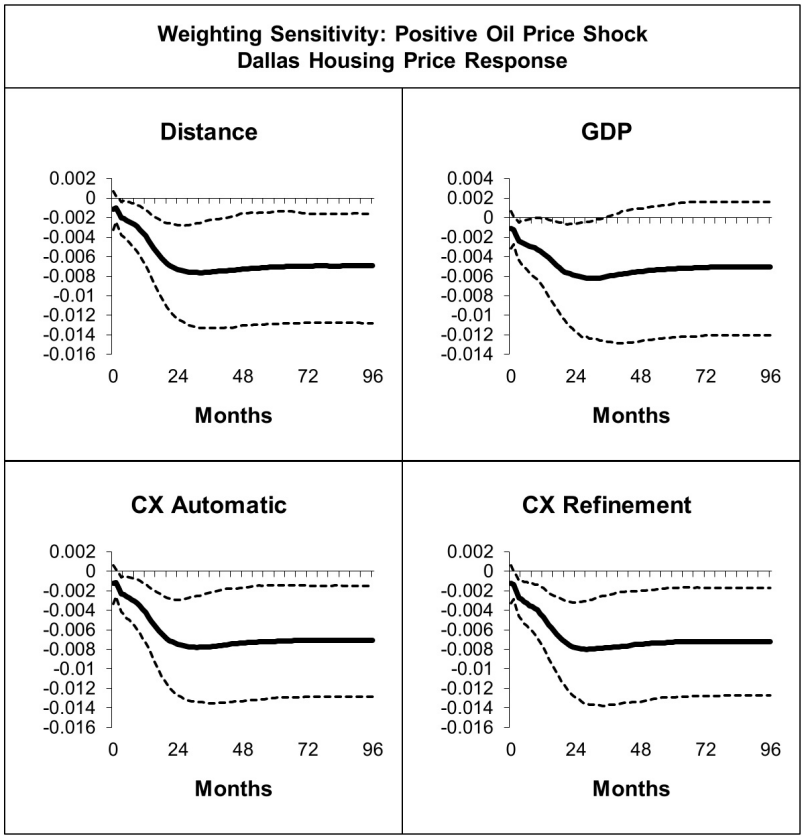

Figure 2.7: GIRFs of a positive (1 s.e.) shock to Oil Prices (bootstrap mean estimates with 90% bootstrap error bounds). Dallas housing price response under each weighting scheme.

Figure 2.7 reports the Dallas housing market response sensitivity across weighting schemes and similarly to the previous shocks, the distance panel and GDP panel directly illustrate a case where the argument of a researcher would be altered just based on the scheme that he chose. Under distance weights, the researcher would say Dallas housing prices are significantly negatively impacted whereas under GDP weights he would not be able to make such a claim. Convex weights and the subsequent optimization procedures provide a straightforward way to clean up the ad hoc

problem in a much more economically characteristic way due to it's basis of maximization. Certainly it could be that responses from the use of convex weights and optimization are still borderline, but in any case, the point is to remove uncertainty and to increase the justification and strength of any argument made from a set of generated results.

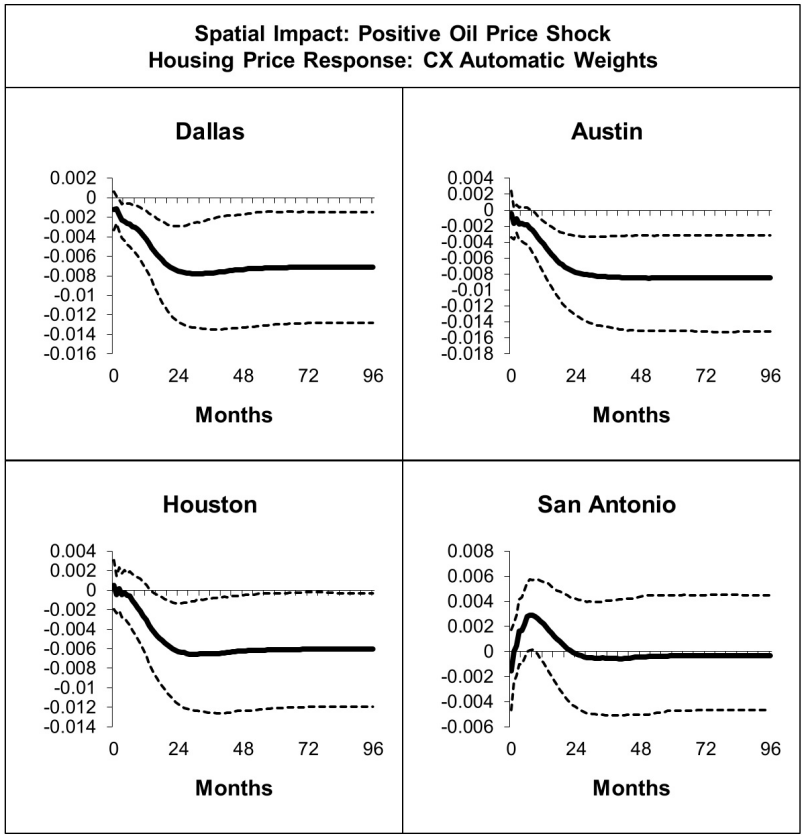

Figure 2.8: GIRFs of a positive (1 s.e.) shock to Oil Prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic optimal model housing price responses.

Similarly to the results of the GDP shock, Figure 2.8 shows a moderate level of response heterogeneity across space from the oil price shock as well. On impact, all metros except for Houston are negatively impacted with effects ranging between -0.04% in Austin to -0.12% in Dallas. Over the first year, San Antonio surprisingly experiences an insignificant positive shock but is otherwise unaffected. Over the same period, Dallas, Houston, and Austin are significantly negatively impacted with effects after twelve months falling around -0.39%. San Antonio is the only metro that appears unaffected. In the long run, the positive oil price shock results in a permanent housing market decline in Dallas, Houston, and Austin which all show negative significant impacts across the horizon.

These findings are in the expected direction, with the exception of San Antonio, particularly in the sense that they can be viewed through the lens of the real GDP response. GDP averages a response of -0.27% through the first year and up to -0.39% for the second and is significantly

negative throughout. Regarding the resilience of San Antonio, one reason could be it's respectively less diversified and larger involvement in oil and gas industry than other areas. It would not come as a surprise that, when the oil price appreciates, places more heavily involved in the extraction and sale of oil would be able to benefit from its appreciation and hence any negative impacts from a contractionary response in the broader market would be offset to a larger extent.

2.4.4 Shock to Short-Term Interest Rates

We now turn attention to a one standard error positive shock to the interest rate, which amounts to a monthly percentage increase of around 10 basis points. In the simplest sense, interest rates directly impact borrowers access to lending and when rates rise, it becomes more difficult to qualify and thus less individuals can secure funding. Ceteris peribus, less individuals with funding means less demand for home ownership and thus we should see a decline in housing prices. In terms of the long run response, this is exactly what we observe across space in the response of housing prices to the positive interest rate shock, save for Austin which appears largely unaffected. Regarding sensitivity however, figure Figure 2.9 reports the Austin housing market response sensitivity to weighting and presents another clear illustration of the case where the responses from each pure weighted model are significantly different. Just as before, if a conclusion were to be drawn about how the housing market in Austin responds to interest rates, the conclusion would be that it is completely resilient if the model were estimated under distance weights but that it is negatively impacted, albeit in a borderline way, if the estimation were under GDP weights. This again illustrates the danger that blind weighting can have and just via visual inspection it is not unreasonable to conclude that blind weighting is equivalent to blind results. The bigger problem with such blind weighting is not the obvious question of whether the response should be interpreted as significant but rather that the question wouldn't even be addressed without considering multiple weighting schemes which of course then begs optimality.

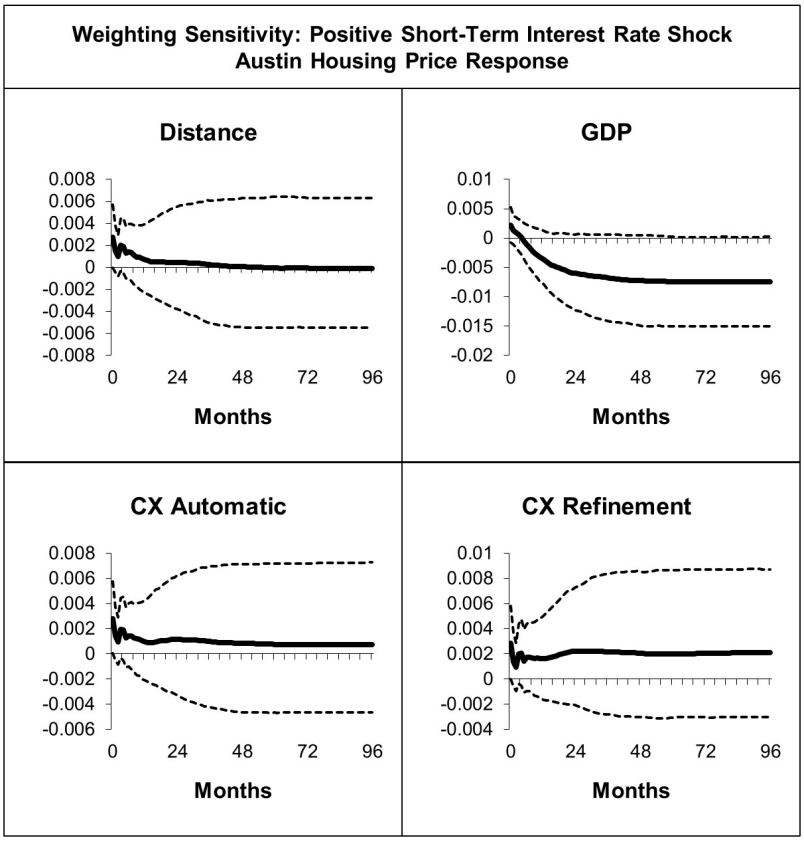

Figure 2.9: GIRFs of a positive (1 s.e.) shock to US Short-Term Interest Rates (bootstrap mean estimates with 90% bootstrap error bounds). Austin housing price response under each weighting scheme.

Regarding heterogeneous transmission, we again observe some differences in the magnitude of the declines across space. While each housing market in Dallas and Houston appears effected only in a borderline way, it is clear that they are effected substantially more than the markets in Austin and San Antonio.

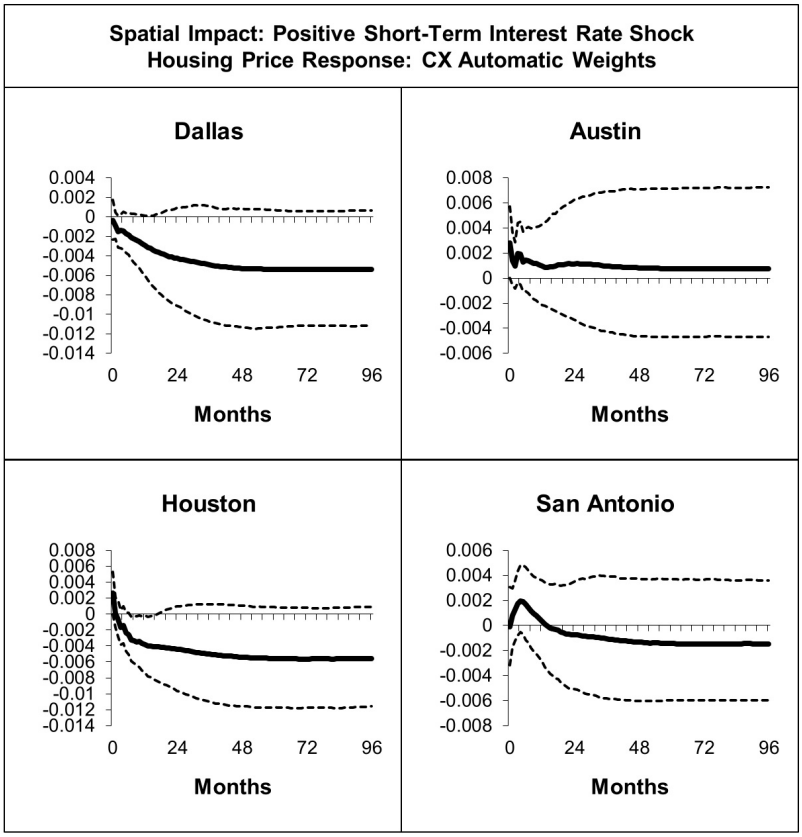

Figure 2.10: GIRFs of a positive (1 s.e.) shock to US Short-Term Interest Rates (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic optimal model housing price responses.

On impact, effects over space are mixed and range range between -0.04% and 0.28%. Over the first twelve months more prominent pattern begins to emerge with housing markets in Dallas and Houston averaging impacts of -0.20% and -0.25% respectively. Interestingly, the average of the point estimates for Austin and San Antonio are both positive over the period but when considering the effect in terms of confidence intervals, it is clear that both metros are not impacted in a way that is significantly different from zero. The impact on the markets in Dallas and Houston however is much closer to significant, especially in through the first twelve months and then borderline across the rest of the horizon.

2.4.5 Domestic Shocks

A variety of global shocks have been considered, all with moderately heterogeneous results falling in the expected directions save for San Antonio. In contrast to global shocks, this section addresses potential spillovers from one metro to another by presenting a one standard error positive shock to housing prices localized solely in Dallas. This housing price shock amounts to a monthly increase in housing prices of around 1.7% in Dallas. First concerning weighting sensitivity, Figure 2.11 presents

the Houston housing market response under each weighting and illustrates again how minor changes in weightings can impact the conclusions that a researcher might reach. In this case, the question is in the duration of significance and in the magnitude of impact. Utilization of distance weights would lead a researcher to conclude that a spillover from the Dallas housing market to the Houston housing market might only last in any significant way for around one quarter and be limited in magnitude to around 31% of the size of the Dallas shock. Utilization of GDP weights on the other hand would lead a researcher to conclude that the spillover might last at least three quarters and be around 106% of the magnitude of the Dallas shock for the first quarter. Again without testing multiple weights, a researcher wouldn't be aware that such differences even exist. To help clarify which effect should be believed, the response from the CX Automatic grid search optimal model has both the duration of significance and the magnitude falling between each distance and GDP weighted model. The CX Automatic optimal model response shows that the housing market in Houston responds positively to the Dallas housing market shock for around 6 months with a magnitude of roughly 65% of the size of the Dallas shock over the first quarter.

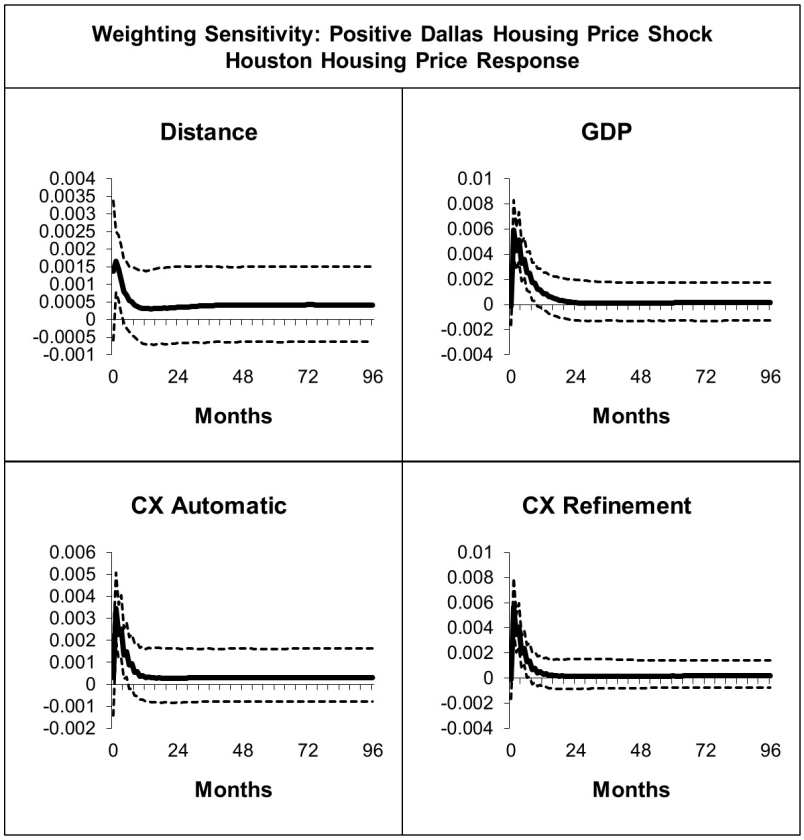

Figure 2.11: GIRFs of a positive (1 s.e.) shock to Dallas housing prices (bootstrap mean estimates with 90% bootstrap error bounds). Houston housing price response under each weighting scheme.

Regarding the broader transmission across space, the other metros in the analysis experience modestly different on impact responses ranging from 0.01% to 0.17%. Over the first year, the

average impacts in Houston and San Antonio are each around 0.12% while the effect in Austin is null. Focusing on the confidence intervals, both Houston and San Antonio are positively impacted in a significant way for around at six months following the onset of the shock whereas Austin seems to be less sensitive and shows an entirely insignificant response.

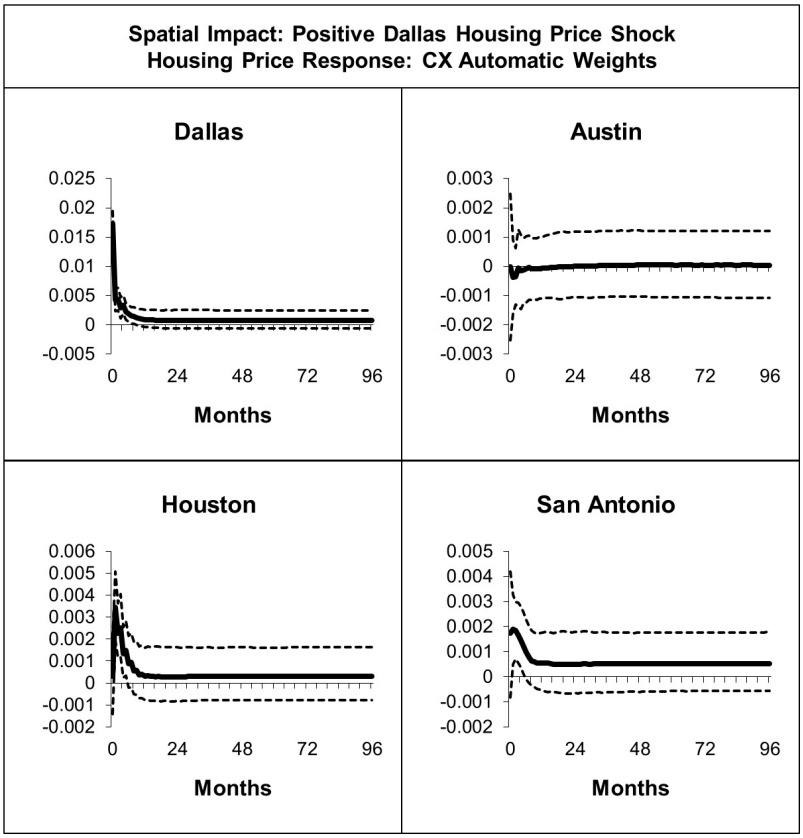

Figure 2.12: GIRFs of a positive (1 s.e.) shock to Dallas housing prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic optimal model housing price responses.

An interesting point of note is the geographical one. Austin lies closer to the origin of the shock than either Houston or San Antonio yet the transmission only appears in a significant way in the latter metros. A potential explanation could be that the prices in Austin's real estate market are substantially higher than the other three metros. Moreover, the Austin market may be meaningfully different enough from the other three metros, perhaps through it's unique positioning in the technology space or through it's building regulations and policies, that the transmission is dampened to the point of insignificance.

2.5 Concluding Remarks

Addressing spatial linkages is an important aspect of modern applied macroeconomics and is central to adequately modelling spillovers and transmission effects. Handling spatial linkages in an imprecise or casual way may yield questionable results, for which extra caution must then be

exercised over their use in forecasting or policy analysis. Interestingly, only a small minority of papers within the GVAR literature have considered the impact of weighting schemes, the primary method for defining such linkages. In all cases presently known to the author, the GVAR literature has yet to address pre-averaging but has instead focused on mixing on variables or post-averaging despite the recent call for further research in the specific area by Martin and Cuaresma (2017). Through the development of convex weighting chapter 1 and the subsequent optimization procedure utilized here, this chapter extends the set of available options for defining spatial linkages in models that handle the curse of dimensionality via compression, such as the GVAR model, and offers a justifiable approach to alleviating uncertainty.

For testing the performance of the newly developed convex weighting scheme, a set of GVAR models on local housing markets are estimated. These models include the four largest metros in Texas, covering roughly 80% of state output, and are estimated over a period of nearly 25 years (from 1990 to 2014). Using these models, forecasting and impulse response sensitivity to weighting scheme choice is specifically investigated. Forecasting performance from models utilizing convexly optimized weights are shown to outperform both naive models but also GVAR models relying on more traditional weighting schemes such as pure distance. Regarding impulse sensitivity, a number of different shocks are passed through the models for which we learn how choices over weighting schemes can critically impact the results and hence shape the arguments that researchers might make. Moreover, by substituting optimization for choice, it is illustrated that convex weighting is well position to solve this complication and hence should allow researchers to be more confident in their findings.

Beyond the technical contribution of this chapter are the substantive impulse response results. In the preceding section, it is shown that Texas housing prices are impacted in moderately heterogeneous ways across space over a set of national level shocks including GDP, equity prices, the oil price, and the short term interest rate. To investigate spillovers and domestic transmission effects to other metros, a localized shock to housing prices only in Dallas is passed through the model. Highlights from the results are as follows: (1) housing markets in Dallas and Houston respond in a highly similar way to all national level shocks with the Austin market being effected at roughly half the magnitude and the San Antonio market being largely unaffected, (2) housing markets in Houston and San Antonio experience a significant short run spillover effect of roughly 50% the magnitude of a Dallas housing price shock while the Austin market experiences no spillover at all.

Overall, the results suggest that convex weight mixing can perform well on smaller universe type questions for which distance weights are popular. A clear path for future research is on the performance of convex weights at the international level where countries are cross-sectional units and where the assumption that trade weights are sufficient has long been widely accepted. This question is addressed in chapter 3.

CHAPTER 3

International Application on the Global Economy

3.1 Introduction

This chapter presents a full-scale GVAR model of the global economy and represents a major extension to the seminal work by Dees et al. (2007) and to the literature on GVAR weighting scheme selection. The primary focus of this chapter is on providing a real-world understanding of how the choice of weighting scheme can impact inferences drawn from impulse response analysis. In this undertaking, a number of further innovations regarding weighting schemes are developed and demonstrated.

Specifically, the Convex Weights developed in chapter 1 are utilized in combination with the existing weighting scheme concept of mixing on variables that was first introduced in Eickmeier and Ng (2015). As a result, this approach leads to a substantial number of possible weightings over which the optimal is determined by employing a full-search algorithm over a time-series cross-validation of in-sample fit[1]. Results from the optimally weighted GVAR model, referred to as the cross-validated convex weights model (CV-CX), are then compared to a pure trade weighted baseline model for which three distinct findings emerge: (a) that significant difference exist between the impulse responses generated from each model, (b) that the bootstrap confidence bands from the CV-CX model are substantially tighter than the baseline model, and (c) that the results of the CV-CX model tell a more coherent story of the global economy.

Regarding the first finding, another contribution of this chapter is in the development of an inference classification algorithm that has the capability to categorize the direction of impulse responses as significant, borderline, or insignificant for a given horizon. Comparison of classified inferences generated from the baseline and CV-CX models shows that, at any horizon, differences exist in no less than 30% of cases[2]. This is a chilling result when considering just how commonly used global macro data is and that over 85% of the global macro GVAR literature has, up to this point, utilized exclusively trade weights[3]. Generally, the evidence put forth in this chapter

[1] A modified version of the GVAR toolbox is utilized to search over 11,011 GVAR models. The modified toolbox as well as the scripts and functions set up for conducting the grid-search and inference exercises presented in later sections are available from the author on request.

[2] Cases are combinations of countries and variables for which there are 123.

[3] As a starting point for analyzing the commonality of variables in the global macro literature, a review of the papers listed in the survey paper Chudik and Pesaran (2016) revealed the following usage among papers utilizing GVAR models with countries as cross-sectional units: output (85%), inflation (63%), equity prices (47%), exchange rates

suggests that, in those circumstances where the linkage mechanism may be in question, a simple re-evaluation under the CV-CX approach may prove fruitful.

As for the second finding, it does not come as a surprise that the CV-CX model results in tighter confidence intervals since it ultimately fits the data better. Such a finding may prove extremely useful for financially oriented GVAR studies where any form of worst-case scenario analysis might be utilized. Additionally, the tightening of confidence intervals as a result of the cross-validated improvement procedure is beneficial to our understanding since it is the confidence intervals by which inference is commonly derived.

Touching on the third finding, the CV-CX model tells a more economically intuitive story of global financial markets. Particularly, in response a negative US equity price shock, the baseline model paints a counter-intuitive picture of equity and bond markets while the CV-CX model results in the expected inverse relationship.

In terms of this chapter's place in the literature, it represents a parallel and an alternative to Gross (2018) who has suggested a fully endogenous method for determining the optimal weighting in GVAR models. The method put forth in this chapter represents an appealing balance between the fully endogenous weighting estimation approach and the specification of a single linkage mechanism that the literature has so heavily relied upon. The mixed-on-variables convex weights approach alleviates the strenuous assumption implicit in the single specification approach while also preserving the economic context that is lost with the fully endogenous method. Specifically, in retaining the economic context of the linkage mechanism, the global model in this chapter revealed that economic variables are best linked by trade while financial variables are best linked by a combination of a county's portfolio investment and it's distance from other countries.

The rest of this chapter is laid out as follows: Section 3.2 presents the model and the standard diagnostics and details the procedure underlying the weighting scheme. Section 3.3 presents the impulse response inference classification algorithm and reports the results of the dynamic simulations and inference sensitivity analysis while Section 3.4 concludes.

3.2 Global Application (1979-2016)

The GVAR models developed in this chapter mirror that of the model in Dees et al. (2007, DdPS) but with expanded temporal coverage. Whereas the DdPS model is estimated over the period 1979(2)-2003(4), the models in this chapter are estimated over an extended period which is 13 years longer now ending at 2016(4). This addition is equivalent to a 52% extension to the total length of the time series. Additionally, the new coverage includes the periods over which

(64%), short-term interest rates (63%), long-term interest rates (41%), oil prices (66%). The same literature is also used to assess the percentage of papers utilizing trade weights.

the financial crisis transpired and thus as an added bonus to the main findings, each intermediate result is compared with it's DdPS equivalent and can be interpreted through the lens of the financial crisis' impact. The GVAR model in this chapter includes 30 countries accounting for roughly 82% of global output as is measured by 2016 Purchasing Power Parity GDP. The coverage here is the same as the DdPS coverage except for the exclusion of the Latin American countries that experienced hyperinflation during the late 1980's[4]. The decision to exclude these countries was made to increase the stability of the model since the overall purpose of this chapter is to provide a sound evaluation of the impact that convex weighting procedures can have at the global level. Table 3.1 presents each of the countries in the model with each under it's respective regional categorization. As was done in DdPS, the 8 included European countries that joined the European Union in 1999 are aggregated together and modeled as a single cross-sectional "Euro" unit.

Table 3.1: Countries and Regions in the GVAR Model

USA	**Euro Area**	**Rest of W. Europe**
China	Austria	Norway
Japan	Belgium	Sweden
UK	Finland	Switzerland
	France	
Other developed economies	Germany	
Australia	Italy	
Canada	Netherlands	
New Zealand	Spain	
Rest of Asia	**Latin America**	**Rest of the World**
Indonesia	Chile	India
Korea	Mexico	Saudi Arabia
Malaysia		South Africa
Philippines		Turkey
Singapore		
Thailand		

[4] Specifically Argentina, Brazil, and Peru are excluded from this chapter's analysis. Collectively, these countries account for less than 4% of the total coverage (2016 PPP-GDP) of those included in the model. As is documented in Kiguel and Liviatán (1995), annual inflation peaked at around 3100%, 2900%, and 7500% for each Argentina, Brazil, and Peru respectively between 1989-1990.

Concerning the variables in model, the standard macroeconomic collection is included. More specifically,

$$y_{it} = \ln(GDP_{it}/CPI_{it}), \qquad p_{it} = \ln(CPI_{it}), \qquad q_{it} = \ln(EQ_{it}/CPI_{it}),$$
$$e_{it} = \ln(E_{it}), \qquad \rho^S_{it} = 0.25 * \ln(1 + R^S_{it}/100), \qquad \rho^L_{it} = 0.25 * \ln(1 + R^L_{it}/100)$$

where the variables are defined exactly as in DdPS. That is, GDP_{it} is nominal Gross Domestic Product, CPI_{it} is the consumer price index, EQ_{it} is the nominal equity price index, E_{it} is the exchange rate in terms of US dollars, R^S_{it} is the short-term interest rate, and R^L_{it} is the long-term interest rate. All data is from the 2016 Vintage release of the GVAR database compiled by Mohaddes and Raissi (2018).

In the DdPS model, country-specific foreign variables, y^*_{it}, π^*_{it}, q^*_{it}, ρ^{*S}_{it}, ρ^{*L}_{it} were constructed exclusively from trade weights. This chapter shares the same weighting scheme in it's baseline model that utilizes fixed trade weights computed as the average of trade flows over three years 2012-2014. Justification for the use of fixed weights is twofold. First, in the context of these exact variables, it has been shown that the correlation between country-specific foreign variables constructed from fixed and time-varying weights is so high that there would be practically no loss from utilizing the fixed specification[5]. Second, because this paper expands on the ideas of Martin and Cuaresma (2017) who investigate a wide range of linkages constructed from bilateral financial flows data that only have available data extending back to around 2005.

In total, the models in this chapter include 23 cross-sectional units[6]. With regard to model specification, the DdPS specification is followed exactly. Specifically, each country model except for the US includes all six country-specific foreign variables, y^*_{it}, π^*_{it}, q^*_{it}, ρ^{*S}_{it}, ρ^{*L}_{it} an the log of oil prices (p^o_t), as weakly exogenous while the US model excludes foreign equity prices and interest rates.

3.2.1 Weighting Schemes

The primary focus of this chapter is on the sensitivity of inference to weighting schemes in the global setting. To investigate this while managing complexity, only the best performing weighting schemes that have been utilized in the literature are included. Specifically, the most comprehensive study on weighting scheme sensitivity to date, Martin and Cuaresma (2017), test the following nine schemes:

- Bilateral trade flows
- Inward portfolio investment

[5] See DdPS table VIII for exact details. The average correlation for all focal foreign variables is 0.96 in levels and is .89 in first differences. These correlations were so high that the authors of DdPS determined that the utilization of fixed weights was justifiable.

[6] Models include 22 individual countries as well as the Euro unit which is made up of a 2012-2014 PPP-GDP weighted average of the 8 European countries specified in Table 3.1

- Outward portfolio investment
- Inward foreign direct investment
- Outward foreign direct investment
- Inward banking claims
- Outward banking claims
- Trade costs
- Geodesic Distance.

In their analysis of pre-crisis forecasting performance, models utilizing banking claims and trade costs were never the winning model and thus those are excluded from the analysis here[7]. As for the remaining schemes, I compress the inward and outward versions of foreign direct investment and portfolio investment respectively by elementwise averaging. Justification for this is that cross-correlation coefficients between inward and outward matrices are very high[8]. After throwing out banking claims and trade costs for poor performance in Martin and Cuaresma (2017) and combining inward and outward matrices for similarity, we are left with four distinct weighting schemes.

Regarding the details of each weighting scheme used in this chapter, trade weights are constructed as the average of the import and export annual figures provided by the IMF Direction of Trade Statistics. The data for 2012-2014 exports and imports is averaged and then used to construct the trade weights. Bilateral portfolio investment data are obtained from Table 8 of the IMF Coordinated Portfolio Investment Survey (CPIS) and bilateral foreign direct investment data is from Table 6.1-o of the IMF Foreign Direct Investment Survey (CDIS). Portfolio investment and foreign direct investment data over the period 2012-2014 is averaged in the construction of these respective weighting schemes similarly to how trade weights are constructed. For the same reasons stated in Martin and Cuaresma (2017), fixed weights are used to avoid time related bias when comparing the performance of specifications. As for the fourth weighting scheme, geodesic distances have been derived form CEPIIs GeoDist dataset. Lastly, I construct a fifth weighting scheme as a global vector weighting of GDP[9]. Cross-correlations between each of these five matrices are presented in Table 3.2. This table reveals that the maximum similarity between matrices is between trade and foreign direct investment (0.74) and the minimum is between distance and GDP (0.05) with all others falling in-between. This collection of matrices provides a reasonable balance between complexity and variation.

[7] See Martin and Cuaresma (2017) Table 4 for details

[8] Cross-correlations coefficients are computed as $corr(vec(A),vec(B))$ where A and B are matrices.

[9] See Section 1.3 for the description of how "global vector weights" are constructed.

Table 3.2: Weighting Scheme Correlation Matrix

	trade	fdi	port	dist	gdp
trade	1.00				
fdi	0.74	1.00			
port	0.56	0.58	1.00		
dist	0.51	0.38	0.22	1.00	
gdp	0.65	0.61	0.73	0.05	1.00

Note: Entries represent correlation coefficients between weighting schemes constructed from bilateral trade flows, foreign direct investment, portfolio investment, geodesic distance, and PPP GDP respectively.

A key component of this chapter is the concept of mixing on variables. This concept was introduced by Eickmeier and Ng (2015) and further extended by Martin and Cuaresma (2017). Mixing on variables refers to the usage of different weighting schemes for different variables in the GVAR context. As the number of variables included increases, obviously so also does the number of combinations. Put more elegantly, the concept of mixing on variables can be expressed as a k-permutation with repetition for which the number of possible k-permutations with repetition of n objects is

$$P'_{n,k} = n^k$$

where, in this case, n is the number of matrices and k is the number of variables in the GVAR model. Considering the standard macroeconomic variable mix in combination with just two matrices, the number of permutations is equal to 64. In the case of Martin and Cuaresma (2017) who use all six variables but consider nine matrices, the number exceeds half a million[10]. To combat this level of complexity, researchers have made use of categorizations and have restricted the scope of mixes. By categorizing variables as either economic or financial, the number of permutations can be drastically reduced. The literature has tended to categorize GDP and inflation as economic while assigning equity prices, interest rates, and exchange rates as financial. Under this categorization, the number of permutations is now limited to the square of the number of matrices. This reduction however may still result in an undesirably high (81) number of permutations as was the case in Martin and Cuaresma (2017) who further reduce the number by only considering mixes with trade weights on the economic variables which ultimately gave them 15 mixes. In this chapter, I categorize the variables in the same way but because we are interested in combining the concept of mixing on variables with the concept of convex weighting, the number of matrices when viewed through the

[10] The number of possible k-permutations for nine matrices and six variables in the case of Martin and Cuaresma (2017) is $P'_{9,6} = 6^9 = 531{,}441$

lens of mixing on variables is actually equal to six since we add an additional 'convex' matrix that allows for any combination of the five pure weight matrices. The mixes on variables are shown in Table 3.3. The number of permutations for this scenario is equal to $P'_{6,2} = 36$ but since each convex entries houses all of the pure matrices, the full set of permutations can be illustrated in fewer rows.

Table 3.3: Weighting Scheme Concept: Mix on Variables

Model	Economic vars.	Financial vars.
01 (Baseline)	Trade	Trade
02	Trade	Convex
03	FDI	Convex
04	Port	Convex
05	Dist	Convex
06	GDP	Convex
07	Convex	Trade
08	Convex	FDI
09	Convex	Port
10	Convex	Dist
11	Convex	GDP
12	Convex	Convex

Note: Rows illustrate the different combinations for mixing on variables. Economic variables are GDP and Inflation and Financial variables are Exchange Rates, Equity Prices, and Interest Rates (short and long).

To understand this more clearly, consider that each convex entry in Table 3.3 actually represents a convex hull rather than a single matrix. Abstracting from the element-wise notation in Equation (1.12), consider a convex weightings in terms of the constrained convex hull

$$\text{Conv}(W) = \left\{ \sum_{\rho=1}^{|W|} \theta_\rho W_\rho \middle| (\forall \rho : \theta_\rho \in \Theta) \wedge \sum_{\rho=1}^{|W|} \theta_\rho = 1 \right\}$$

such that:

$$W = \{\text{Trade, FDI, Port, Dist, GDP}\}$$
$$\iota = 0.1$$
$$\Theta = \{0\iota, 1\iota, 2\iota, \ldots, n\iota\} \text{ with } n\iota = 1$$

where θ are convexity parameters, W are weighting matrices, and $|W|$ represents the cardinality of the set of weighting matrices. The first condition specifies that the weighting matrices are the ones that are included in this chapter. The second defines the interval size and the third specifies that the weights must conform to the interval size. These three conditions together characterize the full set of convex matrices for which a grid search can be conducted to test each unique combination on the grounds of a selection criterion. In this chapter, the selection criterion is the sum of squared

errors over the entire GVAR model. Another factor to consider is lag length selection for which, in this chapter, only AIC is considered so as to mirror Dees et al. (2007).

Table 3.4: Weighting Scheme Grid Search Results

Model	Trade	FDI	Port	Dist	GDP	SSE
01-trade-trade	0	0	0	0	0	36.56553
02-trade-convex	0	0	0.4	0	0.6	**33.41677**
03-fdi-convex	0	0.3	0.7	0	0	37.84087
04-port-convex	0	0.3	0	0	0.7	38.11174
05-dist-convex	0	0	0	0	1	35.14017
06-gdp-convex	0	0	0.6	0	0.4	34.24129
07-convex-trade	0.4	0	0	0.1	0.5	35.89563
08-convex-fdi	0.6	0	0	0	0.4	34.57329
09-convex-port	0.7	0	0	0	0.3	33.64182
10-convex-dist	0	0	0	1	0	44.34379
11-convex-gdp	0.7	0.1	0	0	0.2	33.48794
12-convex-convex	0.3	0	0.1	0	0.6	34.39359

Note: SSE is the error sum of squares obtained as the sum of the squared residuals from each GVAR model. Bold value represents minimum. Model names are encoded as follows:
number - econ. var. weights - fin. var. weights

A full search of the convex hull under each mix on variables combination in Table 3.3 is conducted. The results from this procedure are presented in Table 3.4 which reports the best performing combinations of each mix along with it's respective weighting coefficients and the overall sum of squared errors (SSE). From the above results, the winning model in terms of overall in-sample fit is the model with trade weights on economic variables, convex weights being made up of 40% portfolio investment and 60% GDP on financial variables, and using AIC to determine lag length. The fit of this model is around 9% better than the baseline pure trade weighted AIC model that is estimated in DdPS. A point of critical importance when selecting models based on in-sample metrics is how robust the selection is. To check the robustness, I conduct a form of time series cross-validation whereby I incrementally move the end of the sample backwards in time. The results of this procedure are presented graphically in Figure 3.1 where it can be seen that the performance of the winning model is consistent. In results not reported, the convexity parameter values are also surprisingly consistent over the cross-validation[11]. In the results and discussion that follows, references to the cross-validated convex model (CV-CX) are regarding the best fit combination of mixing on variables and convex weighting; specifically the 02-trade-convex model.

[11] In all periods except for the last, the parameters on the 02-trade-convex model were 0.5 port and 0.5 gdp. The parameters of the last period are found in Table 3.3.

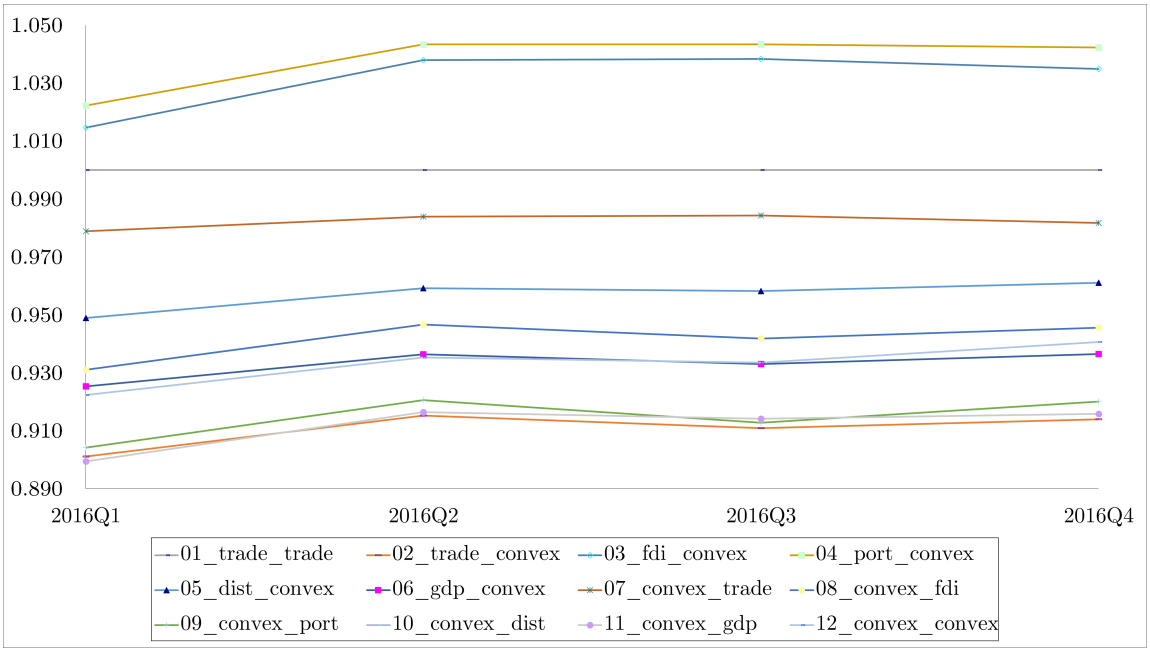

Figure 3.1: Time series cross-validation of relative performance of models across samples with different ending periods. Relative performance is the ratio of the select model SSE to the baseline model SSE.

Table 3.5 provides a tabular representation of the actual baseline pure trade weighting matrix for the same eight focal economies (seven countries plus the euro area itself, composed of eight countries) as DdPS, with a Rest category showing the trade shares with the remaining 7 countries in our sample. Note that since the table is row standardized (rows sum to unity), each row represents a country and each column represents the share of weight a corresponding country gets.

First considering the USA, we can see that the China accounts for the slightly over 18% of trade while the included Euro area countries account for under 15%. Comparing this to the older trade weights from the averaged period 2001-2003 that are reported in DdPS, the share of trade with China has grown substantially from just over 7% at the expense of trade with Japan that is now down to only 7% from over 12%.Trade with the Euro area countries has only minimally declined from 15.5% to its current level. Looking at the last column and comparing to the earlier weights, the trade share of focal countries for the US has picked up slightly.

Regarding the Euro area, trade with the US has fallen substantially over the last decade from nearly 23% to now only around 15%. The Euro area's trade share with China however has grown from only around 5.5% to more recently over 16%. As for the other focal countries, a similar pattern can be observed. China has gained trade share across the board, in most cases having more than doubled, at the cost of lessened trade from the Euro area and the USA.

Table 3.5: Trade Weights Based on Direction of Trade Statistics

Country/Region	USA	Euro area	China	Japan	UK	Rest of Europe			Rest[a]
						Sweden	Switz.	Norway	
USA	0.000	0.146	0.181	0.070	0.037	0.006	0.017	0.004	0.540
Euro area	0.152	0.000	0.164	0.043	0.188	0.057	0.098	0.038	0.259
China	0.198	0.173	0.000	0.144	0.026	0.006	0.012	0.003	0.439
Japan	0.168	0.097	0.268	0.000	0.018	0.003	0.012	0.003	0.432
UK	0.110	0.512	0.076	0.022	0.000	0.028	0.051	0.045	0.156
Sweden	0.059	0.550	0.054	0.016	0.096	0.000	0.013	0.123	0.089
Switz.	0.098	0.654	0.051	0.033	0.053	0.009	0.000	0.004	0.097
Norway	0.065	0.429	0.052	0.018	0.233	0.108	0.011	0.000	0.084

Note: Trade weights are computed as shares of exports and imports displayed in rows by region such that a row, but not a column, sums to one. [a] 'Rest' gathers the remaining countries. The complete trade matrix used in the baseline GVAR model is given in Appendix B.0. Source: Direction of Trade Statistics, 2012 2014, IMF.

Comparing the above trade weights to the convex weights reported in Table 3.6, fairly substantial differences are evident. The convex weights in Table 3.6 are calculated as the elementwise combination of the portfolio investment weighting matrix and the geodesic distance weighting matrix with coefficient values 0.6 and 0.4 respectively. This particular convex weighting scheme applied to the financial variables in combination with pure trade weights on economic variables resulted in the lowest SSE of any mix on variables or convex weighting[12]

Inspecting the US in the first row, the most evident major change is that China now receives a far smaller weight under the convex scheme than under trade weights. Specifically, China's share of the US under the convex weighting is only 11% whereas it was 18% under pure trade which equates to china's impact being lessened by roughly 40%[13]. This roughly 40% reduction in China's importance holds across all countries as the average share for China under pure trade weights is 15% but only around 9% under convex weights[14]. As for the other focal countries pertaining to the US, under convex weights Japan, the UK, and the other non-EU European countries have all gained in importance. Lastly, regarding the rightmost column for the US, it can be seen that the rest of the world now only accounts for just over 26% which is down from over 50%. This alone is an interesting finding in that particularly in terms of foreign exchange rates, the rest of the world appears to matter less than trade weighting would suggest for the US[15]. This isnt that surprising

[12] The weighting scheme in Table 3.6 is selected from Table 3.4 as the row with the minimum SSE (3rd row).

[13] Note that in this discussion of magnified or lessened importance, any mentioned change is only in reference to the impact of country-specific foreign financial variables. This is due to the selected mix on variables which specifies trade weights for economic variables and convex weights for financial variables.

[14] See Appendix B.0 for the complete weighting matrices.

[15] Recall though that the only US-specific foreign variables included in the model are GDP, Inflation, and the Exchange Rate of which only the latter is classified as financial. Hence, the convex weighting for the US model will impact the total SSE less than other countries that have the full set of foreign variables.

though because one would expect the major countries to matter more in terms of exchange rates.

Table 3.6: Convex Weights Based on Portfolio Investment and Distance

Country/Region	USA	Euro area	China	Japan	UK	Rest of Europe			Rest[a]
						Sweden	Switz.	Norway	
USA	0.000	0.250	0.118	0.190	0.113	0.017	0.024	0.025	0.263
Euro area	0.334	0.000	0.108	0.144	0.144	0.020	0.038	0.031	0.181
China	0.350	0.182	0.000	0.090	0.082	0.009	0.012	0.015	0.262
Japan	0.385	0.203	0.096	0.000	0.104	0.010	0.016	0.021	0.165
UK	0.341	0.244	0.090	0.101	0.000	0.013	0.018	0.022	0.169
Sweden	0.271	0.277	0.087	0.104	0.060	0.000	0.026	0.037	0.137
Switz.	0.379	0.194	0.087	0.079	0.073	0.017	0.000	0.025	0.145
Norway	0.269	0.240	0.087	0.133	0.072	0.040	0.023	0.000	0.136

Note: Convex weights are computed as the linear combination of portfolio investment and geodesic distance (coefficient values given in Table 3.4) and are displayed in rows by region such that a row, but not a column, sums to one. [a] 'Rest' gathers the remaining countries. The complete convex matrix used in the convex optimal GVAR model is given in Appendix B.0. Sources: Distance from CEPIIs GeoDist database; Portfolio Investment from International Financial Statistics (IFS) Table 8, 2012-2014, IMF

Concerning the country shares for the Euro Area, a similar set of changes is observed from pure trade to convex weights. First, the share attributed to the US has more than doubled under the convex weights to more than 33% which highlights the extent to which the major financial markets are linked and is indicative that purely trade based weights don't fully capture the relationships between the major economies in terms of financial markets which are now almost exclusively digital. Other changes from trade to convex weights include the drastic over threefold increase in the share attributed to Japan which under convex weights is over 14%. The shares from China, the UK, and the other non-EU European countries having declined fairly substantially.

Other focal countries experience similar changes primarily with the United States and Euro Area receiving greater shares than under trade weights and China receiving less. Overall, apart from any individual differences, the information within the previous two tables suggests that trade does not properly characterize the cross country linkages between financial markets and that the fit can be improved by utilizing both concepts of mixing on variables and convex weight mixing.

3.2.2 Unit Root Tests

Following the standard techniques of the GVAR literature and what is done in chapter 2, it is necessary to test the integration properties of the variables in the GVAR. To carry this out, Park and Fuller (1995) type weighted symmetric ADF tests are implemented for which, in accordence with the $I(1)$ assumption in Pesaran et al. (2004), we would like to see all variables fail to reject

the null in levels and reject the null in first differences, indicating that all variables are $I(1)$[16].

Table 3.7: Unit Root Test Statistics for Domestic Variables

Domestic	Country/Region							
Variables	U.S.	E.A.	China	Japan	U.K.	Sweden	Switz.	Norway
y	-1.63	-1.23	-2.46	-0.40	-2.36	-3.19	-3.34	-0.65
Δy	-4.98	-4.83	-3.84	-5.75	-3.71	-5.16	-4.59	-6.13
Δp	-1.47	-1.92	-3.33	-3.18	-2.38	-3.79	-5.53	-3.03
$\Delta^2 p$	-10.57	-8.87	-7.75	-8.21	-8.86	-7.50	-8.50	-8.95
eq	-2.18	-2.44		-1.94	-1.83	-2.84	-2.07	-3.33
Δeq	-7.20	-5.74		-7.79	-8.14	-7.73	-7.39	-6.21
ep		-2.19	-1.11	-1.87	-2.21	-2.43	-2.60	-2.32
Δep		-7.86	-7.73	-5.64	-6.24	-7.70	-8.44	-8.13
r	-1.33	-1.25	-1.44	-1.84	-1.13	-1.29	-1.98	-1.43
Δr	-4.13	-4.39	-6.71	-5.46	-7.33	-8.77	-5.51	-9.26
lr	-1.30	-0.65		-0.54	-0.08	-0.36	-0.95	-0.79
Δlr	-6.48	-5.71		-6.08	-8.70	-7.63	-6.63	-8.43

Note: The WS statistics are based on univariate AR(p) specifications in the level of the variables with p$\leq$ 5, and the statistics for the level and first differences of the variables are all computed based on the same sample period, namely, 1979Q2-2016Q4. The WS statistics for all level variables are based on regressions including a linear trend, except for the interest rate variables. The 95% critical value of the WS test for a regression with a linear trend is -3.24, and for a regression with an intercept only is -2.55. The unit root test statistics for all the countries are available upon request.

Table 3.7 reports WS statistics for the domestic variables in the focal countries for which the only specific issues lie with inflation. Where Δp was $I(1)$ for all countries in DdPS, there are several countries for which it doesn't appear to be $I(1)$, but fortunately it is for the larger focal countries. Generally though, for the domestic variables the test results broadly support the unit root hypothesis.

Since this chapter estimates both a baseline and a CV-CX model, it is important to check the integration properties of both the baseline trade weighted country-specific foreign variables and their convex weighted counterparts. Table 3.8 reports the test statistics for the country-specific foreign variables computed via trade weights. These tests reveal that all variables excluding several country-specific foreign inflation variables are $I(1)$. The country specific inflation variables that appear to be borderline $I(0)/I(1)$ are the US and Japan. Again, generally the test results for country-specific foreign variables generated from trade weights support the unit root hypothesis[17].

[16] The null hypothesis of the ADF tests is that there is a unit root. Statistics smaller in absolute value than 3.24 for variables with a trend (y, Δp, eq, qp) or 2.55 for those without a trend (r, lr) signify rejection. Failing to reject variables in levels and rejecting in first differences means that variables become stationary in first differences and are thus $I(1)$. Also note that inflation is already in first differences. This is because p was $I(2)$ in DdPS and in order to make proper comparisons, I stick to the specification used there.

[17] Regarding inflation, comparing the unit root results for all countries in this chapter to those reported in DdPS reveals that any issues encountered are shared. Updating the sample to 2016 does not appear to change the integration

Table 3.8: Baseline Model Unit Root Test Statistics for Foreign Variables

Domestic	Country/Region							
Variables	U.S.	E.A.	China	Japan	U.K.	Sweden	Switz.	Norway
y	-2.49	-2.66	-1.04	-1.76	-1.46	-1.36	-1.35	-1.87
Δy	-6.18	-5.54	-6.00	-5.11	-5.35	-5.36	-5.09	-5.06
Δp	-3.58	-3.04	-1.80	-3.41	-2.40	-2.56	-1.97	-2.08
$\Delta^2 p$	-9.29	-7.62	-8.87	-7.94	-8.01	-8.09	-7.85	-7.86
eq	-2.99	-2.34	-2.59	-2.61	-2.39	-2.30	-2.27	-2.21
Δeq	-8.11	-7.97	-7.95	-8.12	-8.03	-8.13	-8.00	-7.96
ep	-2.14	-2.64	-2.11	-2.20	-2.29	-2.27	-2.21	-2.25
Δep	-8.08	-8.24	-7.86	-7.40	-7.94	-8.02	-7.99	-8.02
r	-1.18	-1.12	-0.87	-0.97	-1.07	-0.84	-1.02	-1.09
Δr	-6.75	-5.74	-5.56	-6.41	-5.14	-6.91	-5.20	-4.97
lr	-0.51	-0.34	-0.22	-0.22	-0.55	-0.46	-0.53	-0.29
Δlr	-5.70	-6.59	-6.44	-6.83	-6.04	-6.00	-5.96	-6.26

Note: The WS statistics are based on univariate AR(p) specifications in the level of the variables with p$\leq$ 5, and the statistics for the level and first differences of the variables are all computed based on the same sample period, namely, 1979Q2-2016Q4. The WS statistics for all level variables are based on regressions including a linear trend, except for the interest rate variables. The 95% critical value of the WS test for a regression with a linear trend is -3.24, and for a regression with an intercept only is -2.55. The unit root test statistics for all the countries are available upon request.

Lastly, Table 3.9 reports the unit root test statistics for the country-specific foreign variables generated with the cross validated convex model. Note that only the financial variables are included in Table 3.9 because the CV-CX model mixes on variables such that convex weights are only applied to the foreign financial variables. Hence, the foreign economic variables are still generated with trade weights and will yield the same results as the previous table because they are unchanged. In terms of integration, all financial variables are $I(1)$ under convex weights.

properties for any variables, including inflation.

Table 3.9: CV-CX Model Unit Root Test Statistics for Foreign Variables

Domestic	Country/Region							
Variables	U.S.	E.A.	China	Japan	U.K.	Sweden	Switz.	Norway
eq	-2.43	-2.34	-2.42	-2.26	-2.38	-2.31	-2.31	-2.30
Δeq	-8.01	-7.76	-7.93	-7.75	-7.80	-7.83	-7.76	-7.79
ep	-2.14	-2.09	-1.89	-2.50	-2.21	-2.21	-2.29	-2.16
Δep	-8.36	-8.44	-8.17	-8.11	-8.20	-8.31	-8.24	-8.31
r	-0.89	-1.38	-1.00	-0.94	-0.98	-0.90	-1.00	-1.06
Δr	-5.17	-4.43	-4.24	-5.33	-5.04	-5.08	-4.97	-4.60
lr	-0.36	-0.52	-0.55	-0.62	-0.63	-0.53	-0.60	-0.50
Δlr	-5.83	-6.39	-6.05	-6.22	-6.12	-6.10	-6.07	-6.13

Note: The WS statistics are based on univariate AR(p) specifications in the level of the variables with p$\leq$ 5, and the statistics for the level and first differences of the variables are all computed based on the same sample period, namely, 1979Q2-2016Q4. The WS statistics for all level variables are based on regressions including a linear trend, except for the interest rate variables. The 95% critical value of the WS test for a regression with a linear trend is -3.24, and for a regression with an intercept only is -2.55. The unit root test statistics for all the countries are available upon request.

Overall, there is support for the unit root hypothesis among domestic variables and foreign variables under both model weightings (baseline and CV-CX). Due to this, the ability to distinguish between short-run and long-run relations and interpret the long-run relations as cointegrating remains intact.

3.2.3 Specification and Estimation of the Country-Specific Models

I begin the modelling exercise under the two core assumptions of the literature. First, that country-specific foreign variables are weakly exogenous $I(1)$ variables, and second that the parameters of the individual models are stable over time. Each of these assumptions are considered formally in turn in sections that follow. Also of note, in following what has been presented thus far in this chapter, since two models are estimated, results for each are presented. The first model which represents the baseline is estimated under pure trade weights and is the 01-trade-trade model from Table 3.4. The second and more interesting model is the CV-CX model that corresponds to the 02-trade-convex specification of Table 3.4. In each model, the variable selection is the same and is matched to the original DdPS specification while the lag orders and ranks are allowed to differ due to being determined by selection the AIC selection criteria and formal statistical tests respectively.

Specifically regarding model specification, individual country models are allowed to have dif-

fering variables[18]. All country-specific models except for the US contain y, Δp, eq, ep, ρ^S, ρ^L[19]. Additionally, they include foreign aggregates of these variables and the oil price as weakly exogenous. As for the US model, due to its status as a globally dominant economy, the oil price is included as endogenous to allow the evolution of global macroeconomic variables to influence it. Also, the US model excludes eq^*, ρ^{S*}, and ρ^{L*} as they are unlikely to significantly impact the financial sector of the US. Lastly, ep is excluded because the dollar exchange rate is determined outside of the US model. However, the US model includes both y^* and Δp^*, as in DdPS, so as to allow global feedback effects to impact the US.

After specification, VAR models are estimated and their rank is determined. The order of the individual country VARX*(p_i,q_i) models is determined by AIC where p_i is the lag order on domestic variables and q_i is the lag order on foreign variables. Note that p_{max} is set to 2 while q_{max} is fixed to one. After the selection of lag orders, the number of cointegrating relations for each individual country specific VARX* model is determined by Johansens trace statistics as set out in Pesaran et al. (2000) for models with exogenous $I(1)$ regressors for the case of unrestricted intercepts and restricted trend coefficients (case IV).

Table 3.10: Baseline Model VARX* Order and Number of Cointegration Relationships in the Country-Specific Models

	VARX*(p_i , q_i)		# Cointegrating relationships
Country	p_i	q_i	
USA	2	1	2
Euro area	2	1	1
China	1	1	2
Japan	2	1	2
UK	2	1	1
Sweden	2	1	3
Switzerland	2	1	2
Norway	2	1	2

Note: The rank of the cointegrating space for each country is computed using Johansen's trace and maximal eigenvalue statistics as set out in Pesaran et al. (2000) for models with weakly exogenous $I(1)$ regressors, in the case where unrestricted constants and restricted trend coefficients are included in the individual country error correction models.

The orders of the VARX* models and the number of cointegration relationships for the focal countries are reported for the baseline model in Table 3.10 and for the CV-CX model in Table 3.11.

[18] This is due to data availability, particularly with long-term interest rates. However, the primary reason is to match as close to the DdPS specification as possible.

[19] Owing to data availability, the following countries exclude ρ^L: China, Chile, Indonesia, Malaysia, Mexico, Philippines, Saudi Arabia, Singapore, Thailand, and Turkey. Furthermore, eq is excluded from China, Indonesia, Mexico, Saudi Arabia, and Turkey and ρ^S is excluded from Saudi Arabia all for the same reason.

A VARX*(2,1) is selected for all baseline models except Canada, China, and Malaysia for which a VARX*(1,1) was selected. In the CV-CX model, the same orders are found except for Canada which had a VARX*(2,1). As for the number of cointegrating relations, there are more differences between the baseline and CV-CX models. However, the USA, Euro area, and China had 2, 1, and 2 cointegrating relations respectively in both cases. For the other focal countries though, Japan, the UK, and Switzerland had one fewer cointegrating relation under CV-CX than under the baseline and Sweden had one more.

Table 3.11: CV-CX Model VARX* Order and Number of Cointegration Relationships in the Country-Specific Models

	VARX*(p_i , q_i)		# Cointegrating relationships
Country	p_i	q_i	
USA	2	1	2
Euro area	2	1	1
China	1	1	2
Japan	2	1	3
UK	2	1	2
Sweden	2	1	1
Switzerland	2	1	3
Norway	2	1	2

Note: The rank of the cointegrating space for each country is computed using Johansen's trace and maximal eigenvalue statistics as set out in Pesaran et al. (2000) for models with weakly exogenous $I(1)$ regressors, in the case where unrestricted constants and restricted trend coefficients are included in the individual country error correction models.

3.2.4 Testing Weak Exogeneity

As was mentioned earlier, the core assumption of the GVAR modeling strategy is that the country-specific foreign variables are weakly exogenous[20]. To formally test this assumption, this chapter follows the same procedure as the prior application by estimating the Johansen (1992) and Harbo et al. (1998) style auxiliary equations for the country-specific foreign variables, $\mathbf{x}^*_{it}$. Specifically, for each lth element of $\mathbf{x}^*_{it}$, the following regression is estimated:

$$\Delta x^*_{it,l} = \mu_{it} + \sum_{j=1}^{r_i} \gamma_{ij,l} ECM^j_{i,t-1} + \sum_{k=1}^{s_i} \varphi'_{ik,l} \Delta \mathbf{x}_{i,t-k} + \sum_{m=1}^{n_i} \vartheta'_{im,l} \Delta \tilde{\mathbf{x}}^*_{i,t-m} + \epsilon_{it,l}$$

where $ECM^j_{i,t-1}$, j = 1,2,...,r_i are the estimated error correction terms corresponding to the r_i cointegrating relations found for the ith country model and $\Delta \tilde{\mathbf{x}}^*_{it} = (\Delta \mathbf{x}'^*_{it}, \Delta(e^*_{it} - p^*_{it}), \Delta p^o_t)'$. The

[20] Beyond country-specific foreign variables, the oil price variable that is included in every individual country model is treated as weakly exogenous in all models except for the US model. In results not reported, the oil price is weakly exogenous in all cases.

test for weak exogeneity is an F-test of the joint hypothesis that $\gamma_{ij,l} = 0$, $j = 1,2,...,r_i$ in the above regression. Due to the arguments made in DdPS and for simplicity, I fix the lag orders s_i and n_i to be the same as p_i and q_i of the underlying country-specific VARX* models. The test is carried out separately for the baseline and for the CV-CX model, of which the results are reported in Table 3.12 and Table 3.13 respectively.

Table 3.12: Baseline Model F-statistics for Testing Weak Exogeneity of Foreign Variables

Country		Foreign Variables						
		y^*	Δp^*	q^*	ρ^{*S}	ρ^{*L}	p^o	$e^* - p^*$
USA	F(2,136)	0.59	3.46†					0.57
Euro area	F(1,134)	1.91	1.52	1.32	0.19	0.80	0.01	
China	F(2,135)	0.31	0.65	0.18	1.29	1.92	0.90	
Japan	F(2,133)	0.82	1.04	0.17	2.05	1.11	0.18	
UK	F(1,134)	2.04	2.15	0.07	0.03	3.36	3.71	
Sweden	F(3,132)	0.75	0.35	0.41	0.83	0.08	0.16	
Switzerland	F(2,133)	1.80	2.43	1.50	0.06	0.59	0.96	
Norway	F(2,133)	3.15†	6.05†	2.07	0.03	0.25	3.01	
All Countries	10%	9%	26%	9%	9%	0%	5%	0%

Note: † denotes statistical significance at the 5% level.

Under the baseline model, there are several noteworthy findings. First and perhaps not surprisingly, the weak exogeneity of foreign inflation in the US model is rejected[21]. The question of foreign variable influence on the US model was discussed specifically in DdPS of which doubt was expressed by the authors as to whether the weak exogeneity of the foreign economic variables would be rejected in the US model. As it turned out in their shorter sample, which shares the same specification, weak exogeneity of foreign inflation in the US fails to be rejected. This is in contrast to the rejection in Table 3.12 and thus is evidence that in more recent years, most likely due to the financial crisis, there may now be long-run feedback from Δp_{0t} to Δp^*_{it} which implies that Δp^*_{0t} has become long-run forcing for Δp_{0t}, such that the error correction term of the US VECM now enters the marginal model of Δp^*_{0t}[22]. Having too many rejections is problematic for the validity of inference, however with regard to the purpose of this chapter the overall number of rejections is still within reason. The last row in table Table 3.12 reports the percentage of countries for which rejections occur. Across all variables and countries, the weak exogeneity assumption is violated 10% (14 out of 135 cases) of the time if measuring at the 5% level. The total proportion of rejections drops 3% (4 out of 135 cases) if measured at the 1% level.

[21] In the baseline model weak exogeneity test, the p-value for Δp^* for the US is 0.034 which indicates that the rejection may only be moderately strong.

[22] Let 0 represent the US model

Table 3.13: CV-CX Model F-statistics for Testing Weak Exogeneity of Foreign Variables

Country		Foreign Variables						
		y^*	Δp^*	q^*	ρ^{*S}	ρ^{*L}	p^o	$e^* - p^*$
USA	F(2,136)	0.01	5.01†					0.82
Euro area	F(1,134)	1.06	0.80	3.78†	1.51	0.03	0.70	
China	F(2,135)	0.18	0.54	0.57	1.91	2.48	1.25	
Japan	F(2,133)	1.23	0.57	0.28	2.78	2.51	2.16	
UK	F(1,134)	0.94	1.78	0.05	0.18	3.80	4.61†	
Sweden	F(3,132)	2.44	0.00	0.64	0.04	0.43	0.06	
Switzerland	F(2,133)	1.53	2.43	2.12	1.43	0.27	1.22	
Norway	F(2,133)	2.96†	3.94†	1.02	1.03	0.65	3.16†	
All Countries	10%	9%	17%	9%	9%	5%	14%	0%

Note: † denotes statistical significance at the 5% level.

Under the CV-CX model, very similar results are achieved with more rejections being observed on oil prices rather than inflation. The total proportion remains the same with exactly 14 out of 135 cases being rejected at the 5% level. At the 1% level, the number of rejections falls below 3% with only 3 cases being rejected. Broadly speaking, the weak exogeneity results in both models are acceptable for the purpose of this chapter. If the objective were instead to generate a model measuring and interpreting specific spillovers and transition effects, the problematic country-specific foreign variables would need to be handled. Due to the fact that it is necessary to have identical specifications across both models to isolate the effect of changing the weighting scheme and because weak exogeneity rejections don't occur on the same country-variable combinations for both baseline and CV-CX models, the level of rejection is viewed as acceptable.

3.2.5 Testing for Structural Breaks

Another key issue to be investigated is the incidence of structural breaks and parameter stability. To investigate this, the standard suite is applied which includes the following: Ploberger and Krämer (1992) maximal OLS cumulative sum (CUSUM) statistics, denoted PK_{sup}; its mean square variant, denoted PK_{msq}; Nyblom (1989) tests for parameter constancy against non-stationary alternatives, denoted $\Re$; the Wald form of Quandt (1960) likelihood ratio statistics, denoted QLR; the mean Wald statistics of Hansen, denoted MW; and Andrews and Ploberger (1994) Wald statistics based on exponential average, denoted APW. Robust versions of the last three are also considered. The first two tests are tests of parameter stability while the last three are Wald type tests utilizing a single break at an unknown point.

Table 3.14: Baseline Model Number of Rejections of the Null of Parameter Constancy Per Variable Across the Country-Specific Models at the 5% Level

Alternative test statistic	Domestic variables						Numbers(%)
	y	Δp	eq	ep	ρ^S	ρ^L	
PK_{sup}	3(13.0)	4(17.4)	0(0.0)	2(8.7)	3(13.0)	0(0.0)	12(9.8)
PK_{msq}	1(4.3)	1(4.3)	0(0.0)	2(8.7)	3(13.0)	0(0.0)	7(5.7)
$\Re$	1(4.3)	7(30.4)	4(17.4)	10(43.5)	6(26.1)	7(30.4)	35(28.7)
robust-N	0(0.0)	2(8.7)	1(4.3)	5(21.7)	4(17.4)	3(13.0)	15(12.3)
QLR	9(39.1)	12(52.2)	11(47.8)	16(69.6)	18(78.3)	11(47.8)	77(63.1)
robust-QLR	2(8.7)	4(17.4)	9(39.1)	6(26.1)	5(21.7)	7(30.4)	33(27.0)
MW	5(21.7)	9(39.1)	9(39.1)	12(52.2)	8(34.8)	7(30.4)	50(41.0)
robust-MW	0(0.0)	6(26.1)	5(21.7)	8(34.8)	3(13.0)	5(21.7)	27(22.1)
APW	9(39.1)	11(47.8)	12(52.2)	16(69.6)	18(78.3)	11(47.8)	77(63.1)
robust-APW	2(8.7)	5(21.7)	7(30.4)	6(26.1)	5(21.7)	7(30.4)	32(26.2)

Note: The test statistics PK_{sup} and PK_{mwq} are based on the cumulative sums of OLS residuals, $\Re$ is the Nyblom test for time-varying parameters and QLR, MW, and APW are the sequential Wald statistics for a single break at an unknown change point. Statistics with the prefix 'robust' denote the heteroskedasticity-robust version of the tests. All tests are implemented at the 5% significance level.

Table 3.14 reveals a small number of rejections for the PK tests, with the rejection range falling around 6-10%, which indicates that the parameters are relatively stable. The Nyblom tests and Wald type statistics, on the other hand, show a very high number of rejections even when considering the robust versions that allow for possible changes in error variances. Comparing the number of rejections between the robust and non-robust versions indicates that a sizable number of rejections is due to breaks in error variance. Specifically, allowing for possible changes in error variances in these tests (going from non-robust to robust) brings the rejection range down from around 28-63% to around 12-27%. The rejection range of the robust versions is somewhat closer to being in line with that of the PK and if those specific tests were to be believed, there is little statistical evidence with which to reject the hypothesis of coefficient stability in the case of 90% of the equations comprising the GVAR model. Considering the worst case Wald tests, one might conclude that the coefficients in only 73% of the equations comprising the GVAR model are stable. Comparing Table 3.14 to Table V. in DdPS, it is clear that the financial crisis almost certainly is a major contributing factor to the new-found instability. Further inspection of each table reveals very little evidence of instability (PK tests) or structural breaks (robust Wald type tests) in economic variables over the earlier sample. Considering the full sample of this chapter, which includes the financial crisis, we see greater evidence of parameter instability and structural breaks particularly manifesting more so for inflation than GDP. As for the financial variables, there is substantially more evidence of structural breaks specifically concentrated among equity prices and the bond markets. There also exists additional evidence of instability among exchange rates and short-term

interest rates but to a much lesser degree. Generally speaking, Table 3.14 provides a picture of where the financial crisis was most impactful and with hindsight, the predominant effects being on the bond market and then equity markets makes perfect since in the context of the reasons for the crisis[23].

Lastly, Table B.0.1 in Appendix B.0 presents the suite of stability tests on the CV-CX model. Comparing across Table B.0.1 and Table 3.14, there are no meaningful differences and thus we can conclude that both the baseline and the CV-CX model are equal in terms of parameter stability. Overall, just as in DdPS, there is evidence of instability but to a heightened degree now due to the sample including the financial crisis. Due to this, the focus of the main results is not on point estimates but on bootstrap means and confidence bounds.

3.2.6 Contemporaneous Effects of Foreign Variables on their Domestic Counterparts

Table 3.15 presents the contemporaneous effects from the baseline model of foreign variables on their domestic counterparts together with robust t-ratios computed using Whites heteroskedasticity-consistent variance estimator where the estimates can be interpreted impact elasticities. Looking across variables, we can see that output is fairly elastic in the focal countries and is significant in all focal countries except Norway. Impact elasticities for USA and the Euro Area both have values around 0.5% which implies that a 1% change in foreign real output will result in a 0.5% change in domestic output on impact. Comparing the output elasticities of Table VI in DdPS to those reported in Table 3.15, there are two meaningful changes. First, China has gone from, being insignificantly impacted to being among the most elastic. This can be interpreted as evidence of China's increased global connectedness in recent years[24]. Second, Norway's elasticity has now become completely insignificant.

Regarding inflation, the elasticity among focal countries that are significantly impacted ranges from 0.28% to 0.94% with an average of 0.6%. The level and variance of elasticities among inflation is highly similar to that of output and is evidence that economic variables behave similarly and in contrast to financial variables[25]. This could be a potential reason why pure trade weights are assigned to economic variables in the CV-CX model. Comparing the inflation elasticities reported in DdPS, the only meaningful change is that the UK inflation elasticity has gone from insignificant to significant. This suggests, in combination with the finding that the UK short-term interest rate

[23] Any differences between Table 3.14 and Table V. in DdPS can be interpreted as impacts of the financial crisis. In the sense of an event study framework, the longer sample in this chapter isolates the impact of the financial crisis. It is expected that the bond markets would be most impacted by the crisis due to the unprecedented level of securitization of mortgages through collateral debt obligations. See Verick and Islam (2010) for details on the causes, consequences and policy responses of the Great Recession.

[24] See Feldkircher and Korhonen (2014) for a detailed exposition of the role of China in the modern global economy.

[25] It could be the case that the linking of variables by trade weights, which are largely recognized to characterize global business cycles well (see Baxter and Kouparitsas (2005) and Imbs (2004)), is responsible for the similarity between the elasticities of economic variables.

elasticity has also become significant, that the financial crisis has increased the global integration of the UK financially[26].

Table 3.15: Baseline Model Contemporaneous Effects of Foreign Variables on Their Domestic Counterparts

Country	Domestic variables				
	y	Δp	eq	ρ^S	ρ^L
USA	0.53	0.13			
	[5.16]	[1.18]			
Euro Area	0.57	0.28	1.10	0.26	0.73
	[6.63]	[4.04]	[25.45]	[3.24]	[10.04]
China	0.71	0.68		0.04	
	[3.21]	[2.85]		[0.36]	
Japan	0.63	0.07	0.77	-0.03	0.49
	[3.67]	[0.89]	[7.64]	[-0.30]	[5.44]
UK	0.59	0.59	0.79	0.68	0.83
	[4.56]	[5.06]	[14.81]	[3.26]	[7.15]
Sweden	1.19	0.94	1.14	0.47	1.01
	[5.39]	[5.70]	[22.97]	[1.26]	[7.44]
Switzerland	0.50	0.28	0.87	0.28	0.54
	[3.55]	[2.27]	[15.84]	[1.42]	[7.27]
Norway	0.29	0.86	1.08	0.31	0.72
	[1.32]	[2.94]	[11.75]	[0.77]	[4.85]

Note: White's heteroskedastic-robust t-ratios are given in square brackets

Equity price impact elasticities are highly significant and are very close to unity, averaging 0.96% across focal countries. In some cases the equity price impact elasticity even exceeds unity, signifying an overreaction. Specifically, this is observed for Euro Area and Norse equity prices and is an effect that has continued to persist since it's documentation by DdPS. Regarding short-term interest rates, there is a good deal of variance among the elasticities across the focal countries with many effects being insignificant. In the updated sample reported here, the short-term interest rates of Sweden and Switzerland no long appear to be effected by their foreign counterparts while the elasticity for UK short-term rates is now highly significant. As for long-term interest rates, the average impact elasticity among the focal countries is 0.72% with significance across the board. This effect is up from 0.62% in the DdPS sample which again may simply be a symptom of the financial crisis and the subsequent unified global downturn.

Impact elasticities generated from the CV-CX model are not meaningfully different from the baseline model results in Table 3.15. The CV-CX results are presented in Table B.0.2 in Appendix B.0. Overall, there are few differences between the DdPS results and the results reported in this

[26] This effect is most likely due to the fact that global short-term interest rates all dropped to zero as leaders pledged more than 1 trillion dollars to tackle the global financial crisis. (see Ait-Sahalia et al. (2012) for details on the global policy response).

paper. It would appear that the financial crisis has had little effect on domestic-foreign impact elasticities.

3.3 Inference Sensitivity

In a similar vein to the section on impulse response analysis in the regional application of chapter 2, the main results of this chapter are drawn from impulse responses but with increased focus on inference. Inference by definition is a conclusion reached on the basis of evidence and reasoning and in terms of impulse responses analysis, is most commonly a judgment of the sign and significance of a given response based on confidence intervals. Impulse response functions are normally presented graphically because of the intuitive and efficient nature over the alternative tabular representation, and thus readers customarily form judgments over impulse responses visually. When viewing impulse response charts, the focus is usually first on whether the response is significant and then later on any magnitudes. The results here focus on the prior. In determining significance visually, a series of logical deductions are made which can be formally expressed. Algorithm 1, found below, defines the logic for classifying significance visually from an IRF chart. This classification of significance can be thought of as translating from impulse response to inference.

To summarize the classification logic in words, we first look at the direction of the response and then see if the relevant bound falls in the correct direction. For example, if the response is positive (negative), one would look at the lower (upper) bound and assess whether it falls above (below) zero and if it does, the inference would be that the response is positive (negative) significant. Another classification beyond simply significant is what is commonly referred to as "borderline." Visual assessment of a borderline response hinges on the relevant bound falling in the significant direction but being very close to zero, meaning that the strength of the inference is questionable. Assessment of whether the relevant bound falls close to zero critically depends upon the scale of the y-axis in the IRF chart. Typically, well depicted IRF charts will minimize whitespace above and below the upper and lower bounds respectively, thus a reasonable sense of the y-axis can be inferred from the difference between the maximum value of the upper bound and the minimum value of the lower bound across the entire horizon of the graphical depiction. Condition $v.$ in Algorithm 1 defines this difference as CI_{range} and is used in the conditional statements in combination with the borderline scaling parameter θ which serves as a lever for controlling how close a bound must be to zero before the response is classified as borderline. Throughout the results in this chapter, a value of 5% is used for θ which means that the relevant bound must fall between 0 and plus or minus 5% of the CI_{range} for it to be classified as borderline. This value was set by the author and is the value that resulted in the maximum number of correctly classified responses. Values above 5% tended to result in too many responses being classified as borderline when they were in fact significant and values below resulted in the too many responses being classified as significant when they were clearly borderline.

Algorithm 1: Impulse Response Inference Classification Algorithm

Input : Impulse Response Function with Confidence Intervals

Output: Classification of IRF significance

if sign($IRF_{mean.,h}$) = pos. **then**
  if $CI_{LB,h} > \theta CI_{range}$ **then**
    Response$_h$ = Positive significant
  else if $\theta CI_{range} > CI_{LB,h} > 0$ **then**
    Response$_h$ = Positive borderline
  else
    Response$_h$ = Insignificant
else
  if $CI_{UB,h} < (-\theta CI_{range})$ **then**
    Response$_h$ = Negative significant
  else if $(-\theta CI_{range}) < CI_{UB,h} < 0$ **then**
    Response$_h$ = Negative borderline
  else
    Response$_h$ = Insignificant

where:

i. $IRF_{med.,h} =$ average bootstrap mean estimate for horizon h,

ii. $CI_{LB,h} =$ average confidence interval lower bound for horizon h,

iii. $CI_{UB,h} =$ average confidence interval upper bound for horizon h,

iv. θ is the borderline scaling parameter,

v. $CI_{range} = \max CI_{UB,\{h\}} - \min CI_{LB,\{h\}}$ where $\{h\}$ is the set of all horizons

In the results that follow, a framework is presented for analyzing the impact of weighting schemes on inferences. Specifically, this chapter is interested in the number of inference classifications that differ between modeling under the baseline model (pure trade weights) and under the CV-CX model (trade-convex weights). This can be thought of as an empirical parallel to the simulation analysis in Gross (2018) who finds that GVAR parameters are highly sensitive to weighting scheme misspecification and often differ significantly.

GIRFs are generated under both the baseline model and the CV-CX model for all variables of all countries using Equation (1.10) for three different external shocks[27]: (1) a one standard error negative shock to US real equity prices; (2) a one standard error positive shock to US short-term interest rates; and (3) a one standard error positive shock to oil prices. It is important to note that the bootstrapping procedure takes into account sampling uncertainty in selecting

[27] The choice of shocks corresponds to that of Dees et al. (2007).

the convexity parameters θ_ρ. In other words, a complete weighting scheme grid search that is identical to what is described in earlier sections is performed for each bootstrap sample for the CV-CX model. Conducting the CX-procedure on each bootstrap sample has several effects on the impulse responses. First, simply by construction it produces wider confidence intervals than bootstrapping conditional on weighting. Second, the mean estimates are offset in various directions and magnitudes. In comparing impulse responses generated from each approach (conditional and unconditional on weighting), the ultimate results are extremely similar.

Responses are broken into three different time horizons where the short-run is between impact and 4 quarters, the medium horizon is the period between 4 and 12 quarters, and the long-run is between 12 and 24 quarters. Algorithm 1 is run on all impulse responses for each horizon to classify each response as positive or negative significant, borderline, or insignificant. The resulting set of inference classifications for each model (baseline and CV-CX) is then compared against it's counterpart.

In the individual shock analysis sections the same set of figures is presented, each with results corresponding to that particular shock. Select GIRF responses are first presented to illustrate clearly the differences in inference between the baseline and the CV-CX models. The next three figures are bar charts that utilizing the results of running Algorithm 1. The first shows the percentage of responses by variable and at each horizon for which the inference is not the same under each model. Following this is a figure that shows each specific type of inference change as a result of switching from the baseline to the CV-CX model, again broken down by horizon. The last bar chart aggregates the inference results into two categories; strengthening and weakening. An example for strengthening of inference would be the case where a given inference is classified as insignificant under the baseline model but as significant (or borderline) under the CV-CX model. Similarly, a case for weakening of inference would be one for which under the baseline model a given inference is classified as significant but under the CV-CX model it is classified as borderline (or insignificant). This chart also reports results separately by horizon. One of the driving forces behind the results reported in the bar charts is the tightening of confidence bands from increasing the fit (reducing SSE). The chart immediately following the bar charts presents reductions in confidence interval bandwidth by reporting the following formula for each variable:

$$\frac{CI^{\text{CV-CX}}_{range,\{h\}}}{CI^{\text{baseline}}_{range,\{h\}}} - 1.$$

The last chart presents globally aggregated (via PPP-GDP) responses across all countries for each variable from both baseline and CV-CX models.

3.3.1 Shock to US Equity Prices

Consider first the GIRFs for a one standard error negative shock to US equity prices. This shock equates to a fall of around 5-6% in US real equity prices per quarter. Figure 3.2 presents select GIRFs for illustrative purposes of the sensitivity to weighting. Odd rows depict responses from the baseline model and even rows depict responses from the CV-CX model.

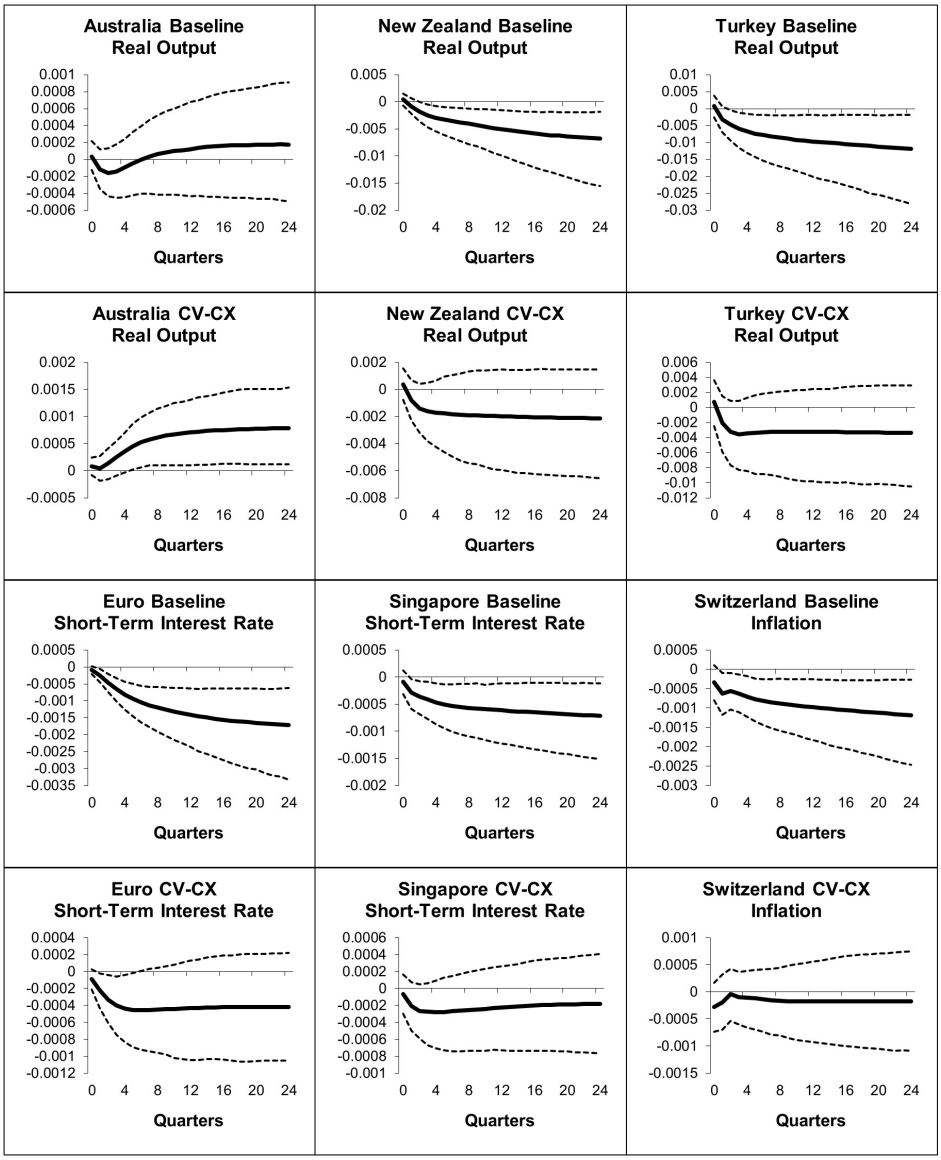

Figure 3.2: Select GIRFs of a negative unit (1 s.e.) shock to US real equity prices. Rows 1 and 3 generated from the baseline model. Rows 2 and 4 generated from the CV-CX model. (bootstrap mean estimates with 90% bootstrap error bounds)

Each response illustrates a case where an inference change is observed. For example, looking at the panels in the top-left, one would conclude that Australian GDP is not significantly impacted by a negative US equity price shock (at any horizon) but only if he used a model with purely

trade weights. If he were to use the CV-CX model, he would conclude that, while not significantly effected in the short-run, over the medium and long-run the response of Australian GDP appears to be positive and significant. Similarly, by inspecting the top-middle panels for New Zealand GDP, one would conclude negative significance under trade weights but insignificance under CV-CX weights. The top-right panes illustrate the case of borderline negative under trade weights, and insignificant under CV-CX weights. The bottom panels illustrate similar cases. Of particular interest might be the bottom left panel though which depicts the Euro Area short-term interest rate response. In DdPS, which uses exclusively trade weights, and also here under the baseline model the conclusion would be that rates drop significantly in response to a negative US equity price shock. The CV-CX model disagrees though, and shows that the significant portion of the negative response may only be limited to the short-run rather than across all horizons. The bottom-middle panels tell the equivalent story for short-term rates in Singapore, although the baseline response is only borderline. The bottom-right panels show the same sensitivity but for inflation in Switzerland. The responses depicted in Figure 3.2 were selected to illustrate that the sensitivity ranges across all dimensions. Looking at individual panels reveals sensitivity across horizons and directions while the variable and country selection reveals that the sensitivity isn't limited across space or to certain variables.

To shed light on the question of how ubiquitous the sensitivity is across variables, Figure 3.3 presents the percentage of countries where any inference change is observed at each horizon as a result of switching from the baseline model to the CV-CX model[28]. Regarding the short-run, Figure 3.3 shows that the variables most susceptible to an inference change are exchange rates and output, both of which a change is observed in as many as 40% of countries. As for the medium horizon, differences in inference on short-term interest rates are observed in as many as 45% percent of countries. Worse yet, inferences in the long-run differ from baseline and CV-CX models in over 60% of countries for long-term interest rates, 50% for short-term interest rates, 40% for exchange rates, and no less than 20% for other variables. Generally, Figure 3.3 provides evidence that weighting scheme sensitivity on inferences from impulse responses to a US equity price shock is not contained across space, variables, or horizons but instead is widespread.

[28] Any inference change includes all possible combinations of the inference classifications: positive/negative and significant, borderline, or insignificant.

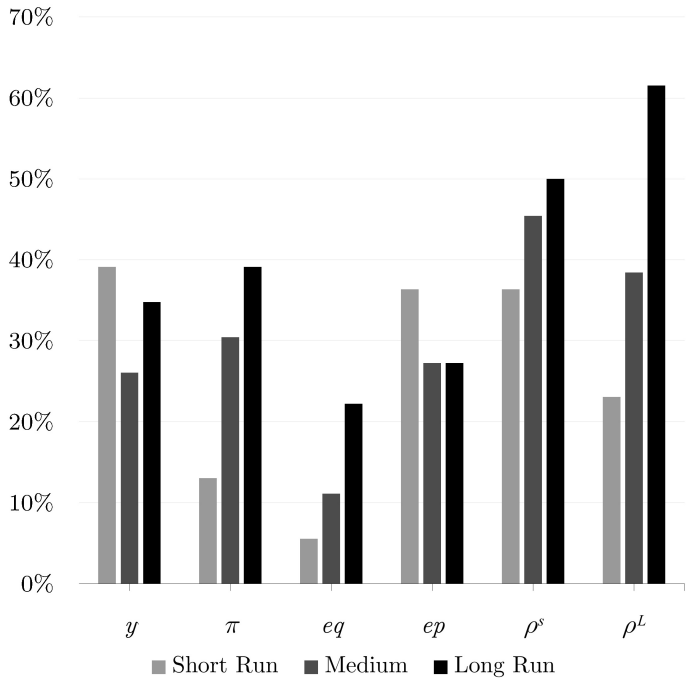

Figure 3.3: Percentage of impulse responses for which the inference differs between baseline and CV-CX models by variable and by horizon. (Results derived from a negative unit (1 s.e.) shock to US real equity prices)

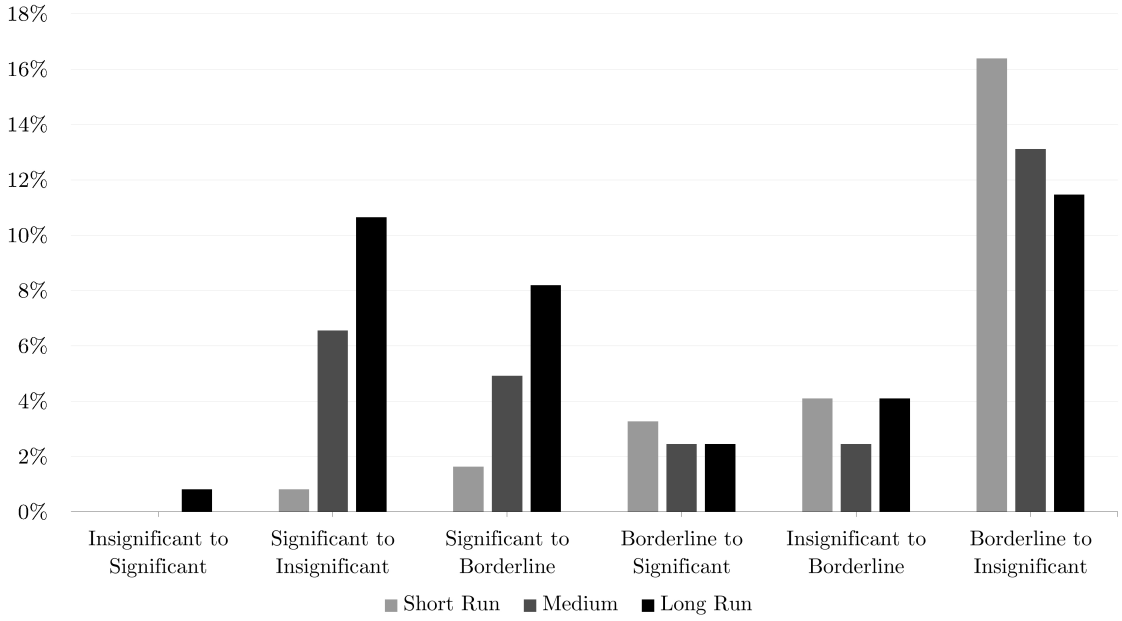

Figure 3.4: Percentage of impulse responses for which the inference differs between baseline and CV-CX models by type of inference change and by horizon. (Results derived from a negative unit (1 s.e.) shock to US real equity prices)

Concerning the types of inference changes that are observed in response to a negate US equity

price shock, Figure 3.4 aggregates across variables and shows the percentage of responses that fall into each of six categories[29]. Recall that each category is interpreted as an inference change that occurs as a result of switching from the baseline to the CV-CX model. In the short-run, the bulk of the action is in the case for which a response was classified as borderline under the baseline model but insignificant under the CV-CX model. Interestingly, this is the case at all horizons. At the medium and long-run horizons however, the distribution of inference changes includes a broader range of cases such as significant to insignificant and significant to borderline with the proportion of cases being relatively consistent across each of these longer horizons. Generally though, as for most types of changes, the percentage of inference changes increases at longer horizons.

Further aggregating similar effects in Figure 3.4 into cases for which the inference is strengthened or weakened, Figure 3.5 illustrates that the majority of inference changes are in the form of weakening[30]. Figure 3.5 also reports the sum of all cases where it can be observed that more

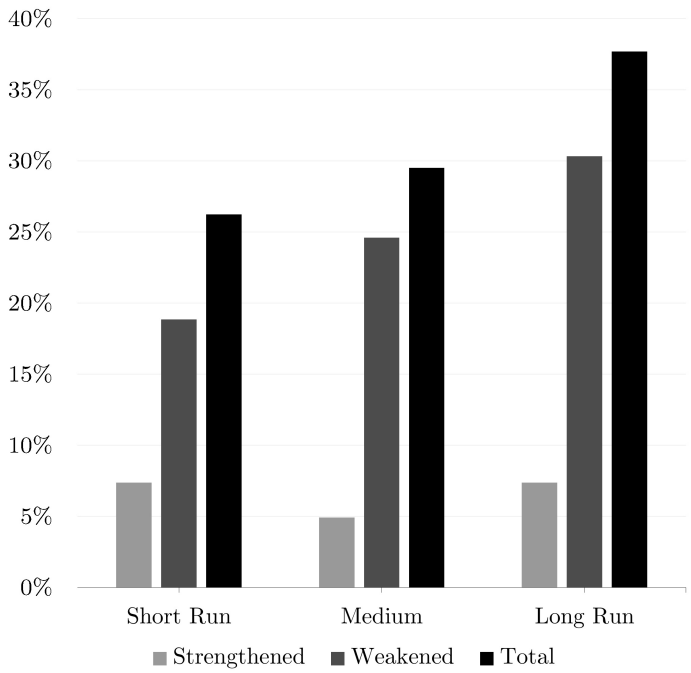

Figure 3.5: Percentage of all inferences strengthened and weakened at each horizon as a result of switching from the baseline to the CV-CX model. (Results derived from a negative unit (1 s.e.) shock to US real equity prices)

changes occur at longer horizons. This is not entirely surprising as the CV-CX model fits better than the baseline and as a result would be expected to have tighter confidence intervals for which

[29] Note that there are no cases for which the extent of the change was great enough that the direction of the finding changed (i.e., in no case does the inference switch from positive significant to negative significant). Due to this, only the cases for which inference changes occurred are reported in Figure 3.4 where each case represents a result either strengthening or weakening but in same direction.

[30] Recall that weakening includes the following cases: significant to borderline, significant to insignificant, and borderline to insignificant. Cases for strengthening are the reverse.

the magnitude of tightening compounds over time.

To investigate the claim of tighter confidence bands under the CV-CX in the context of the negative shock to US equity prices, Figure 3.6 reports the average change in confidence interval bounds range from GIRFs for each variable[31]. Surprisingly, in the short-run, minimal reduction is observed in the bandwidth for most variables and in some cases, such as for inflation and the interest rate variables the bandwidth is actually wider. As the horizon increases though, substantial tightening is observed across all variables with magnitudes varying between 15% tighter for inflation to around 50% tighter for real equity prices in the long-run with the average across all variables falling around 30% tighter. Figure 3.6 directly clarifies the impact that the improved fit has on GIRF confidence intervals. Curiously, if one regresses the data in the bottom panel of Figure 3.6 on a time trend, the coefficient is highly significant and equal to roughly -1%.

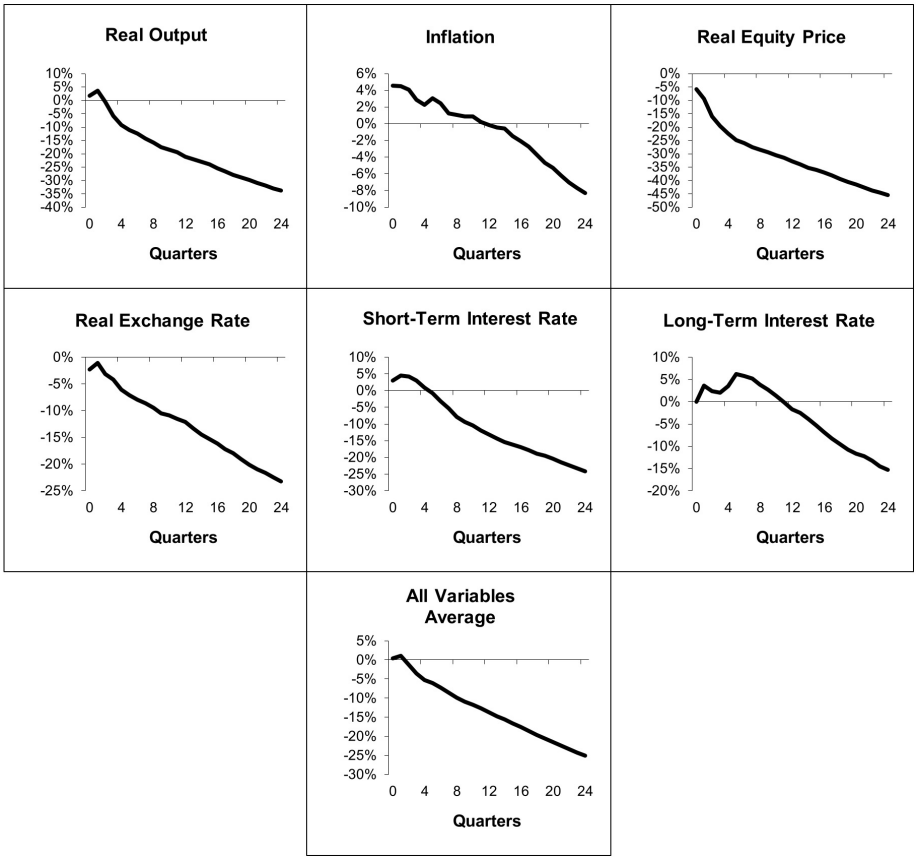

Figure 3.6: Percent change in GIRF 90% bootstrap confidence interval range ($\text{CI}_{UB} - \text{CI}_{LB}$) from baseline to CV-CX model. Each panel displays an averaged result across all countries. Negative values indicate tighter confidence intervals under the CV-CX model. (Results derived from a negative unit (1 s.e.) shock to US real equity prices)

This implies that for each additional quarter beyond impact, one can expect the confidence intervals

[31] The formula used in the computation of Figure 3.6 is given towards the end of Section 3.3.

to be 1% tighter under the CV-CX model than under the baseline model. It may simply be a coincidence that the quarterly tightening is approximately equal to a factor of one tenth of the fitness improvement though. More research would need to be done to strengthen any such claim[32].

A final point of interest in terms of the baseline and CV-CX models is the resulting global responses. Global responses are computed as PPP-GDP aggregations of individual responses. Figure 3.7 reports the global responses of each variable to the negative US equity price shock from each model for which there are several noteworthy points. Beginning with the differences in magnitudes between the models, generally the CV-CX model paints a milder picture of the global response than the baseline model does. However, while both models agree on output, equity prices, and short-term interest rates, key differences exist in responses of inflation, exchange rate, and long-term interest rates. Looking first at inflation, the baseline model response is counter-intuitive. Ordinarily we would expect the decline in short-term rates to cause inflationary pressure, however the baseline model shows a deflationary response. The inflation response from CV-CX model on the other hand falls in the expected direction. The models also disagree on the global response of long-term interest rates.

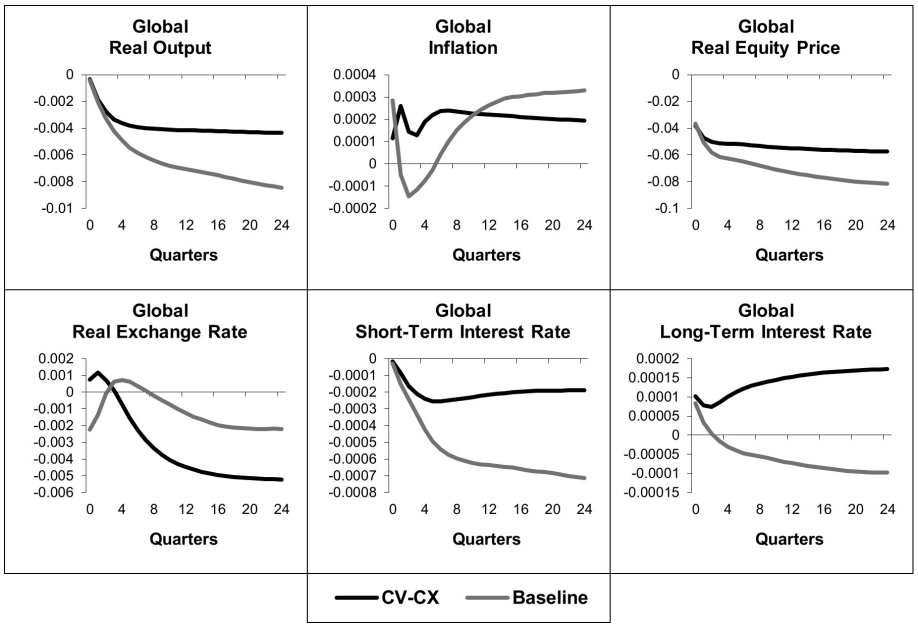

Figure 3.7: Global response to a negative unit (1 s.e.) shock to US real equity prices. Panels report PPP-GDP aggregated responses from the both the baseline and the CV-CX model. (bootstrap mean estimates)

The baseline model has equity prices and long-term interest rates moving together while the CV-CX model has them moving inversely. The upward pressure on the long-term interest rates under the CV-CX model shows that the bond markets tend to react more to inflation expectations rather than

[32] How confidence interval tightness is related to the overall fit in GVAR models is an extremely specific question. If one deemed such a pursuit worthy of further investigation, a simulation study would likely be the best place to start.

to growth prospects. Moreover, under the CV-CX model the bond and equity market reactions are consistent with each other whereas their response is inconsistent under the baseline model.

Generally, a more coherent story of the global economy is produced with the CV-CX model. Starting with the equity price decline, global output falls, and central banks to drop short-term interest rates to combat the economic slowdown. The reduction in short term rates means more money in the economy or increased inflation and thus increased inflation expectations which ultimately creates upward pressure on long-term interest rates and causes the yield curve to steepen. Lastly, global currencies show an appreciation and then a subsequent depreciation against the dollar with the action occurring sooner under the CV-CX model. This is evidence that the rest of the world overreacts to the US equity price decline in terms of monetary policy in the long-run.

3.3.2 Shock to Oil Prices

The second shock of consideration is the positive one standard error shock to oil prices in the US model. A one standard error positive shock results in a 12-13% increase per quarter in the price of oil. Figure 3.8 presents GIRFs for select country-variable combinations for which various inference changes are observed. Just as in the previous section, the GIRFs presented were selected to illustrate the sensitivity both across space and across variables. Several shocks are of particular interest as they were also presented in DdPS and thus their presentation in this section offers an opportunity to compare across time in addition to the comparison to weightings. The top-middle panel presents the Euro area short-term interest rate response to the positive oil price shock. The GIRF response reported in DdPS shows short-run borderline positive response that is comparable to the baseline model's response in this chapter. The CV-CX model on the other hand shows that the response is positive significant across all horizons. This is evidence that perhaps the monetary response in the Euro area to oil prices is more sensitive than was previously suggested. The second noteworthy shock in Figure 3.8 is the bottom-middle panel that depicts the US equity price. Since DdPS primarily focuses on transmission and spillover effects between the US and the EU, this shock is also included in their set of focal responses. Again, the baseline model response reported in this chapter is highly similar to the response in DdPS which shows a negative borderline response in the short-run becoming significant in the medium to long-run. By contrast, the CV-CX model shows a much dampened effect that is insignificant at all horizons which is evidence that the US stock market may be more resilient than has been previously suggested.

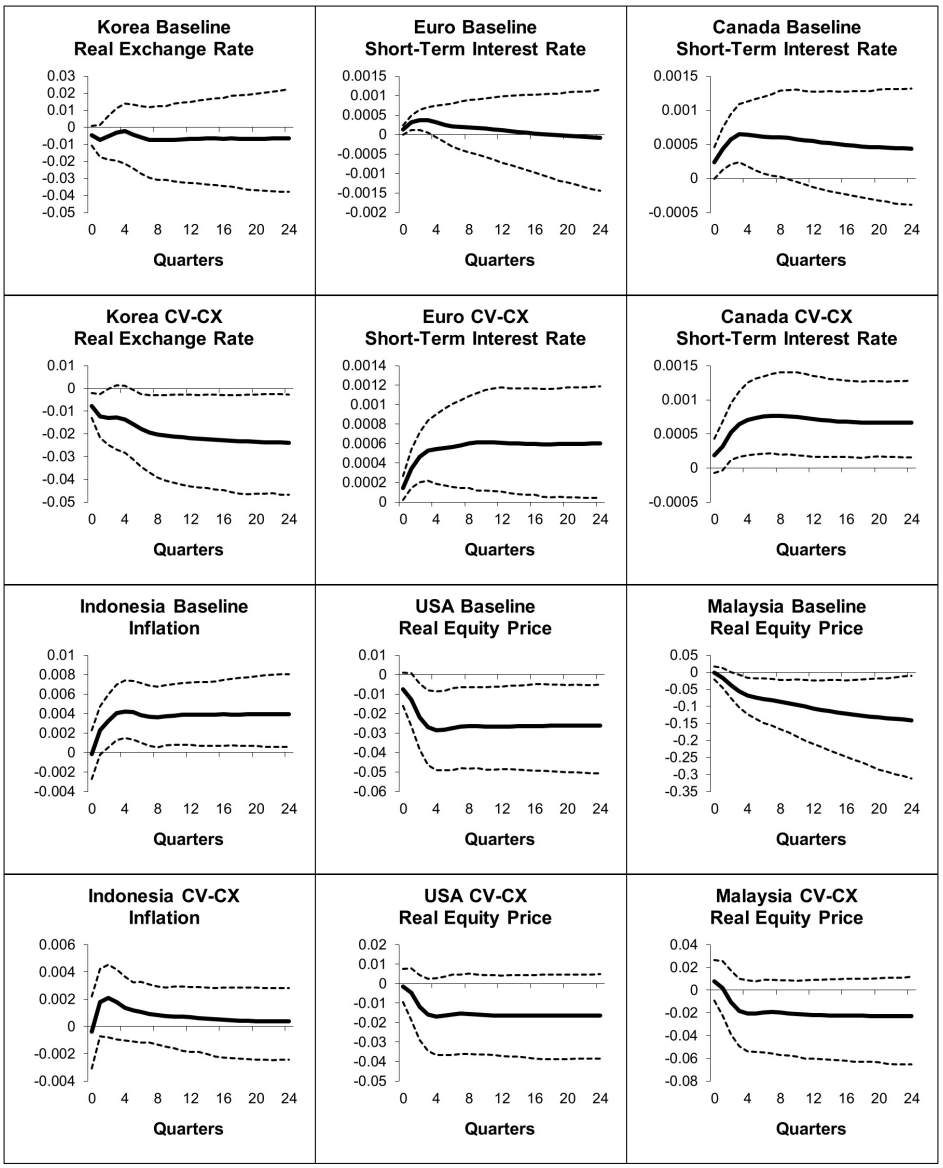

Figure 3.8: Select GIRFs of a positive unit (1 s.e.) shock to oil prices in the US model. Rows 1 and 3 generated from the baseline model. Rows 2 and 4 generated from the CV-CX model. (bootstrap mean estimates with 90% bootstrap error bounds)

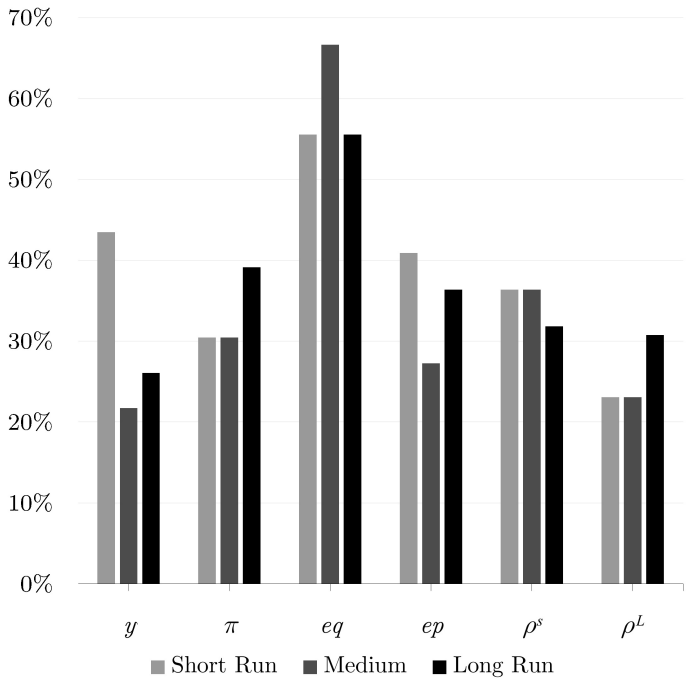

Figure 3.9: Percentage of impulse responses for which the inference differs between baseline and CV-CX models by variable and by horizon. (Results derived from a positive unit (1 s.e.) shock to oil prices in the US model)

Figure 3.9 presents the inference sensitivity by variable. In the short-run, the sensitivity is fairly evenly distributed across variables with the majority of changes occurring in equity prices ($> 50\%$) followed closely by output ($> 45\%$). In the medium horizon, the number of countries for which the equity price inference for baseline and CV-CX models differs is even greater with changes occurring in nearly 70% of countries. Less differences are observed over the medium horizon than in the short run for other variables though. Generally, the prevalence of inference changes between the baseline and CV-CX models for the oil price shock is concentrated among equity prices and is primarily a weakening of the inference which suggests that equity markets are more resilient to oil price movements than prior models have suggested.

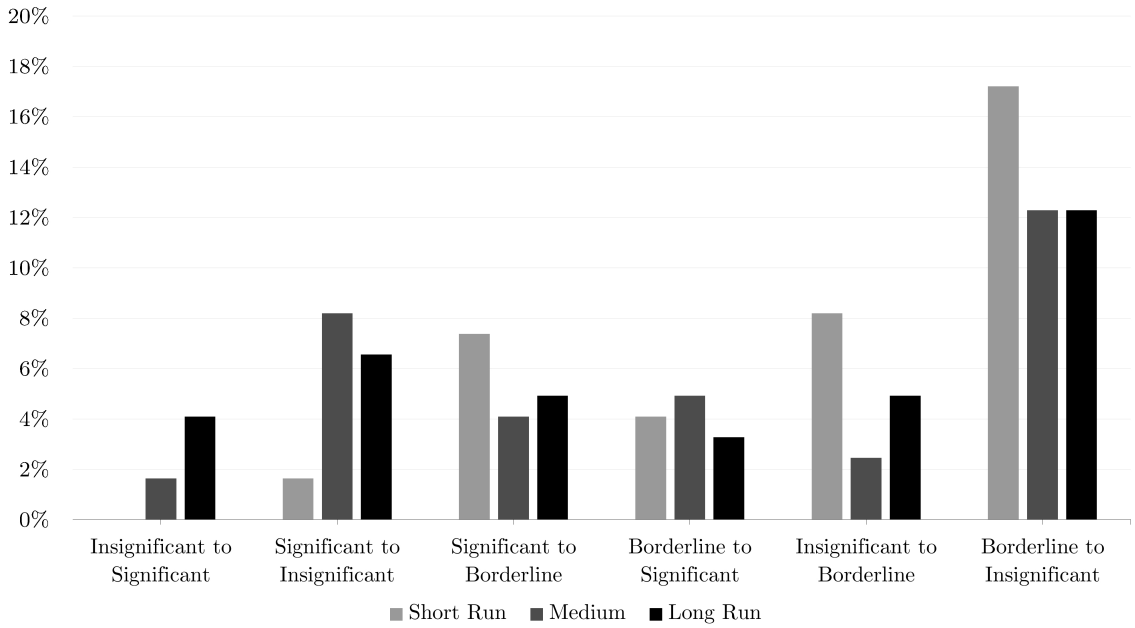

Figure 3.10: Percentage of impulse responses for which the inference differs between baseline and CV-CX models by type of inference change and by horizon. (Results derived from a positive unit (1 s.e.) shock to oil prices in the US model)

Figure 3.10 shows the percentage of oil price shock inferences that differ between baseline and CV-CX models by type of change. Here again, as was the case with the shock to US equity prices, it is shown that the most of the action is in responses being considered borderline under the baseline model and being considered insignificant under the CV-CX model.

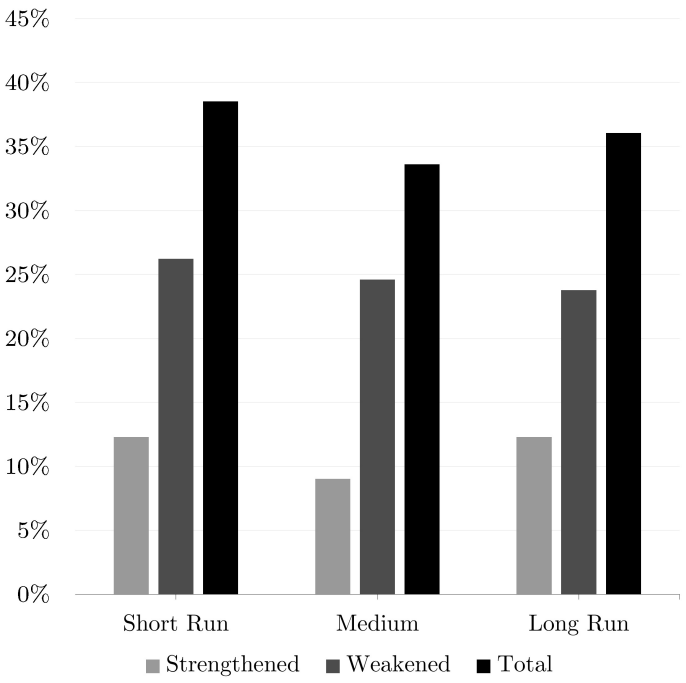

Figure 3.11: Percentage of all inferences strengthened and weakened at each horizon as a result of switching from the baseline to the CV-CX model. (Results derived from a positive unit (1 s.e.) shock to oil prices in the US model)

Concerning the categorization of oil price shock inferences into strengthening and weakening as a result of switching from the baseline to the CV-CX model, Figure 3.11 shows the strengthening/weakening breakdown as well as the total percentage of variable/country combinations for which a difference is observed at each horizon. In response to the oil price shock, it is shown that across all horizons under the CV-CX model roughly a quarter of the inferences drawn from GIRFs are weaker in terms of significance than under the baseline model. On the other hand, oil price shock GIRF inferences drawn from the CV-CX model are only stronger for around a tenth of the variable/country combinations. Overall, the total number of oil price shock GIRF inferences that are difference between the baseline and CV-CX models is around 35% across all horizons. This is in contrast to the inference results from the US equity price shock where there were more differences as the horizon lengthened.

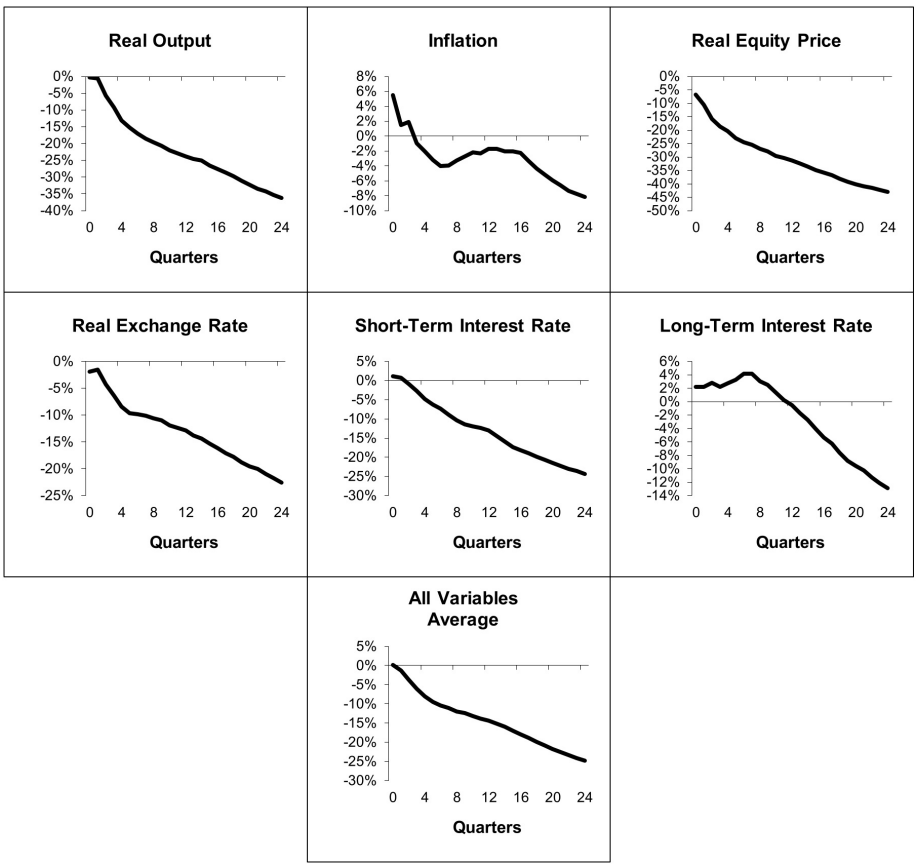

Figure 3.12: Percent change in GIRF 90% bootstrap confidence interval range ($\text{CI}_{UB} - \text{CI}_{LB}$) from baseline to CV-CX model. Each panel displays an averaged result across all countries. Negative values indicate tighter confidence intervals under the CV-CX model. (Results derived from a positive unit (1 s.e.) shock to oil prices in the US model)

Inspecting the bounds range improvement from the CV-CX model over the baseline, Figure 3.12 shows tightening across all variables. Whereas there was only minimal tightening in the short-run to the US equity price shock, there is tightening across all variables even in the short run with magnitudes varying between around 2% tighter to over 10% tighter. Equity prices and output are again the most impacted with the bandwidth being nearly 50% tighter in the long run for equities and 40% tighter for output under the CV-CX model. Again, a similar compounding tightness is observed with the effect being roughly 1% tighter under the CV-CX model than the baseline model per quarter. Broadly speaking, the tightening of bootstrap confidence intervals for oil price GIRFs is similar in scale to the tightening observed for US equity price GIRFs.

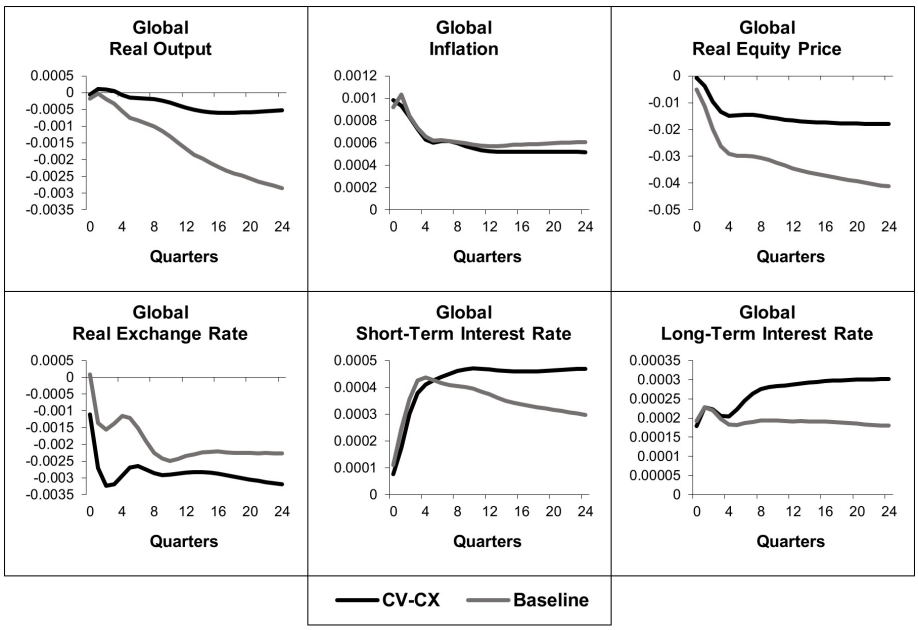

Figure 3.13: Global response to a positive unit (1 s.e.) shock to oil prices in the US model. Panels report PPP-GDP aggregated responses from the both the baseline and the CV-CX model. (bootstrap mean estimates)

Lastly, Figure 3.13 reports PPP-GDP aggregated global GIRF responses to the positive oil price shock. Generally, the results are in the same directions as the baseline but with mixed differences in magnitudes. Compared to the baseline model, the CV-CX model suggests that output and equity prices are less effected by the oil price shock but that interest rates are more effected. Specifically, the CV-CX model suggests that equity markets are more resilient to oil prices while bond markets are more sensitive the previously suggested by trade weights.

3.3.3 Shock to US Short-Term Interest Rate

The last shock considered is a positive one standard error shock to short-term interest rates in the US. This amounts a 0.14% increase in the Fed Funds rate (i.e., around 56 basis points), measured on a quarterly basis. The magnitude of the shock is smaller than that in DdPS due to the extended period of time that the rates remained stable around 0 following the financial crisis.

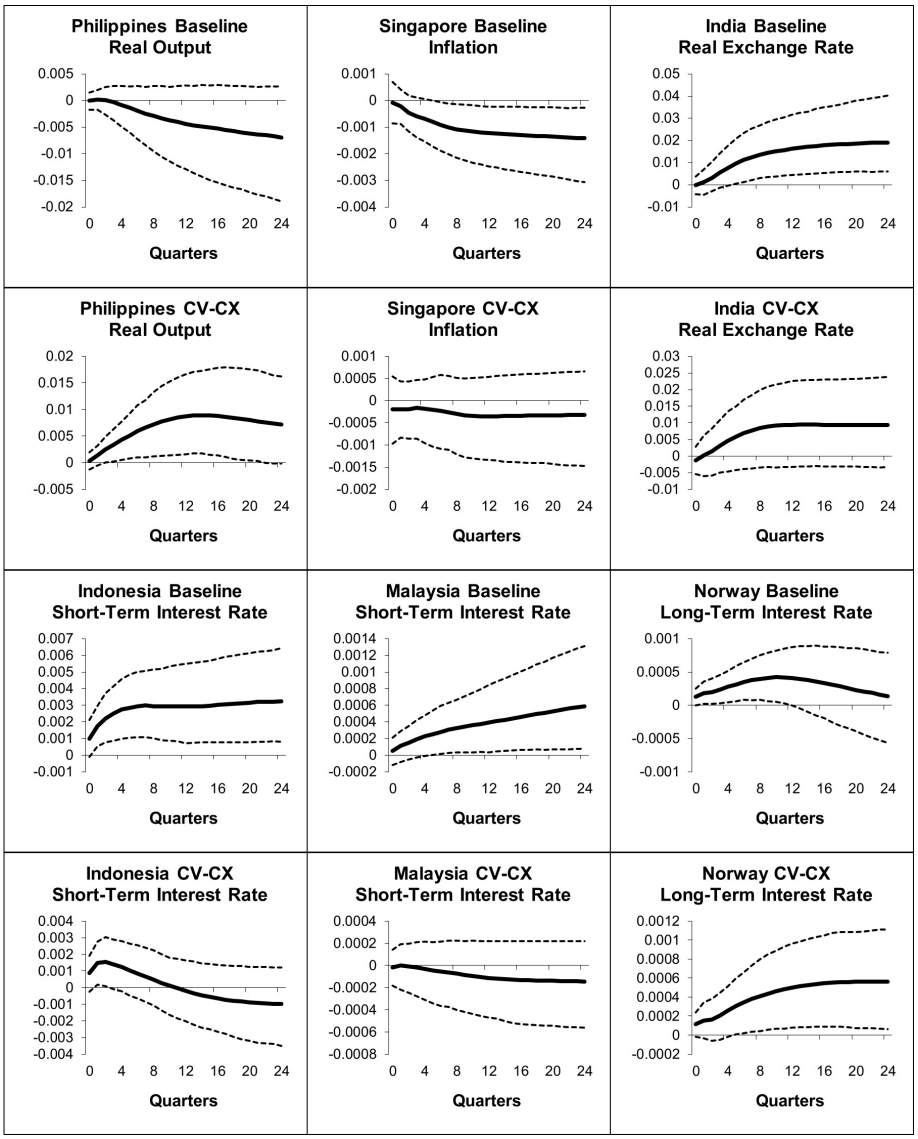

Figure 3.14: Select GIRFs of a positive unit (1 s.e.) shock to US short-term interest rates. Rows 1 and 3 generated from the baseline model. Rows 2 and 4 generated from the CV-CX model. (bootstrap mean estimates with 90% bootstrap error bounds)

Figure 3.14 presents select GIRF responses for which significant differences are observed between the baseline and CV-CX models. The chosen responses, as with the previous shocks, illustrate inference differences across all dimensions. The top panels show that differences exist across economic variables and exchange rates while the bottom panels show differences to interest rates.

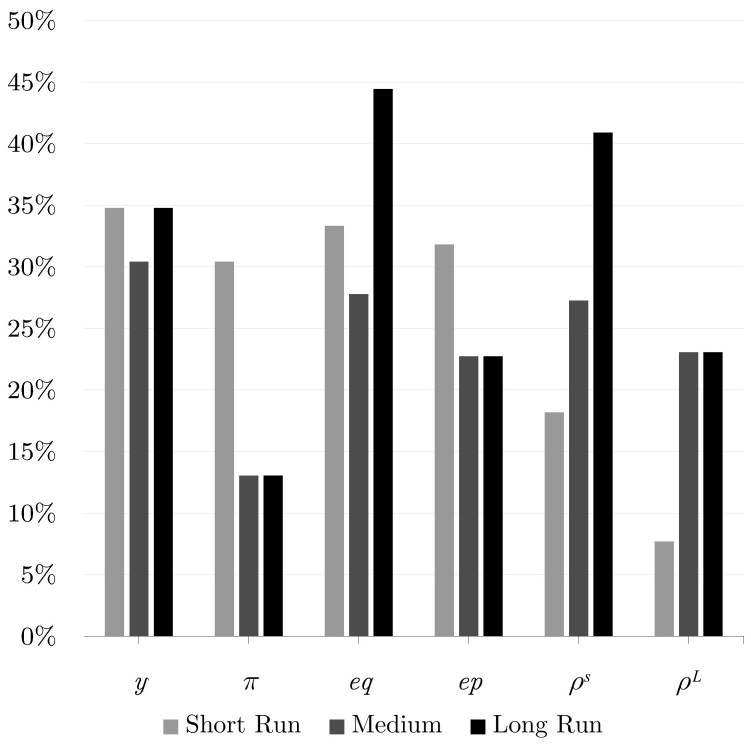

Figure 3.15: Percentage of impulse responses for which the inference differs between baseline and CV-CX models by variable and by horizon. (Results derived from a positive unit (1 s.e.) shock to US short-term interest rates)

Breaking down the changes by variable, Figure 3.15 reveals that the distribution of changes across variables at the short and medium term are fairly similar. In the short-run, the percentage of countries for which the GIRF inference differs is roughly around 30% for all variables except for the long-term interest rate for which the number of observed differences is less than 10%. As for the medium term, the percent of observed differences is between 20-30% for all variables except inflation. In the long-run however, equity price and short-term interest rate inferences drawn from baseline and CV-CX models differ for around 50% and 40% of cases respectively which, as with the rest of the evidence, implies that the linkage of financial variables by trade weights may be questionable. Generally, Figure 3.15 serves to validate the claim that US short-term inference rate shock GIRF inference differences are not simply limited to a single variable but exist across all variables and at each horizon.

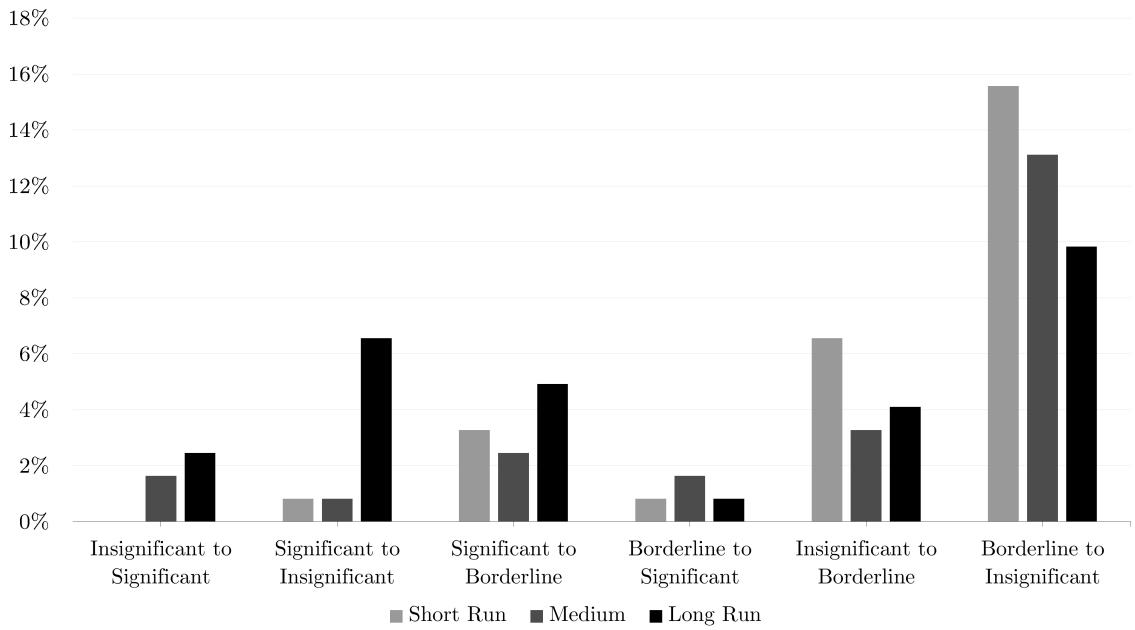

Figure 3.16: Percentage of impulse responses for which the inference differs between baseline and CV-CX models by type of inference change and by horizon. (Results derived from a positive unit (1 s.e.) shock to US short-term interest rates)

Figure 3.16 shows the type of change observed as a result of switching from the baseline to the CV-CX model for the US short-term interest rate shock. Similarly to the other shocks, of the responses where differences exist, the bulk of the action is the case for which responses were borderline under the baseline model and insignificant under the CV-CX model. Regarding the distribution of types of observed inference changes in the short-run, this is overwhelmingly the case with very few other types of changes occurring. As for the medium horizon, there is slightly more variation among represented cases but still most of the changes are of the the weakening case from borderline to insignificant. As for the long-run, the distribution of cases is more uniform but still with the borderline to insignificant case holding the majority. Overall Figure 3.16 shows that, while the majority is held by a single type of inference change at all horizons, other types of inference changes are represented and it is not the case that switching from the baseline to the CV-CX model will alter all of the inferences in the same way.

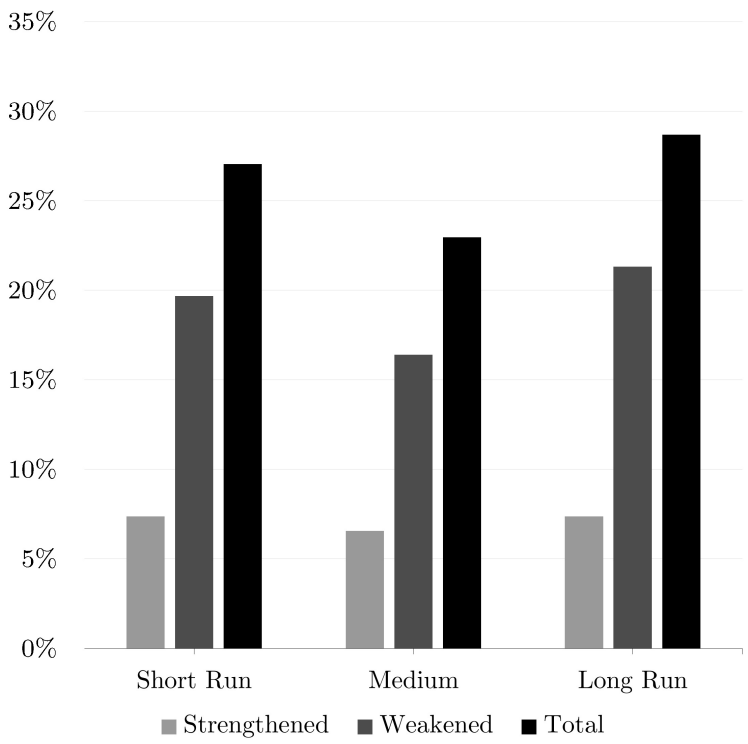

Figure 3.17: Percentage of all inferences strengthened and weakened at each horizon as a result of switching from the baseline to the CV-CX model. (Results derived from a positive unit (1 s.e.) shock to US short-term interest rates)

To further illustrate the variation in inference changes between the baseline and CV-CX models, Figure 3.17 presents the strengthening and weakening categorization of types of changes. It is again shown that the majority of changes at all horizons are those for which the inference is weakened by the CV-CX model. This is particularly the case in the long run where greater than 20% of inferences are less significant than under the baseline model. Across all horizons, less than 10% of inferences are strengthened by changing to the CV-CX model. As for the total number of inferences where a difference is observed, a higher percentage occurs in the long run which may not be unexpected in light of the tightening of confidence bounds.

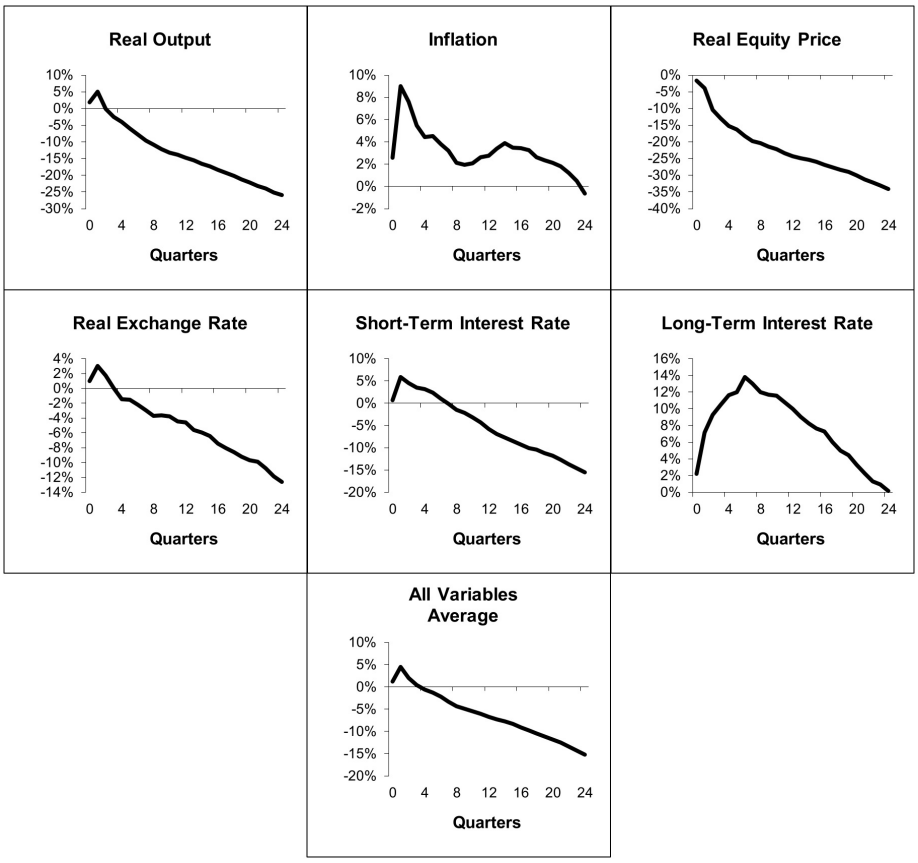

Figure 3.18: Percent change in GIRF 90% bootstrap confidence interval range ($\text{CI}_{UB} - \text{CI}_{LB}$) from baseline to CV-CX model. Each panel displays an averaged result across all countries. Negative values indicate tighter confidence intervals under the CV-CX model. (Results derived from a positive unit (1 s.e.) shock to US short-term interest rates)

Specifically regarding confidence interval bandwidth change, Figure 3.18 shows a pattern for all variables that looks highly similar to that of the bandwidth changes for the first shock to US equity prices. In the short-run the bands are actually wider (values above 0) under the CV-CX model than under the baseline model. This is in line with the short-run findings of Figure 3.16 that show a large percent of impulses that are considered borderline or significant under the baseline model but are considered insignificant under the CV-CX model. The time it takes the bands to become tighter greatly differs across variables with bands starting tighter on impact for equity prices but taking just over 20 quarters for long-term interest rates. The magnitude of improvement in bandwidth also differs across variables with long-run bandwidths being roughly 40% tighter for equity prices and 30% tighter for output while just breaking even for long-term rates. Generally, excluding the odd behavior of the long-term interest rate bounds, fairly consistent tightening is observed across all variables.

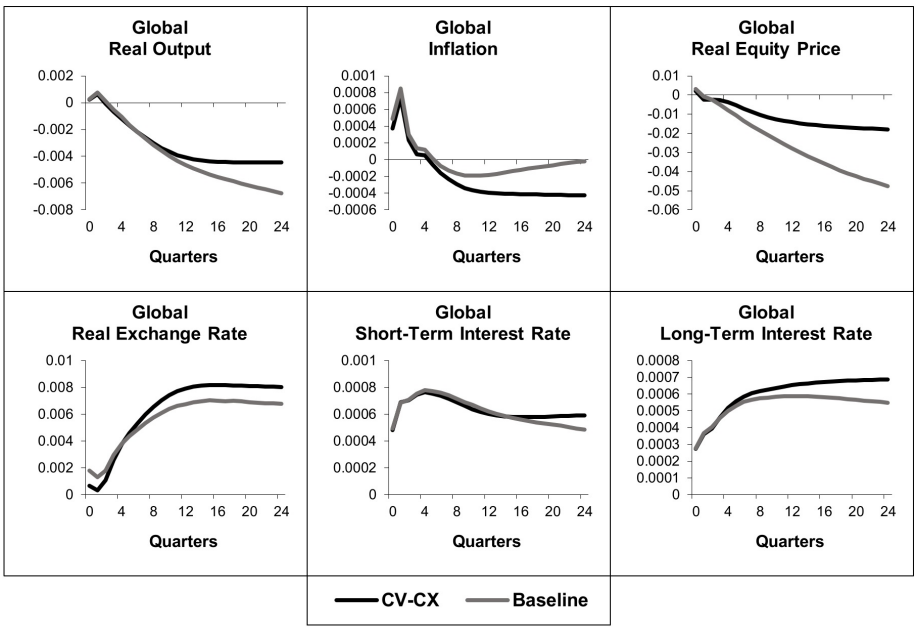

Figure 3.19: Global response to a positive unit (1 s.e.) shock to US short-term interest rates. Panels report PPP-GDP aggregated responses from the both the baseline and the CV-CX model. (bootstrap mean estimates)

Lastly, Figure 3.19 presents the global responses to the US short-term rate hike. Under both the baseline and the CV-CX models, there are no disagreements in the direction of the responses across all variables. Magnitudes differ slightly by variables and in a similar way to the oil price shock with equity prices less impacted and long term rates more impacted. Interestingly, the puzzle referred to in DdPS regarding the positive output and inflation response is still present but with far smaller magnitudes under each case. Finally, global exchange rates appreciate against the dollar which suggests that the global monetary response follows suit but to such a degree as to be considered an overreaction.

3.4 Concluding Remarks

This chapter presents a full-scale empirical application of the global economy with the intent and purpose of demonstrating the sensitivity of generalized impulse response inferences to weighting scheme selection. Together with the innovation of convex weighting presented in chapter 1, this chapter also proposes an extension of the existing weighting scheme literature by combining the concepts of weighting scheme mixing on variables with convex weighting. To illustrate the scope of inference sensitivity to weighting, two separate models are estimated for which inferences from generalized impulse responses are compared. A baseline model that relies on pure trade weights, which is analogous to the model found in the seminal GVAR work by Dees et al. (2007), is run on an updated sample alongside the optimal convex weighted model which is identified via a full search algorithm. Another contribution of this chapter is the impulse response inference classification

algorithm that is capable of categorizing the direction of impulse responses as significant, borderline, or insignificant for any given horizon. This algorithm is used in the analysis of all impulse responses generated from three individual shock scenarios and the resulting inferences of the baseline and convex weighted models are compared along a number of dimensions with the primary focus being on whether any differences exist.

Table 3.16: Any Inference Change - Any Country, Any Variable

	Short Run	Medium Term	Long Run
Neg. shock to US equities	26%	30%	38%
Pos. shock to oil prices	39%	34%	36%
Pos. shock to US short rates	27%	24%	30%
Average	31%	29%	34%

Note: Entries are the percentage of IRFs for which any inference change occurs as a result of switching from the baseline to the CV-CX model.

Table 3.16 presents a tabular representation of the main results where rows are individual shocks and elements are the percentage of GIRFs for which a different level of significance was observed between the baseline and the convex weighted model. The last row presents the average across all shocks and shows that in no less than practically 30% of impulse responses at any horizon, a different inference would be reached and thus likely a different conclusion would be argued. In line with the arguments put forth in chapter 2, if GVAR practitioners are interested in maximizing the strength of their arguments and minimizing the amount of assumptions they make, some form of weighting scheme optimization should be employed, such as the procedures set forth and demonstrated in this dissertation.

Broadly speaking, the GVAR as a model has been used in a wide range of applications, see for reference Chudik and Pesaran (2016), but the vast majority of the literature to date still relies on the specification of a single weighting scheme to link domestic and foreign variables together. While this may be appropriate in select circumstances where the linkage mechanism is abundantly clear, this dissertation suggests that the real risk is similar to omitted variable bias in the classical sense. Choosing to estimate only a single weighting scheme that may have been selected in an ad hoc fashion is to shun the importance of robustness to that dimension. Estimating, analyzing results, and forming arguments from a single weighting schemes is to ignore the potential that the weighting scheme might influence the results. Rejecting such potential in favor of economic context afforded by a user-specified scheme might previously have been preferable to the alternative, which was the devoid of economic context endogenously estimated weighting proposed in Gross (2018), but fortunately by the innovations in this dissertation, a simple go-between has been developed that balances the economic context and considers the sensitivity that weighting schemes may have.

Although future research would do well to shed additional light on the exact nature and magnitude of the sensitivity of various forecasting and inference-type results in GVAR models to the choice of weighting, this paper has provided real-world evidence that the choice can have major implications on both. Researches looking to utilize the GVAR method will no longer, by the methods proposed in this dissertation, need to concern themselves with the exact specification of the linkage mechanism and will no longer have to sacrifice the economic context that alternatives require. This dissertation provides a meaningful contribution to the GVAR literature in the area of weighting scheme choice and lays a plentiful foundation for future applications.

APPENDIX A.0

Table A.0.1: Distance Weights

MSA	Dallas	Austin	Houston	San Antonio
Dallas	0.000	0.217	0.264	0.178
Austin	0.402	0.000	0.415	0.592
Houston	0.309	0.262	0.000	0.230
San Antonio	0.289	0.521	0.321	0.000

Table A.0.2: GDP Weights

MSA	Dallas	Austin	Houston	San Antonio
Dallas	0.000	0.436	0.674	0.432
Austin	0.164	0.000	0.170	0.109
Houston	0.687	0.463	0.000	0.458
San Antonio	0.150	0.101	0.156	0.000

APPENDIX A.1

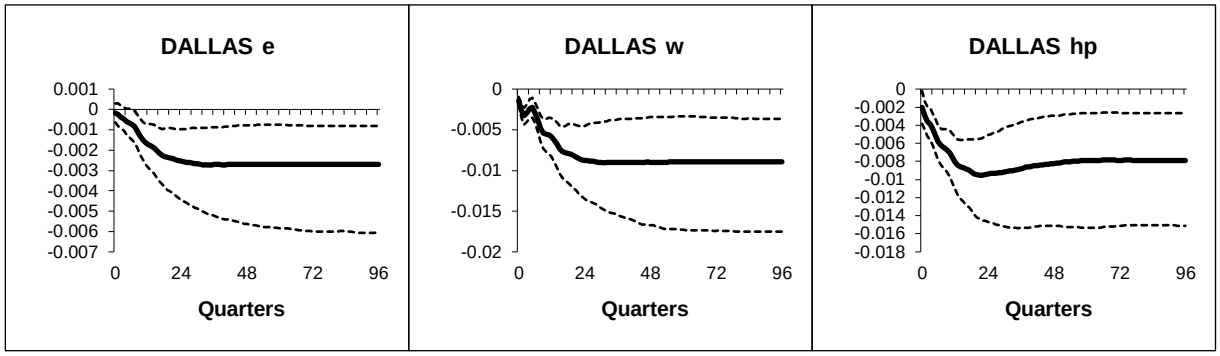

Figure A.1.1: Dallas - Generalized impulse responses of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

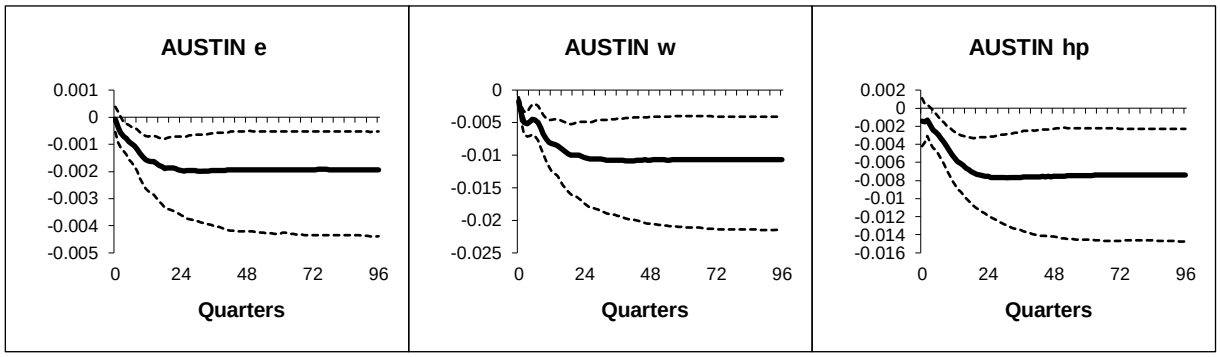

Figure A.1.2: Austin - Generalized impulse responses of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

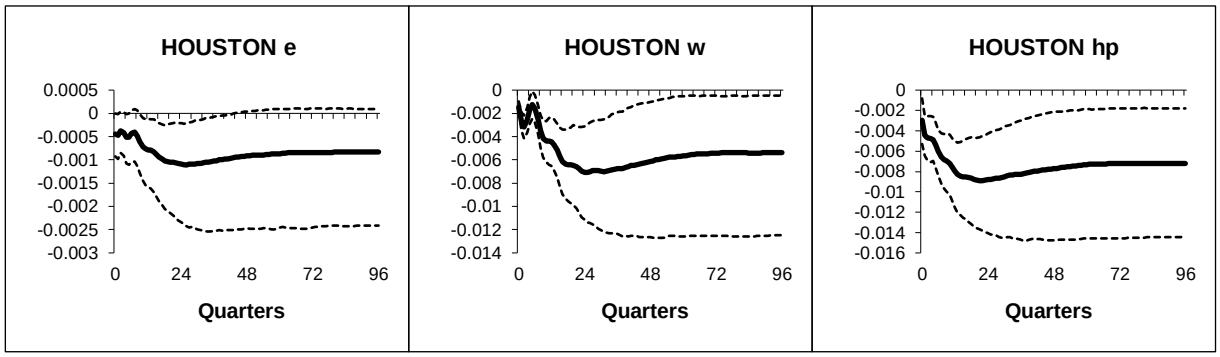

Figure A.1.3: Houston - Generalized impulse responses of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

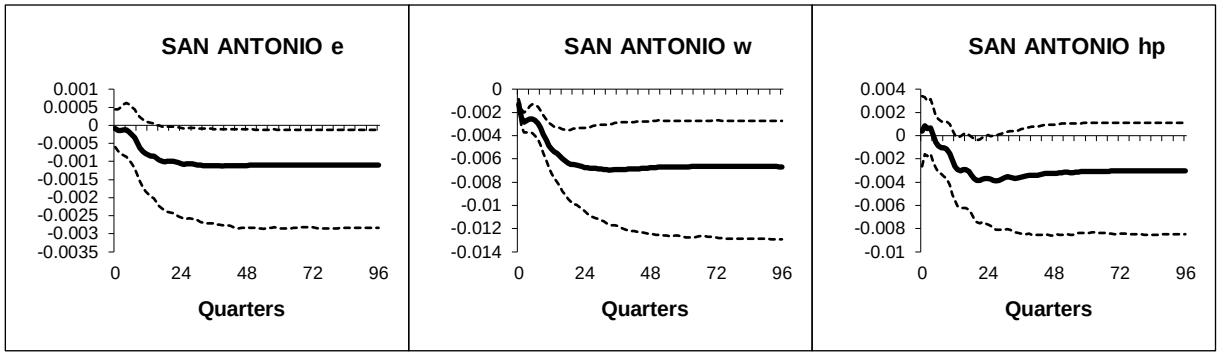

Figure A.1.4: San Antonio - Generalized impulse responses of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

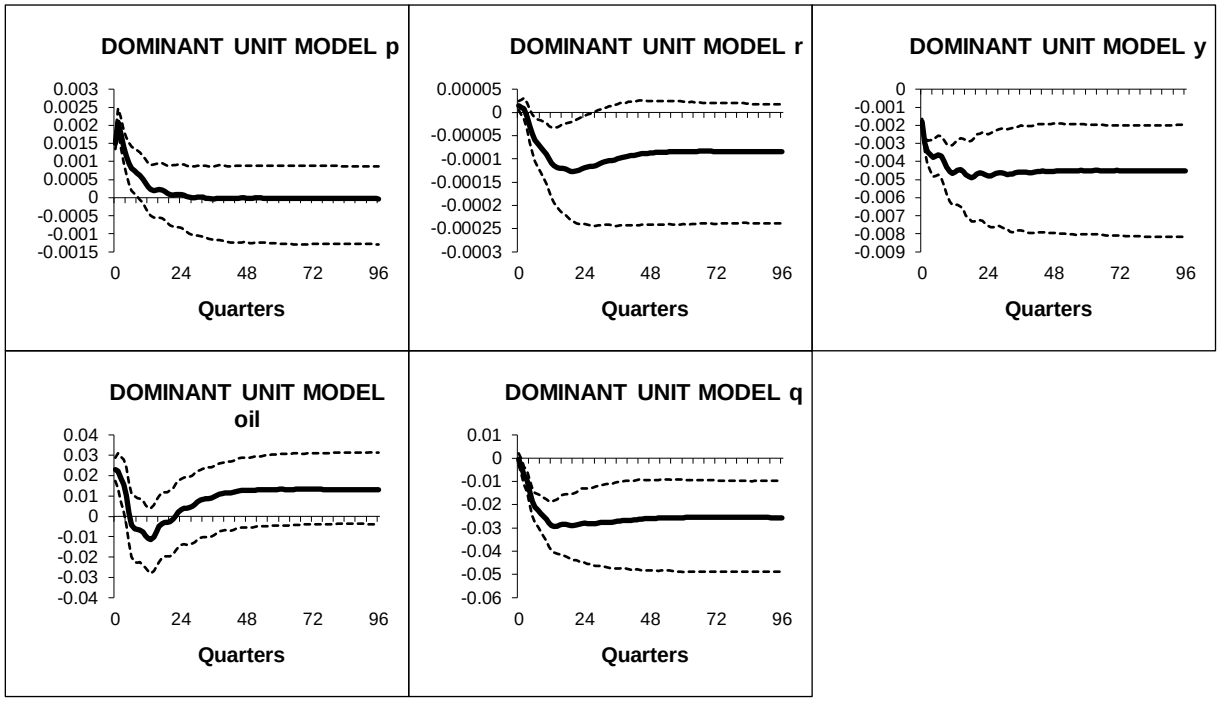

Figure A.1.5: DU - Generalized impulse responses of a negative (1 s.e.) shock to US GDP (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

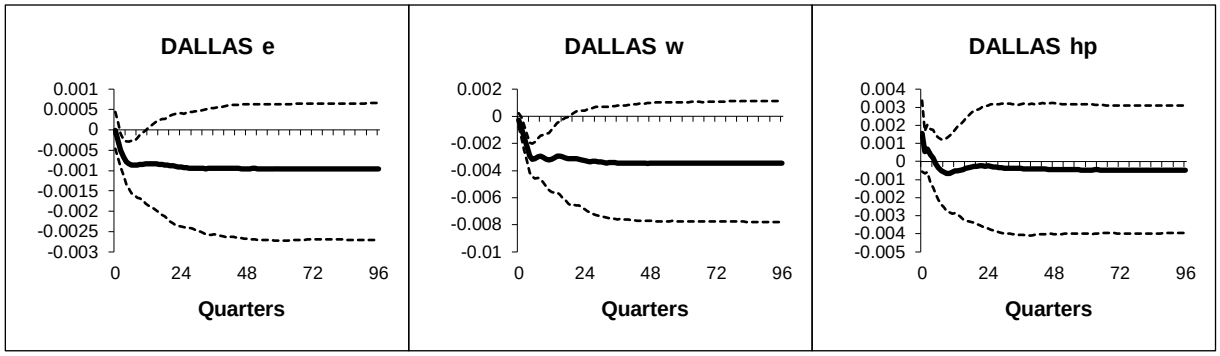

Figure A.1.6: Dallas - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

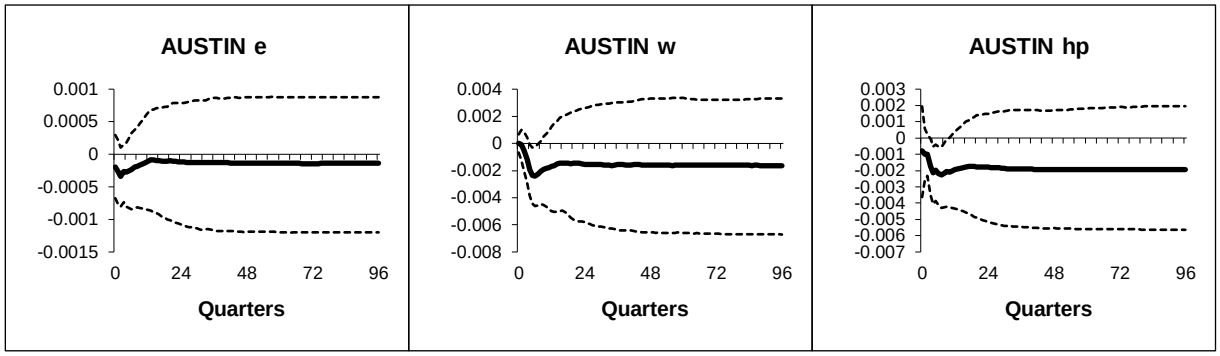

Figure A.1.7: Austin - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

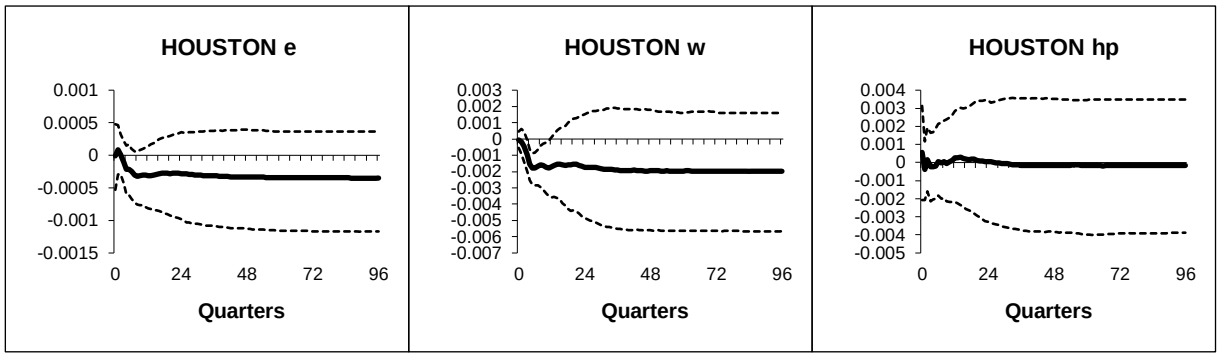

Figure A.1.8: Houston - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

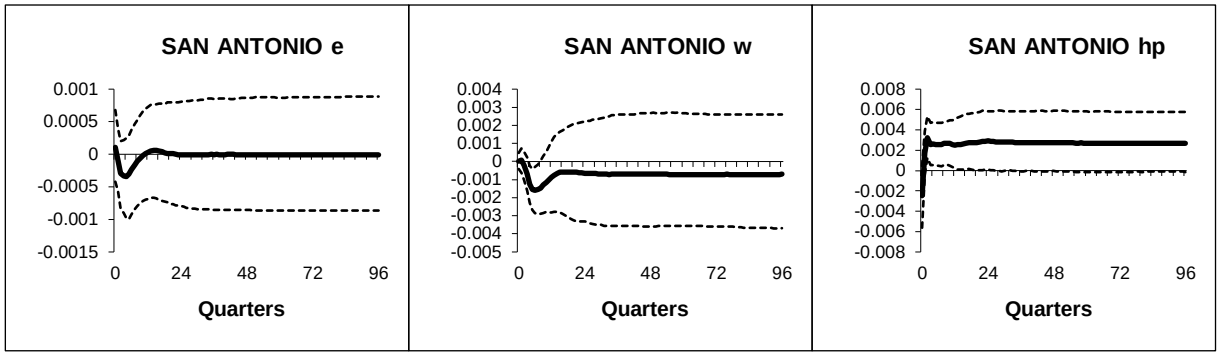

Figure A.1.9: San Antonio - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

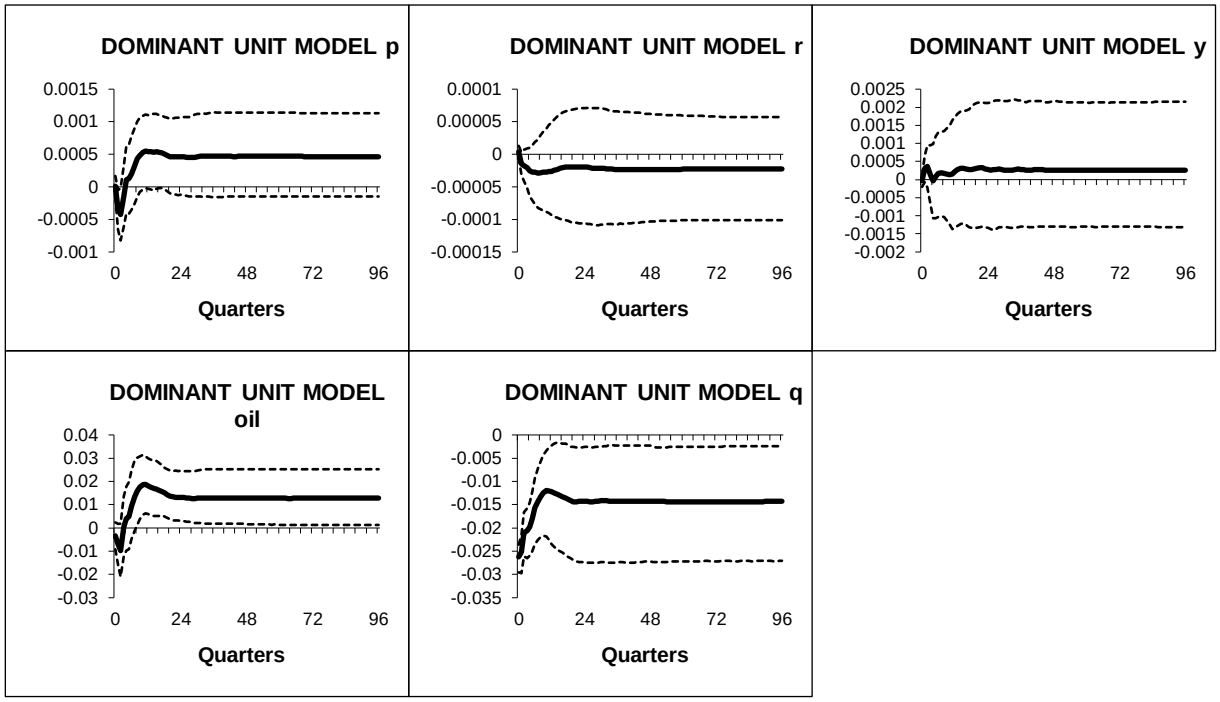

Figure A.1.10: DU - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

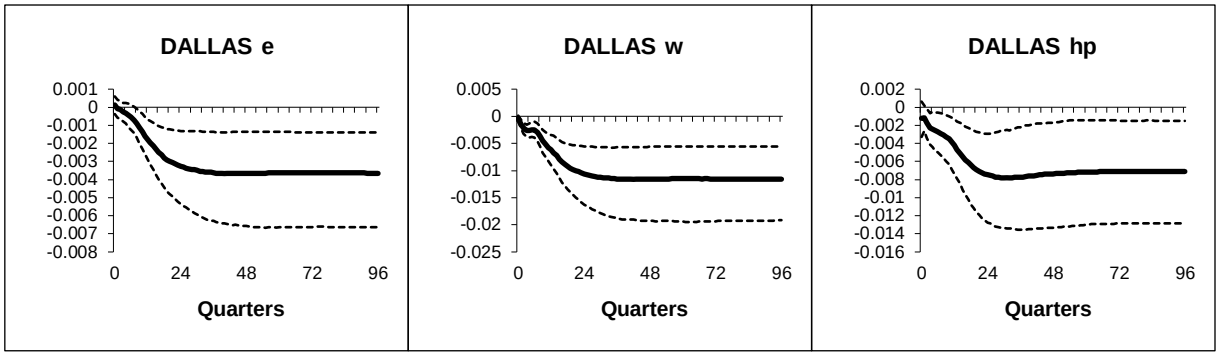

Figure A.1.11: Dallas - Generalized impulse responses of a positive (1 s.e.) shock to oil prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

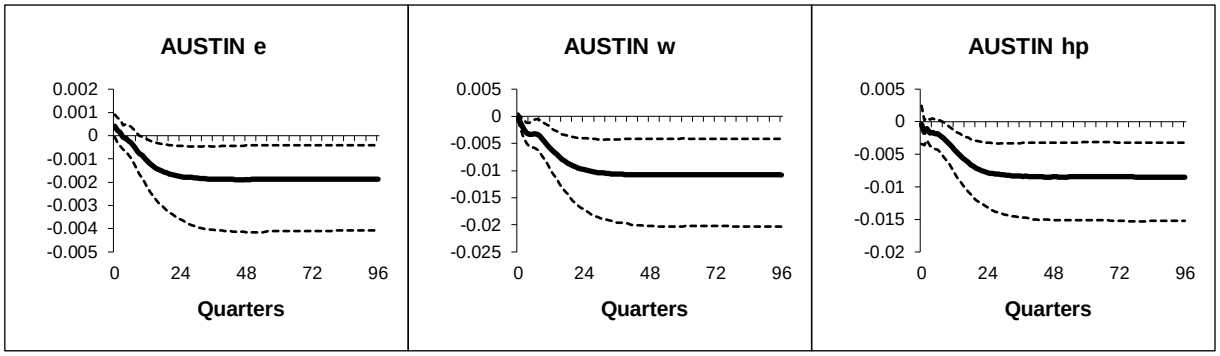

Figure A.1.12: Austin - Generalized impulse responses of a positive (1 s.e.) shock to oil prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

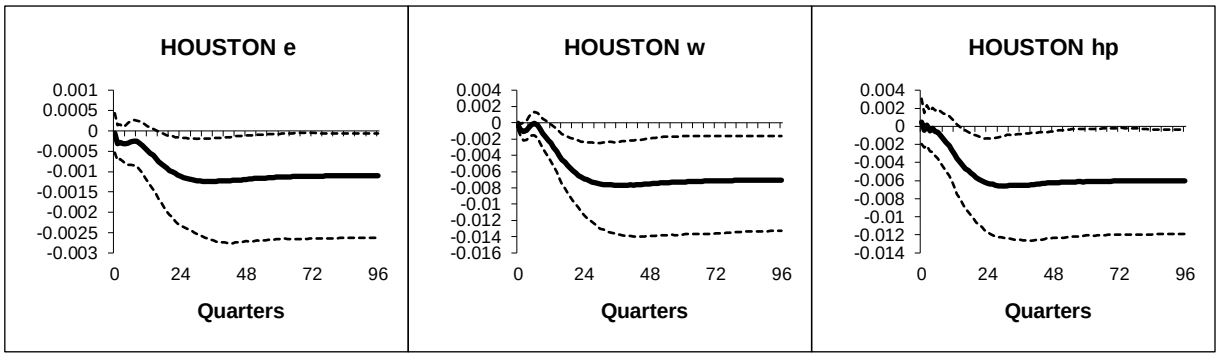

Figure A.1.13: Houston - Generalized impulse responses of a positive (1 s.e.) shock to oil prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

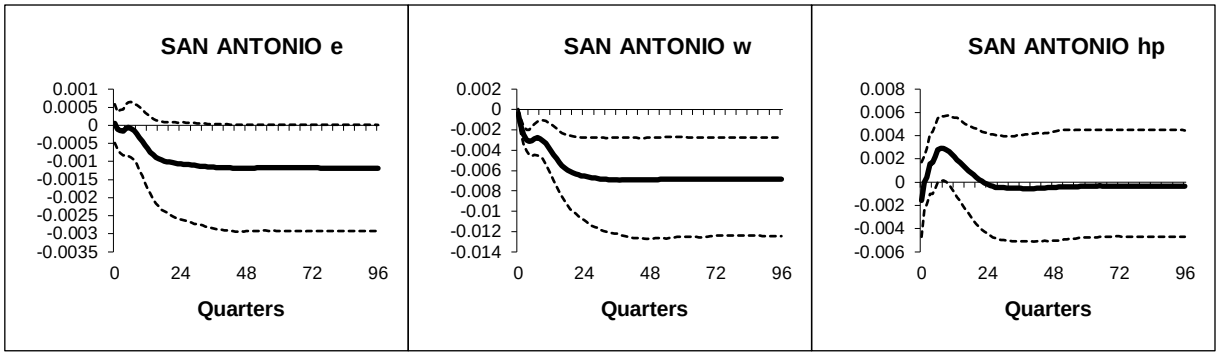

Figure A.1.14: San Antonio - Generalized impulse responses of a positive (1 s.e.) shock to oil prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

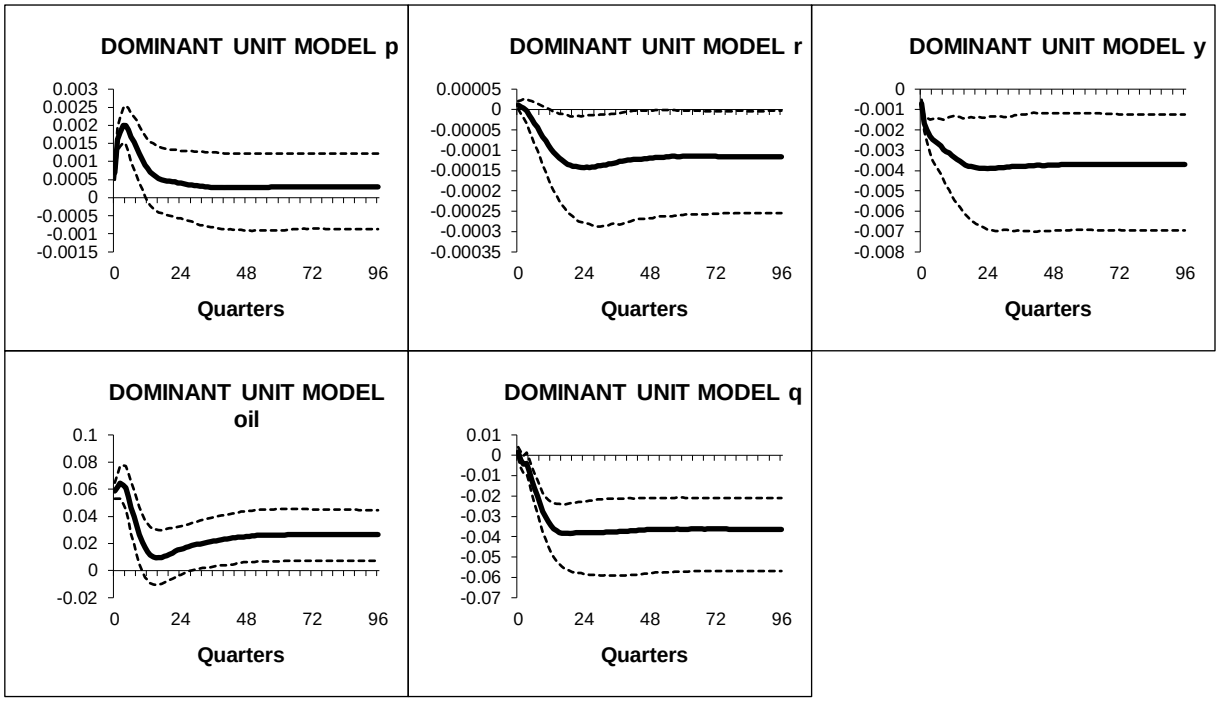

Figure A.1.15: DU - Generalized impulse responses of a positive (1 s.e.) shock to oil prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

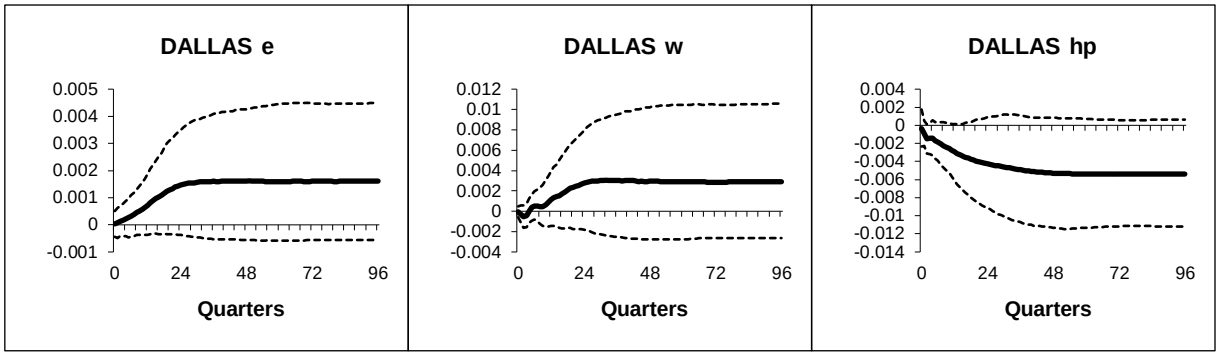

Figure A.1.16: Dallas - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

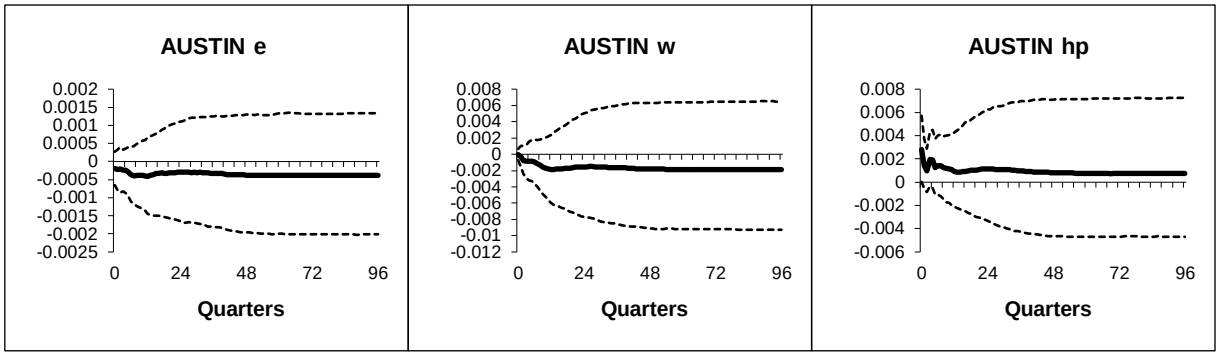

Figure A.1.17: Austin - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

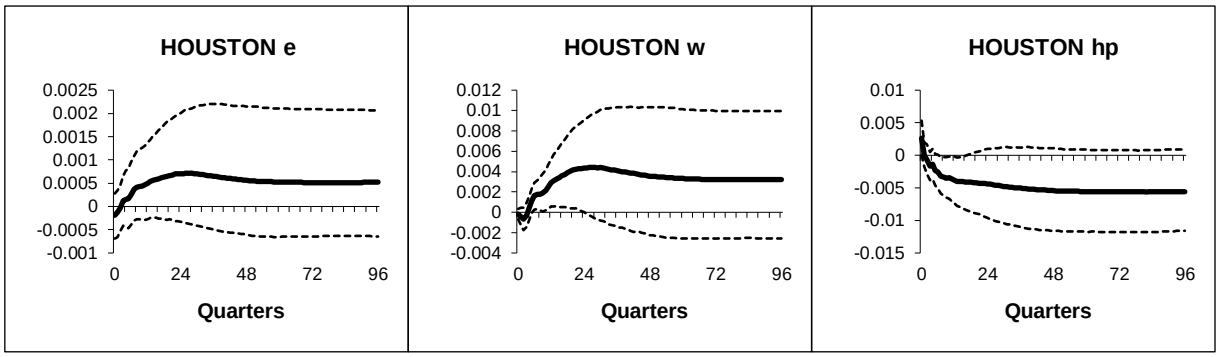

Figure A.1.18: Houston - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

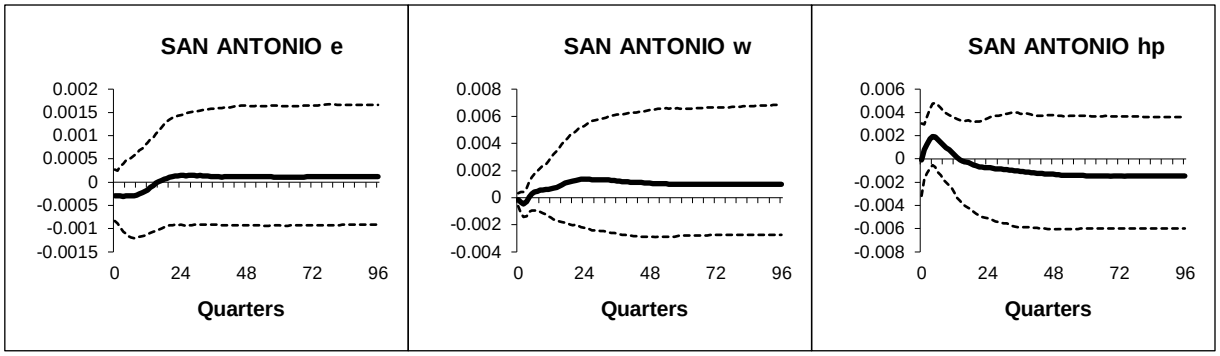

Figure A.1.19: San Antonio - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

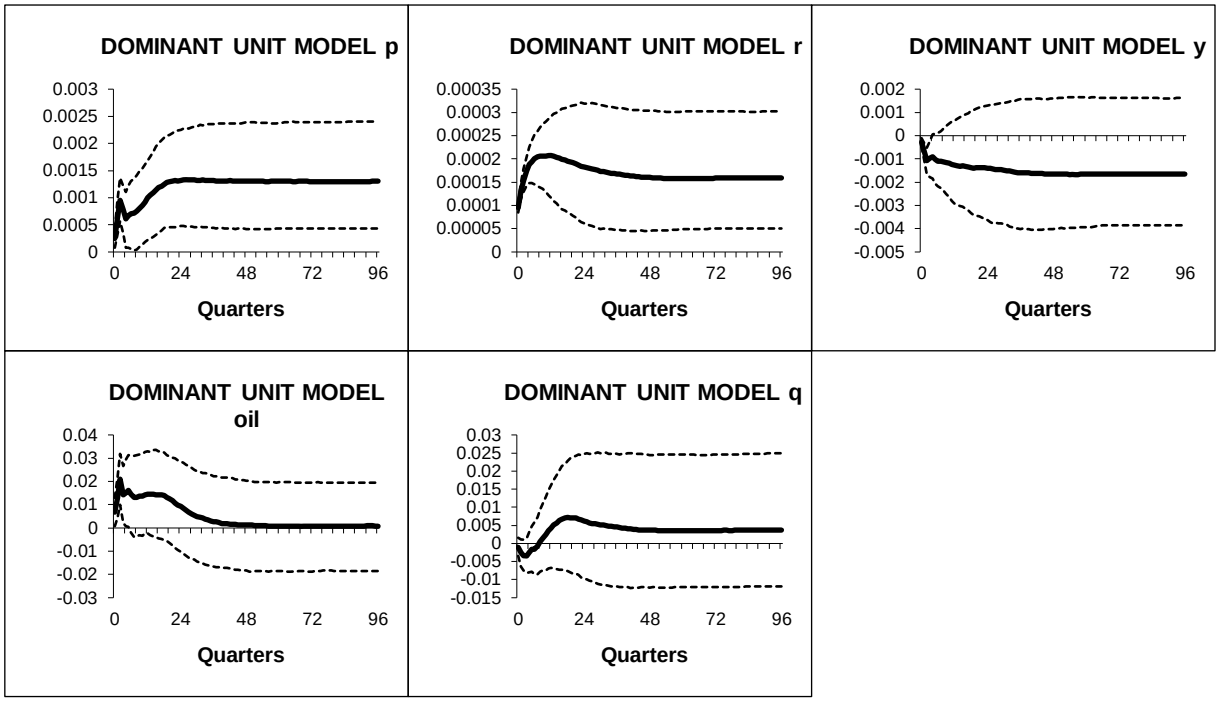

Figure A.1.20: DU - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

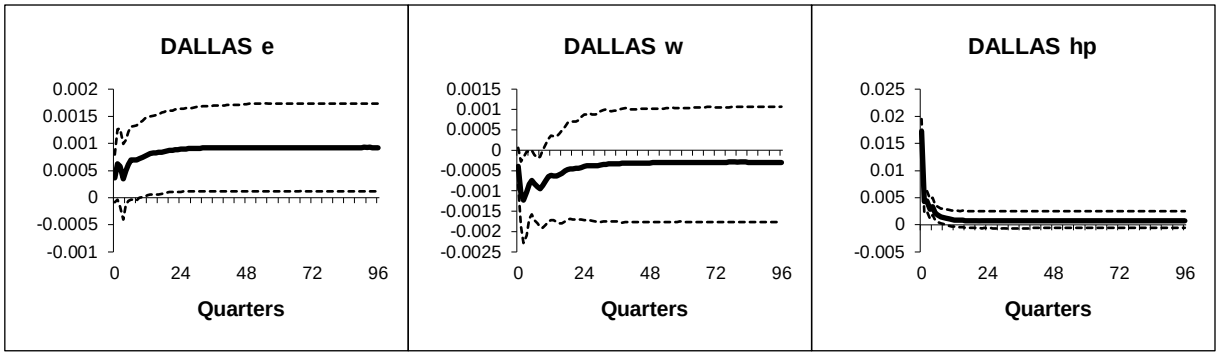

Figure A.1.21: Dallas - Generalized impulse responses of a positive (1 s.e.) shock to Dallas house prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

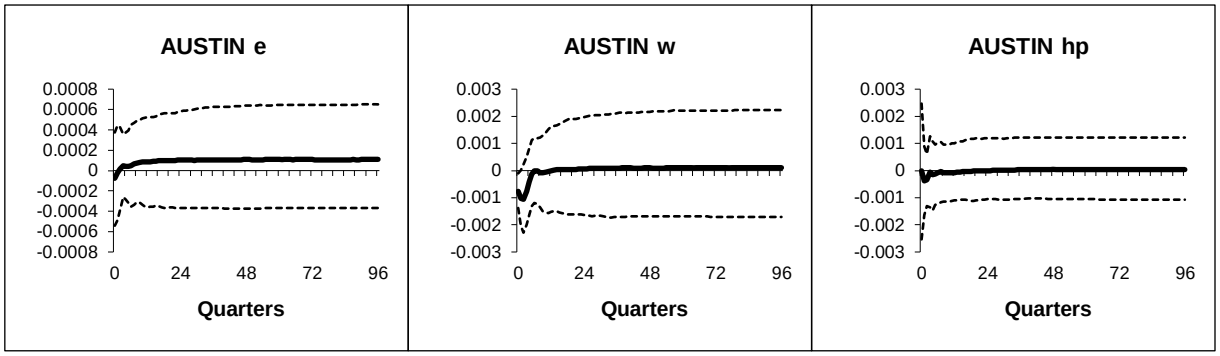

Figure A.1.22: Austin - Generalized impulse responses of a positive (1 s.e.) shock to Dallas house prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

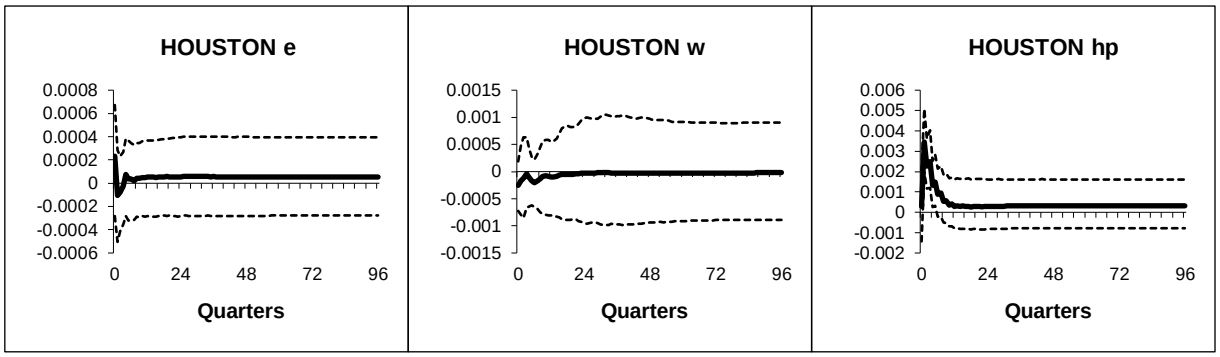

Figure A.1.23: Houston - Generalized impulse responses of a positive (1 s.e.) shock to Dallas house prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

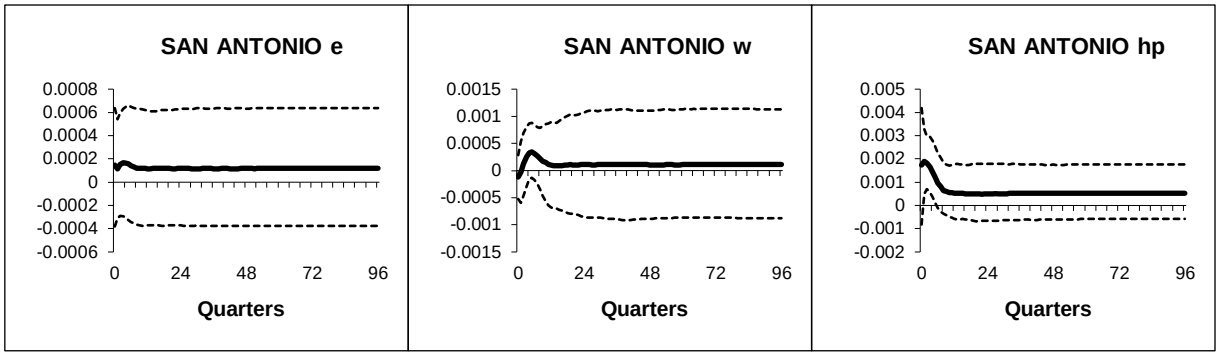

Figure A.1.24: San Antonio - Generalized impulse responses of a positive (1 s.e.) shock to Dallas house prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

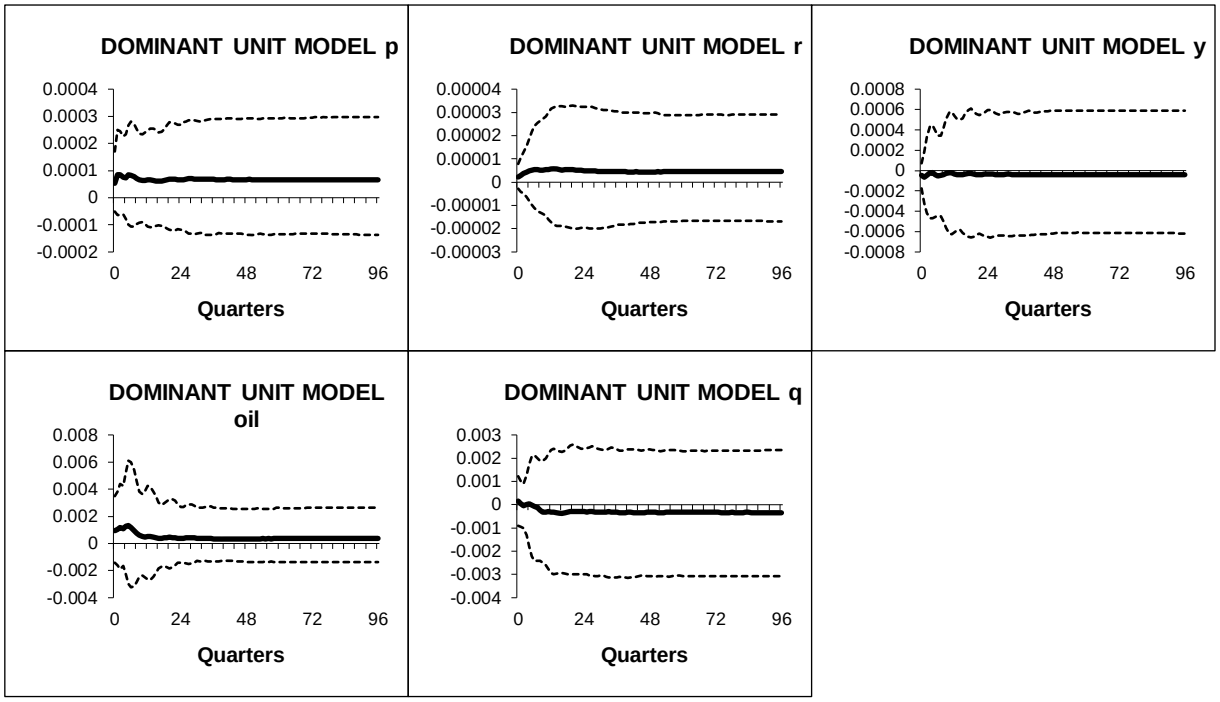

Figure A.1.25: DU - Generalized impulse responses of a positive (1 s.e.) shock to Dallas house prices (bootstrap mean estimates with 90% bootstrap error bounds). CX Automatic model GIRFs.

APPENDIX B.0

Table B.0.1: CV-CX Model Number of Rejections of the Null of Parameter Constancy Per Variable Across the Country-Specific Models at the 5% Level

Alternative test statistic	Domestic variables						Numbers(%)
	y	Δp	eq	ep	ρ^S	ρ^L	
PK_{sup}	4(17.4)	1(4.3)	1(4.3)	2(8.7)	2(8.7)	0(0.0)	10(8.2)
PK_{msq}	3(13.0)	1(4.3)	0(0.0)	2(8.7)	1(4.3)	0(0.0)	7(5.7)
$\Re$	3(13.0)	7(30.4)	5(21.7)	10(43.5)	7(30.4)	7(30.4)	39(32.0)
robust-N	0(0.0)	3(13.0)	2(8.7)	5(21.7)	3(13.0)	3(13.0)	16(13.1)
QLR	10(43.5)	12(52.2)	11(47.8)	16(69.6)	18(78.3)	12(52.2)	79(64.8)
robust-QLR	1(4.3)	4(17.4)	6(26.1)	9(39.1)	3(13.0)	6(26.1)	29(23.8)
MW	4(17.4)	10(43.5)	10(43.5)	13(56.5)	9(39.1)	8(34.8)	54(44.3)
robust-MW	1(4.3)	5(21.7)	5(21.7)	8(34.8)	3(13.0)	5(21.7)	27(22.1)
APW	10(43.5)	12(52.2)	11(47.8)	15(65.2)	18(78.3)	12(52.2)	78(63.9)
robust-APW	3(13.0)	4(17.4)	6(26.1)	9(39.1)	3(13.0)	6(26.1)	31(25.4)

Note: The test statistics PK_{sup} and PK_{mwq} are based on the cumulative sums of OLS residuals, $\Re$ is the Nyblom test for time-varying parameters and QLR, MW, and APW are the sequential Wald statistics for a single break at an unknown change point. Statistics with the prefix 'robust' denote the heteroskedasticity-robust version of the tests. All tests are implemented at the 5% significance level.

Table B.0.2: CV-CX Model Contemporaneous Effects of Foreign Variables on Their Domestic Counterparts

Country	Domestic variables				
	y	Δp	eq	ρ^S	ρ^L
USA	0.56	0.13			
	[5.17]	[1.13]			
Euro Area	0.48	0.30	1.12	0.44	0.61
	[5.55]	[4.78]	[19.91]	[4.94]	[8.13]
China	0.66	0.72		0.06	
	[2.93]	[3.16]		[0.61]	
Japan	0.53	0.15	0.79	-0.06	0.57
	[2.97]	[1.91]	[7.61]	[-0.45]	[5.21]
UK	0.57	0.59	0.89	0.68	0.76
	[4.62]	[4.70]	[18.16]	[3.02]	[8.05]
Sweden	1.28	0.74	1.25	0.62	0.75
	[5.52]	[3.85]	[15.80]	[1.57]	[4.87]
Switzerland	0.45	0.29	1.00	0.10	0.49
	[3.07]	[2.27]	[18.81]	[0.67]	[7.34]
Norway	0.27	0.81	1.06	0.03	0.53
	[1.25]	[3.01]	[9.96]	[0.08]	[3.57]

Note: White's heteroskedastic-robust t-ratios are given in square brackets

Table B.0.3: Trade Weights

Countries	austlia	can	china	chl	euro	india	indns	japan	kor	mal	mex	nor	nzld	phlp	safrc	sarbia	sing	swe	switz	thai	turk	uk	usa
austlia	0.00	0.00	0.05	0.01	0.02	0.03	0.03	0.06	0.04	0.04	0.00	0.00	0.23	0.02	0.02	0.01	0.04	0.01	0.01	0.04	0.01	0.02	0.01
can	0.01	0.00	0.02	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.03	0.02	0.02	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.03	0.21
china	0.28	0.08	0.00	0.28	0.16	0.16	0.16	0.27	0.29	0.17	0.10	0.05	0.18	0.14	0.19	0.17	0.14	0.05	0.05	0.17	0.12	0.08	0.18
chl	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
euro	0.09	0.06	0.17	0.17	0.00	0.18	0.08	0.10	0.08	0.10	0.07	0.43	0.10	0.09	0.26	0.15	0.10	0.55	0.65	0.08	0.48	0.51	0.15
india	0.04	0.01	0.03	0.03	0.03	0.00	0.05	0.01	0.03	0.04	0.01	0.00	0.02	0.01	0.06	0.09	0.04	0.01	0.01	0.02	0.03	0.02	0.02
indns	0.03	0.00	0.03	0.00	0.01	0.05	0.00	0.04	0.04	0.06	0.00	0.00	0.02	0.04	0.01	0.02	0.11	0.00	0.00	0.05	0.01	0.00	0.01
japan	0.16	0.03	0.14	0.10	0.04	0.04	0.16	0.00	0.14	0.14	0.03	0.02	0.08	0.18	0.08	0.14	0.08	0.02	0.03	0.20	0.02	0.02	0.07
kor	0.07	0.01	0.11	0.06	0.02	0.04	0.09	0.08	0.00	0.05	0.02	0.02	0.04	0.08	0.03	0.11	0.07	0.01	0.01	0.04	0.03	0.01	0.03
mal	0.03	0.00	0.04	0.00	0.01	0.03	0.07	0.04	0.02	0.00	0.01	0.00	0.03	0.04	0.02	0.01	0.16	0.00	0.00	0.07	0.01	0.01	0.01
mex	0.01	0.04	0.01	0.04	0.02	0.01	0.00	0.01	0.02	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.16
nor	0.00	0.01	0.00	0.00	0.04	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.01	0.04	0.00
nzld	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
phlp	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.02	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.02	0.00	0.00	0.02	0.00	0.00	0.01
safrc	0.01	0.00	0.02	0.00	0.02	0.03	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.01	0.01	0.01	0.02	0.01
sarbia	0.00	0.00	0.03	0.00	0.02	0.08	0.02	0.05	0.06	0.01	0.00	0.00	0.02	0.04	0.05	0.00	0.03	0.01	0.01	0.03	0.03	0.01	0.02
sing	0.05	0.00	0.03	0.00	0.02	0.05	0.14	0.03	0.04	0.16	0.00	0.01	0.04	0.10	0.01	0.04	0.00	0.00	0.01	0.05	0.00	0.01	0.02
swe	0.01	0.00	0.01	0.01	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.02	0.03	0.01
switz	0.01	0.01	0.01	0.01	0.10	0.07	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.00	0.04	0.03	0.05	0.02
thai	0.04	0.00	0.03	0.01	0.01	0.02	0.05	0.05	0.02	0.07	0.01	0.00	0.03	0.06	0.02	0.02	0.04	0.01	0.01	0.00	0.01	0.01	0.01
turk	0.00	0.00	0.01	0.01	0.05	0.01	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.01
uk	0.03	0.03	0.03	0.02	0.19	0.03	0.01	0.02	0.01	0.01	0.01	0.23	0.04	0.01	0.05	0.01	0.02	0.10	0.05	0.02	0.07	0.00	0.04
usa	0.09	0.68	0.20	0.23	0.15	0.13	0.08	0.17	0.13	0.11	0.69	0.06	0.11	0.16	0.11	0.17	0.11	0.06	0.10	0.10	0.10	0.11	0.00

Table B.0.4: Convex Weights - 0.4∗Portfolio Investment + 0.6∗Distance

Countries	austlia	can	china	chl	euro	india	indns	japan	kor	mal	mex	nor	nzld	phlp	safrc	sarbia	sing	swe	switz	thai	turk	uk	usa
austlia	0.00	0.02	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.08	0.06	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.04
can	0.03	0.00	0.03	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.02	0.03	0.02	0.03	0.04	0.02	0.03	0.07
china	0.09	0.09	0.00	0.09	0.11	0.09	0.09	0.10	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.12
chl	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
euro	0.17	0.15	0.18	0.16	0.00	0.16	0.16	0.20	0.16	0.16	0.17	0.24	0.14	0.16	0.17	0.27	0.16	0.28	0.19	0.16	0.19	0.24	0.25
india	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03
indns	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
japan	0.14	0.09	0.09	0.07	0.14	0.08	0.09	0.00	0.09	0.08	0.09	0.13	0.12	0.09	0.08	0.06	0.10	0.10	0.08	0.08	0.09	0.10	0.19
kor	0.01	0.01	0.03	0.01	0.02	0.02	0.01	0.02	0.00	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.02
mal	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.04	0.00	0.00	0.01	0.00	0.00	0.01
mex	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
nor	0.01	0.01	0.01	0.02	0.03	0.01	0.01	0.02	0.02	0.01	0.01	0.00	0.01	0.04	0.02	0.01	0.01	0.04	0.02	0.02	0.02	0.02	0.02
nzld	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
phlp	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
safrc	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.01
sarbia	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01
sing	0.02	0.00	0.09	0.00	0.01	0.11	0.09	0.00	0.05	0.10	0.00	0.00	0.01	0.05	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.01	0.02
swe	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.04	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.01	0.02
switz	0.02	0.02	0.01	0.01	0.04	0.01	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.03	0.00	0.01	0.01	0.02	0.02
thai	0.01	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
turk	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01
uk	0.07	0.04	0.08	0.07	0.14	0.08	0.06	0.10	0.08	0.08	0.07	0.07	0.07	0.06	0.07	0.24	0.08	0.06	0.07	0.07	0.13	0.00	0.11
usa	0.34	0.47	0.35	0.44	0.33	0.33	0.34	0.39	0.35	0.33	0.42	0.27	0.33	0.33	0.41	0.16	0.36	0.27	0.38	0.35	0.34	0.34	0.00

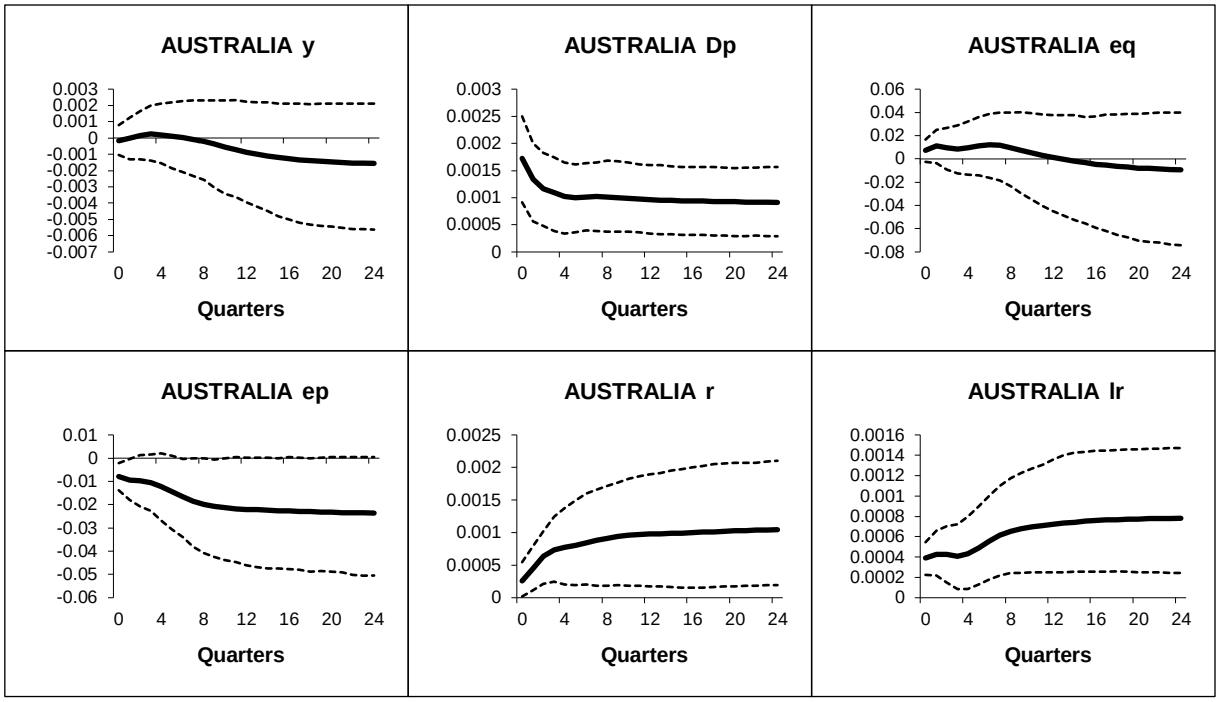

Figure B.1.24: Australia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

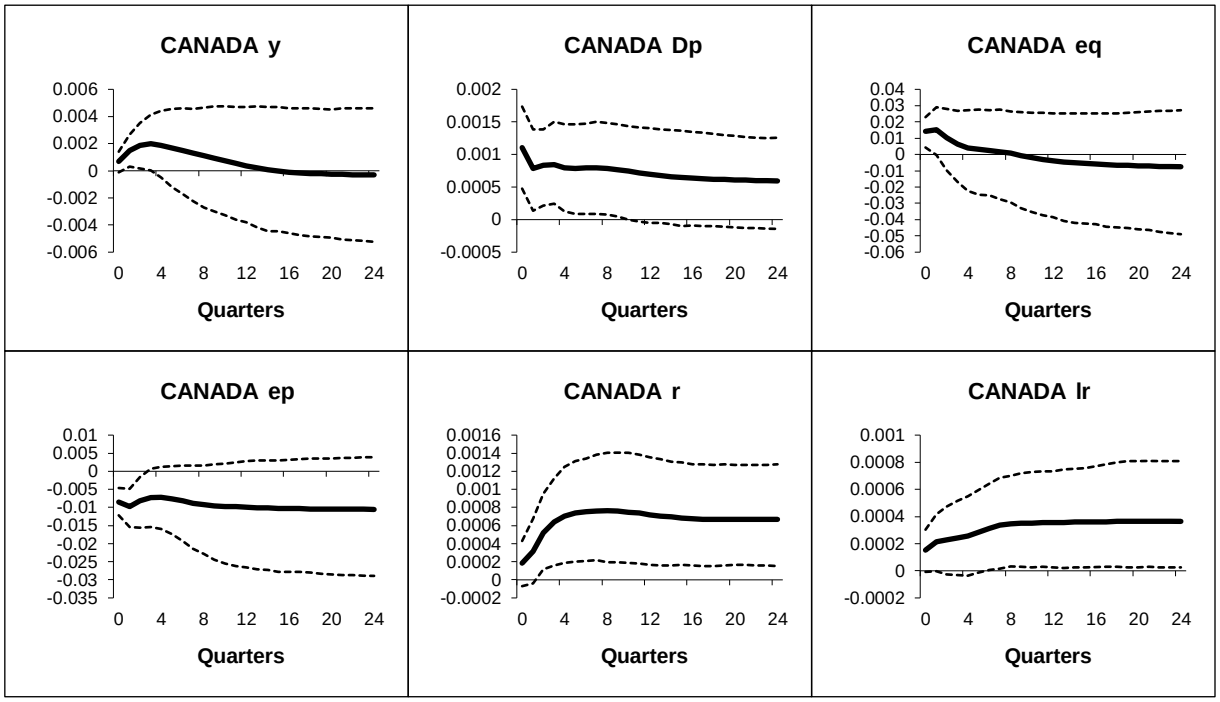

Figure B.1.25: Canada - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

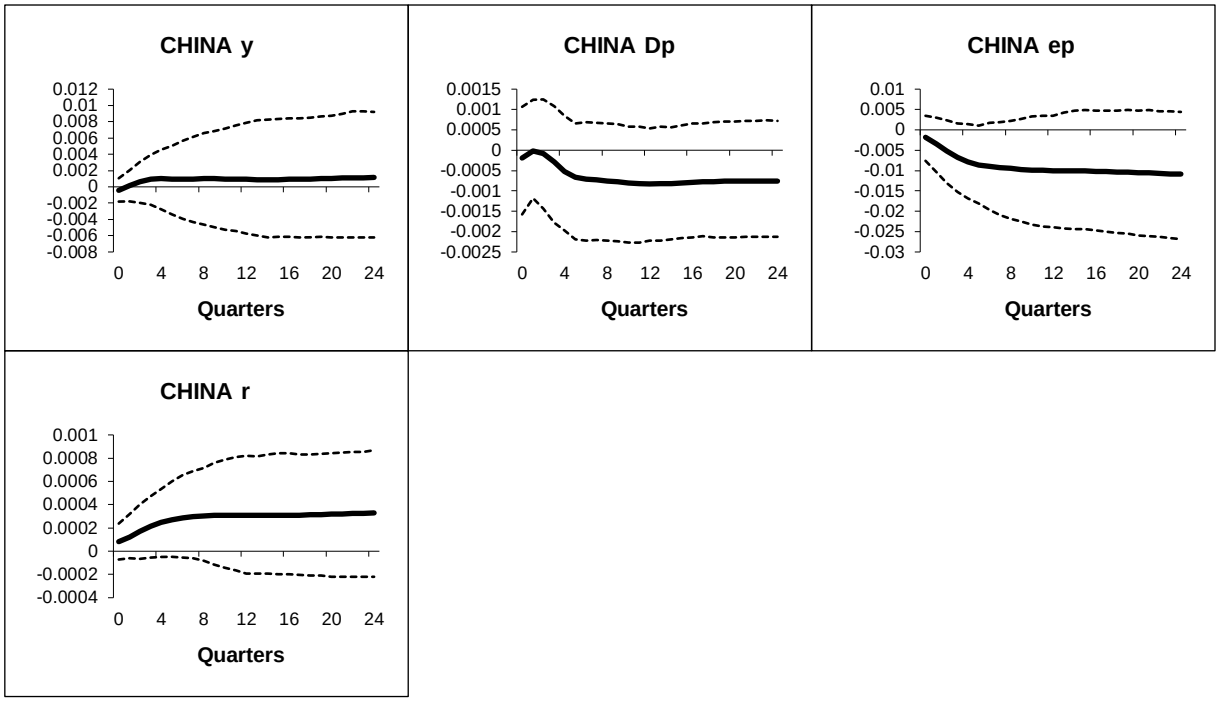

Figure B.1.26: China - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

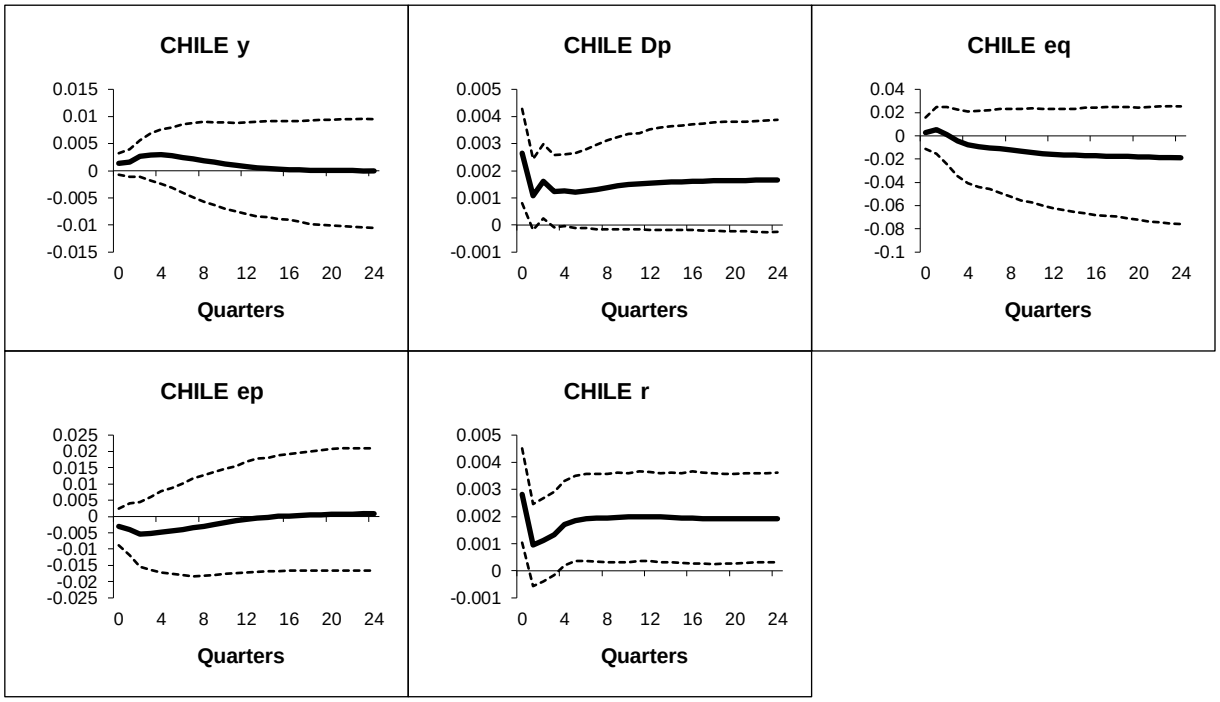

Figure B.1.27: Chile - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

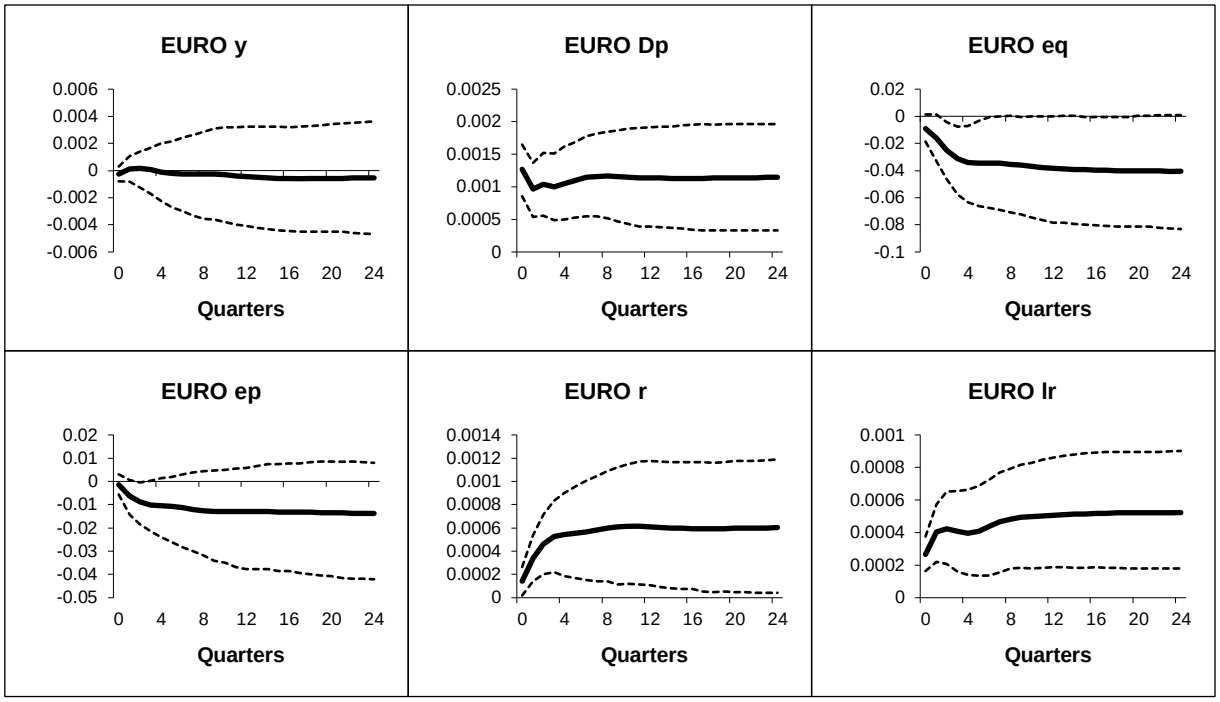

Figure B.1.28: Euro - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

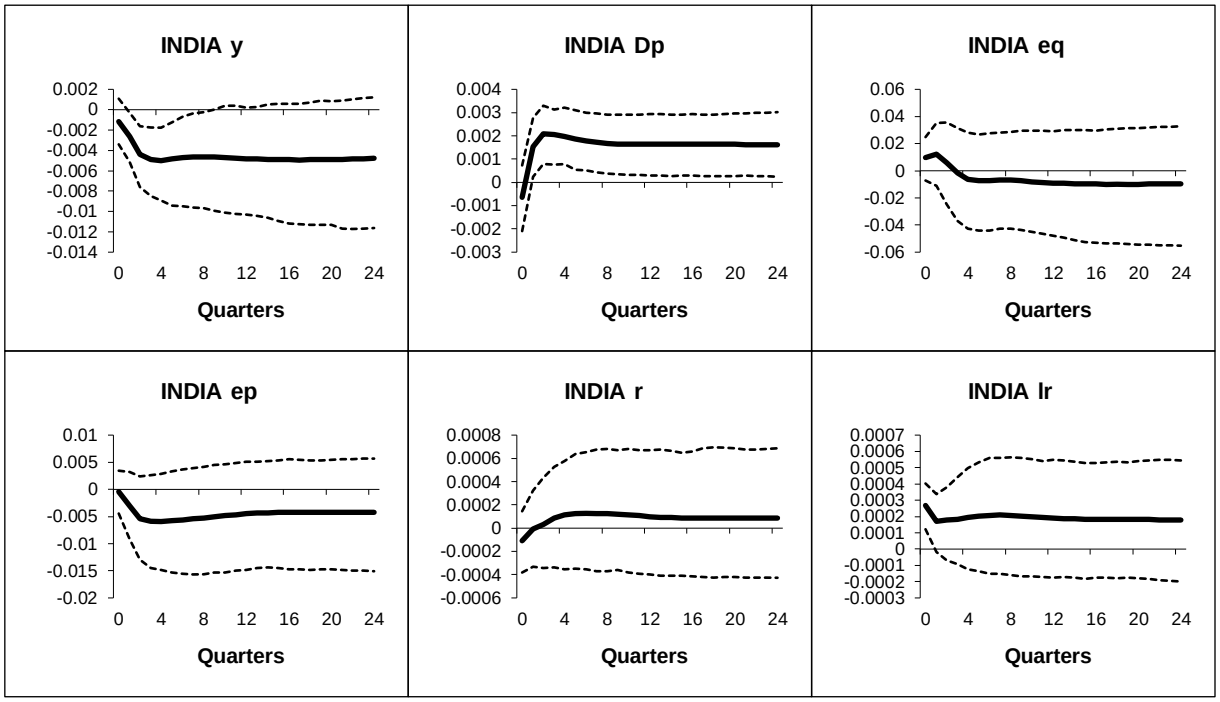

Figure B.1.29: India - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

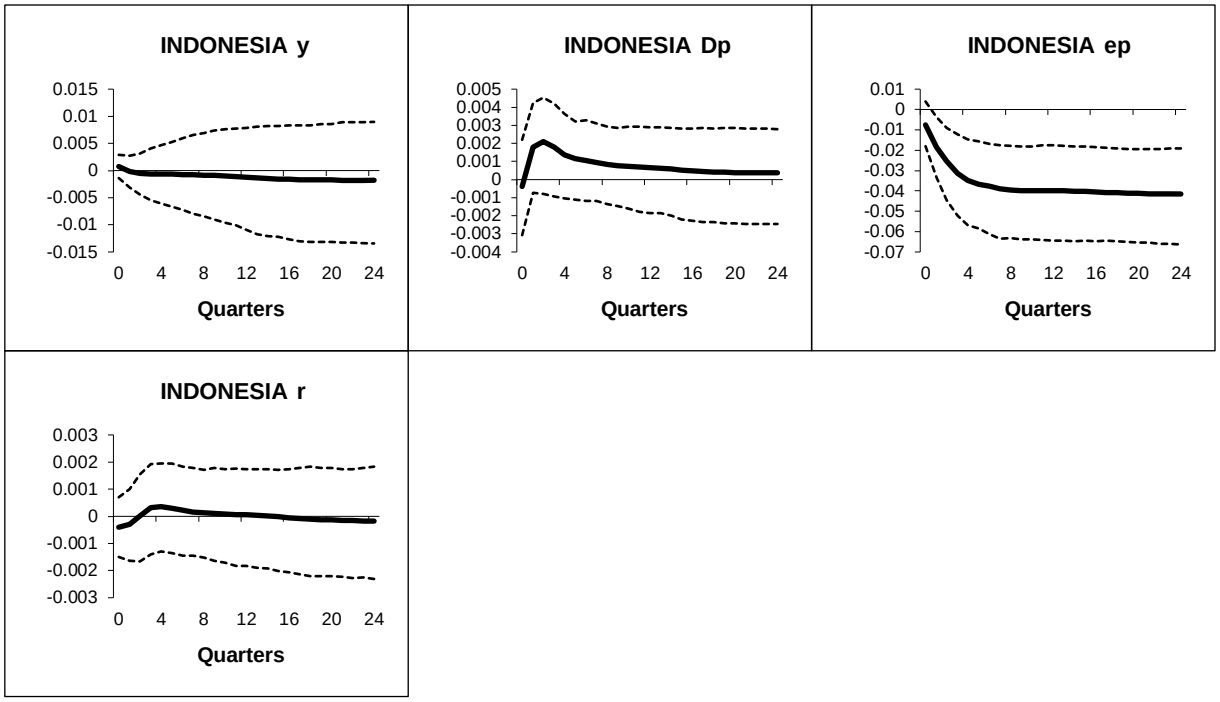

Figure B.1.30: Indonesia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

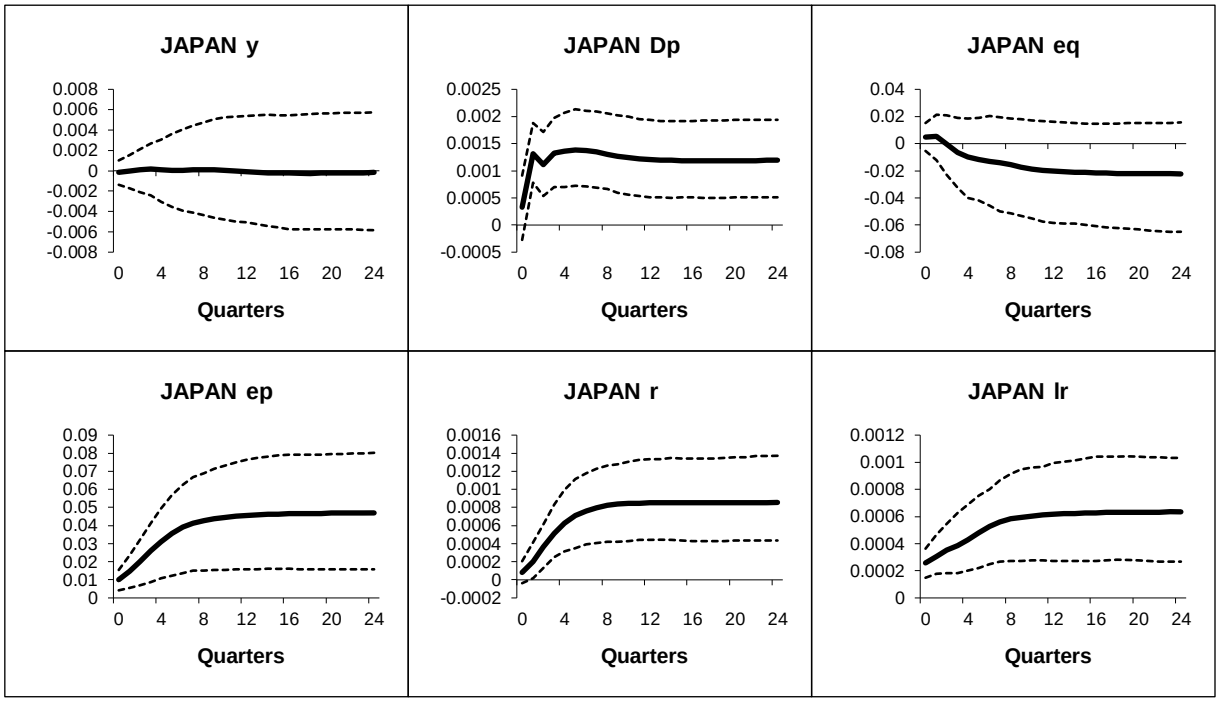

Figure B.1.31: Japan - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

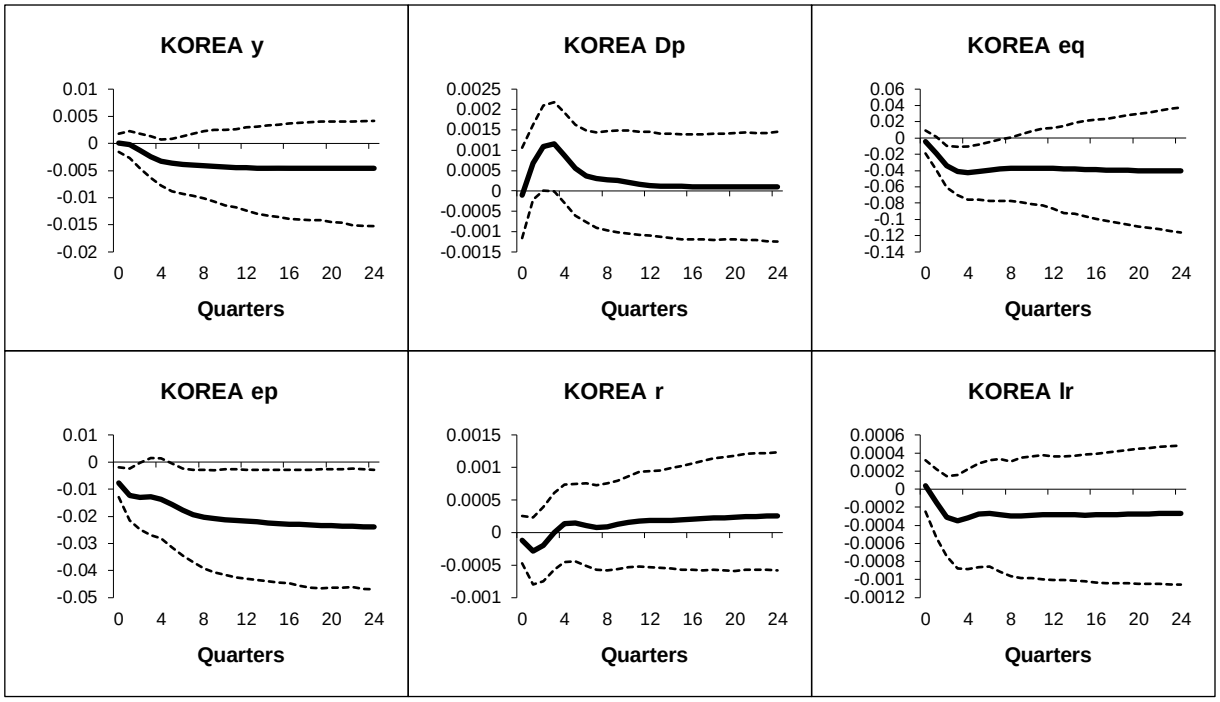

Figure B.1.32: Korea - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

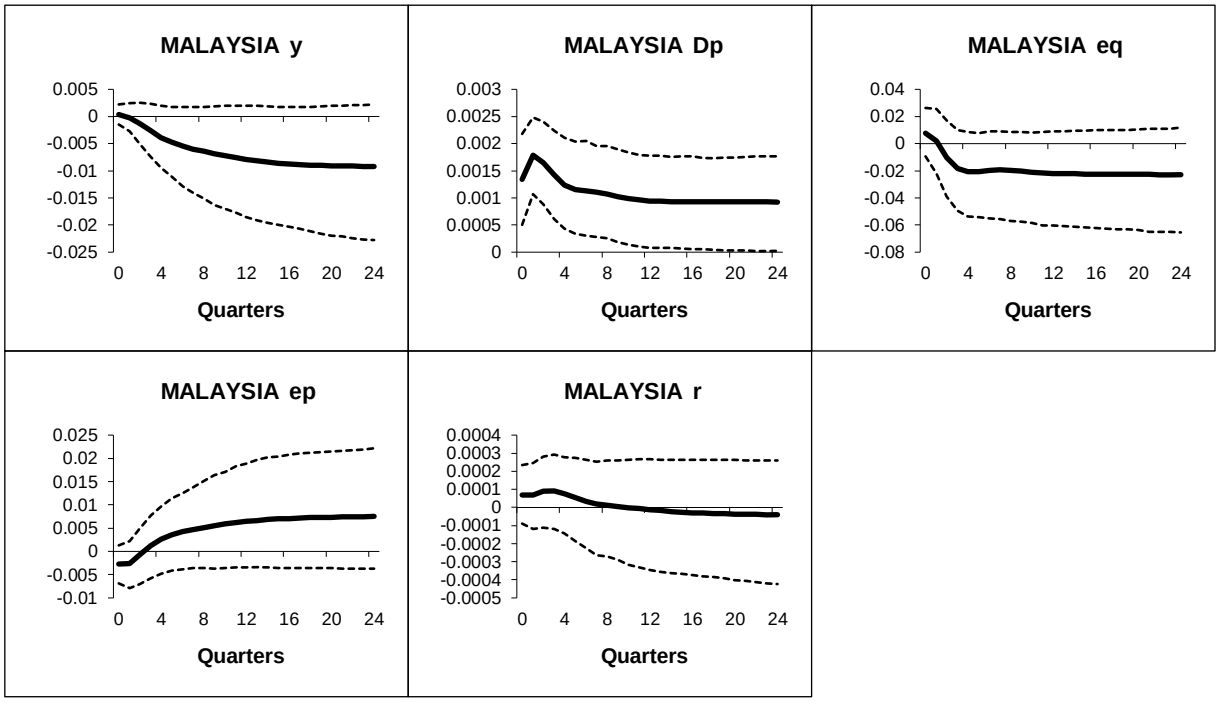

Figure B.1.33: Malaysia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

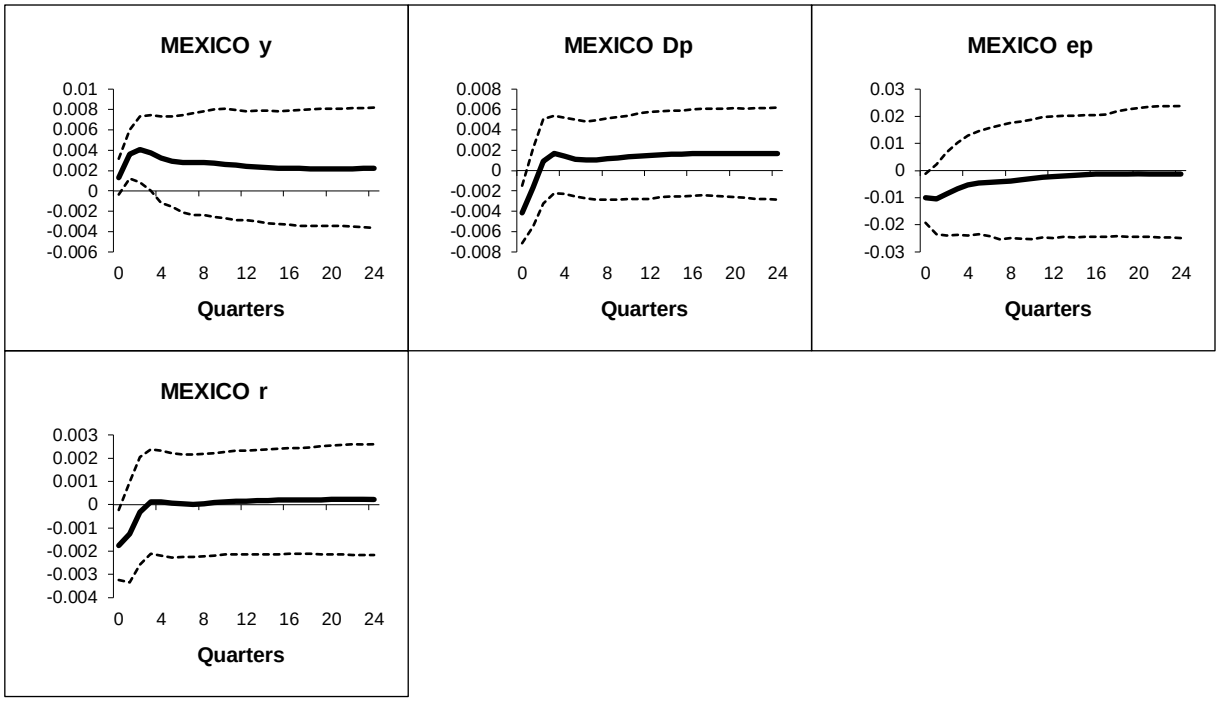

Figure B.1.34: Mexico - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

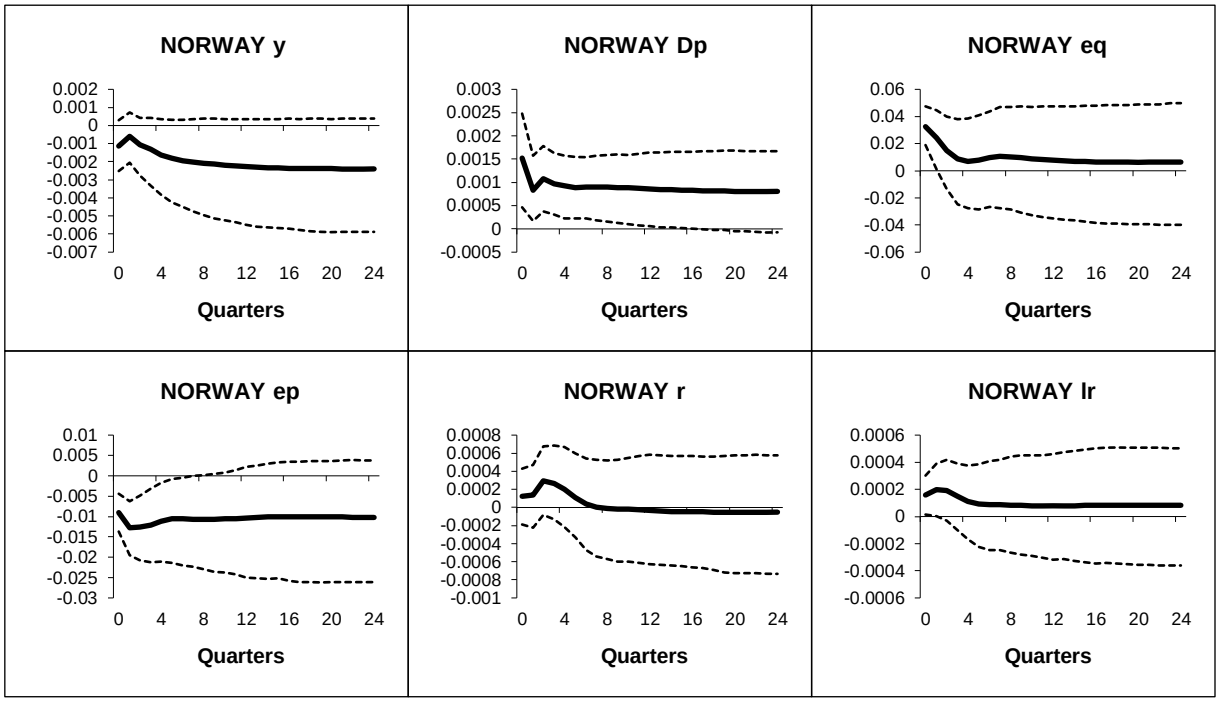

Figure B.1.35: Norway - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

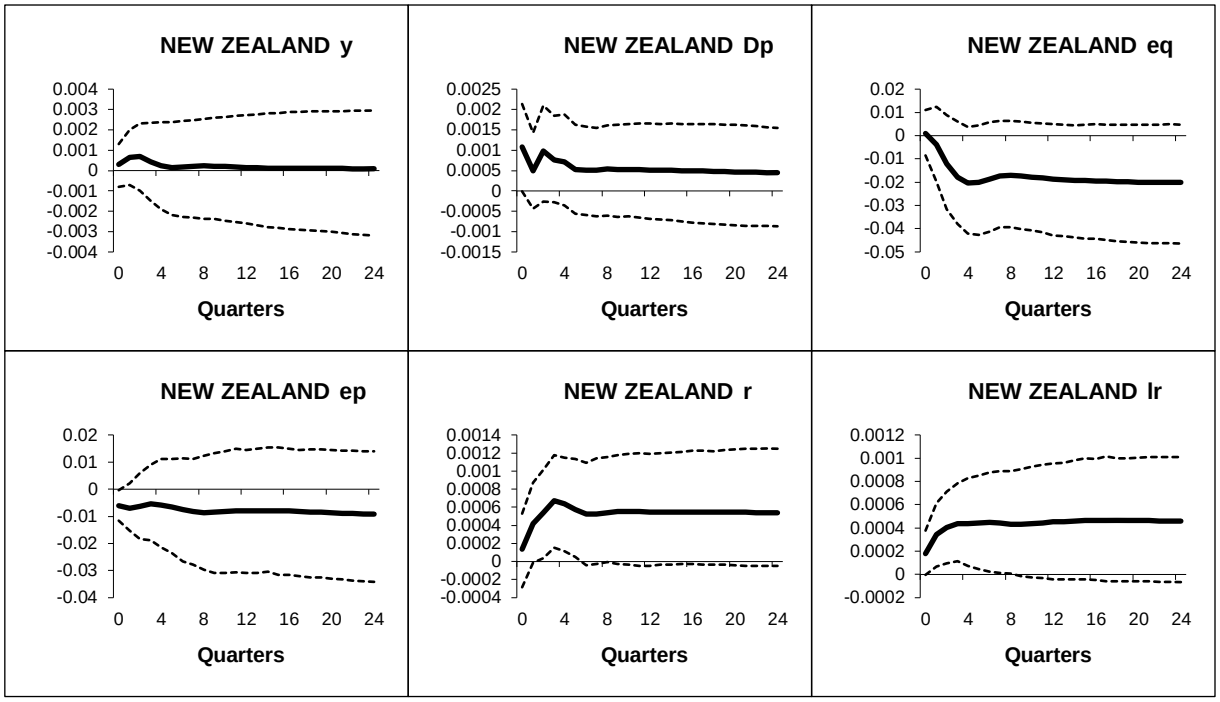

Figure B.1.36: New Zealand - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

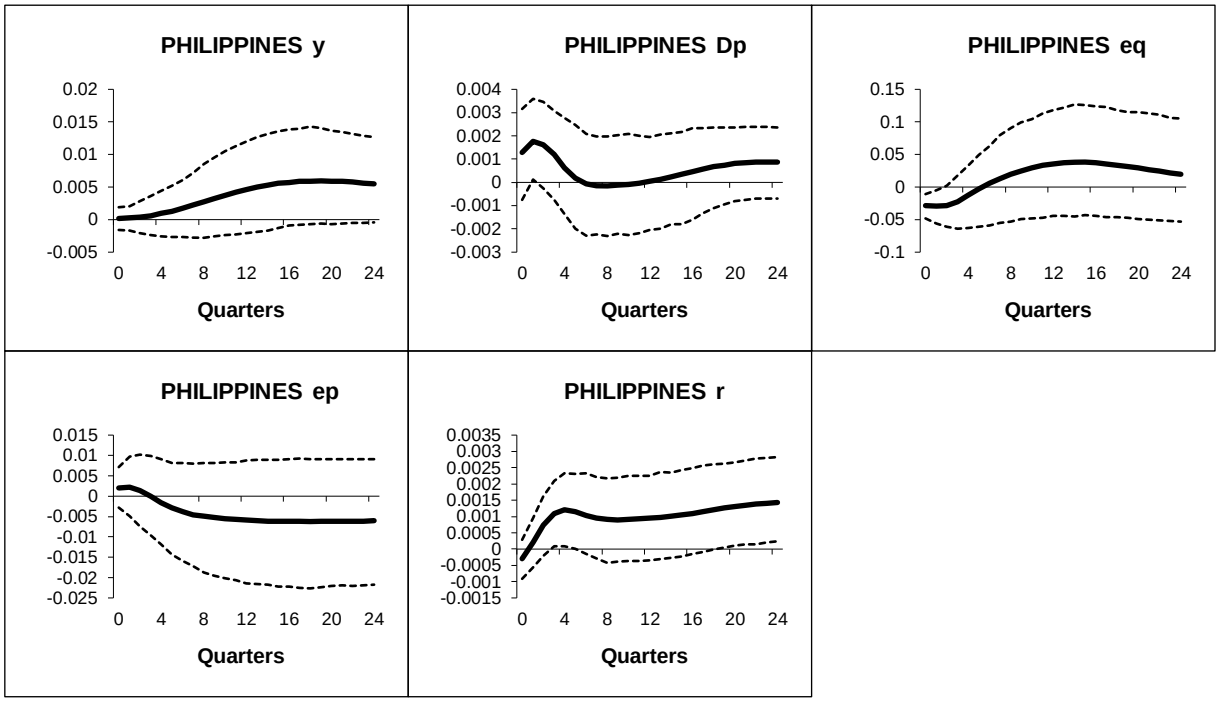

Figure B.1.37: Philippines - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

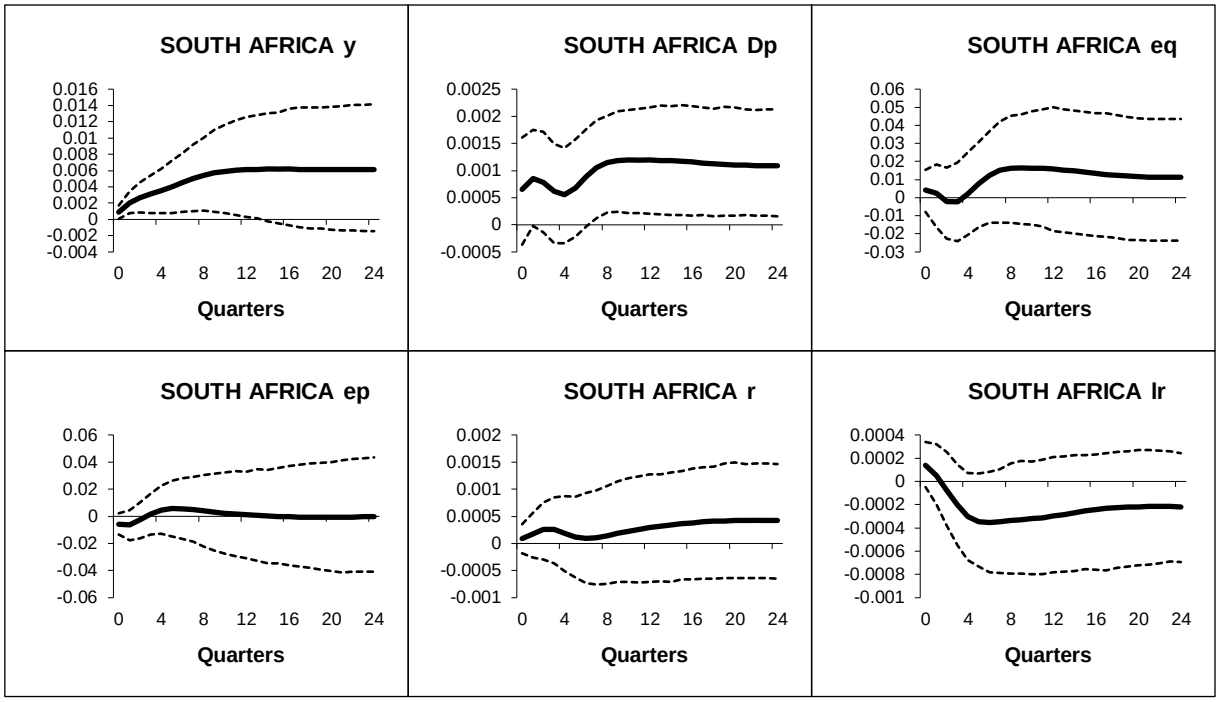

Figure B.1.38: South Africa - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

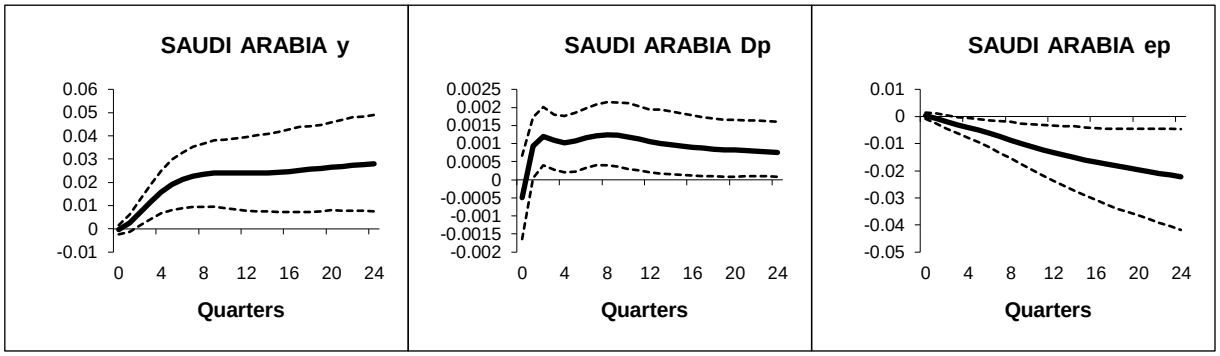

Figure B.1.39: Saudi Arabia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

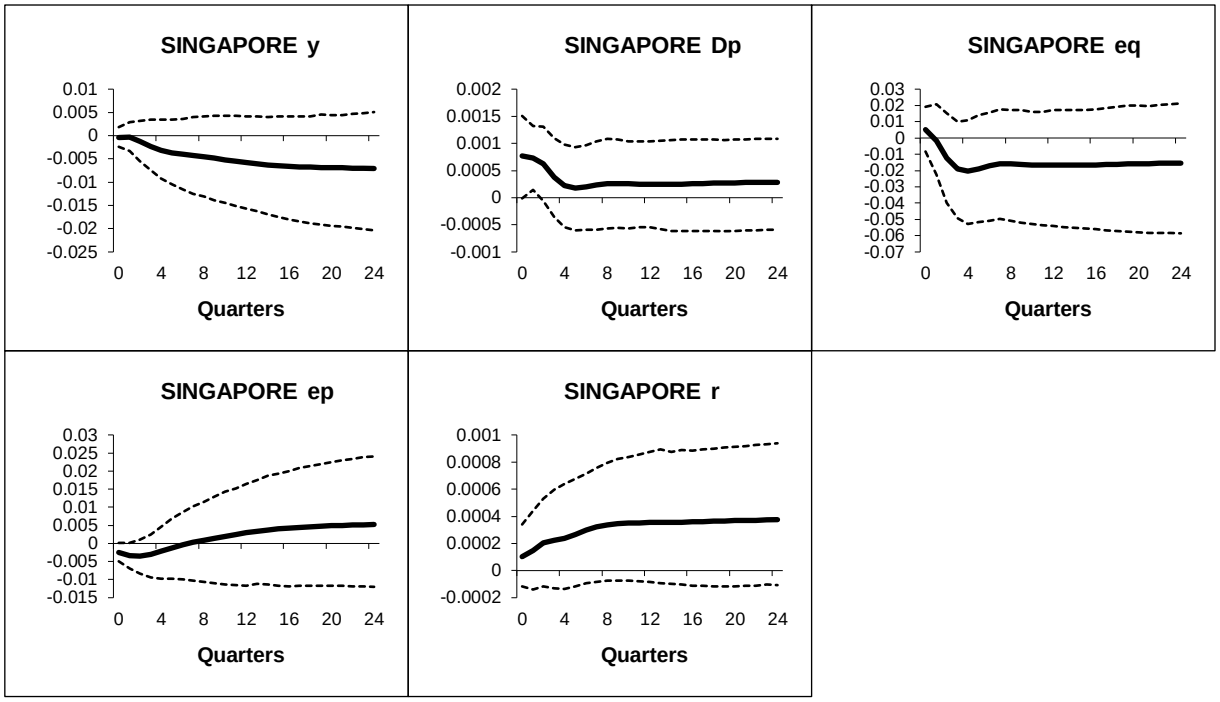

Figure B.1.40: Singapore - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

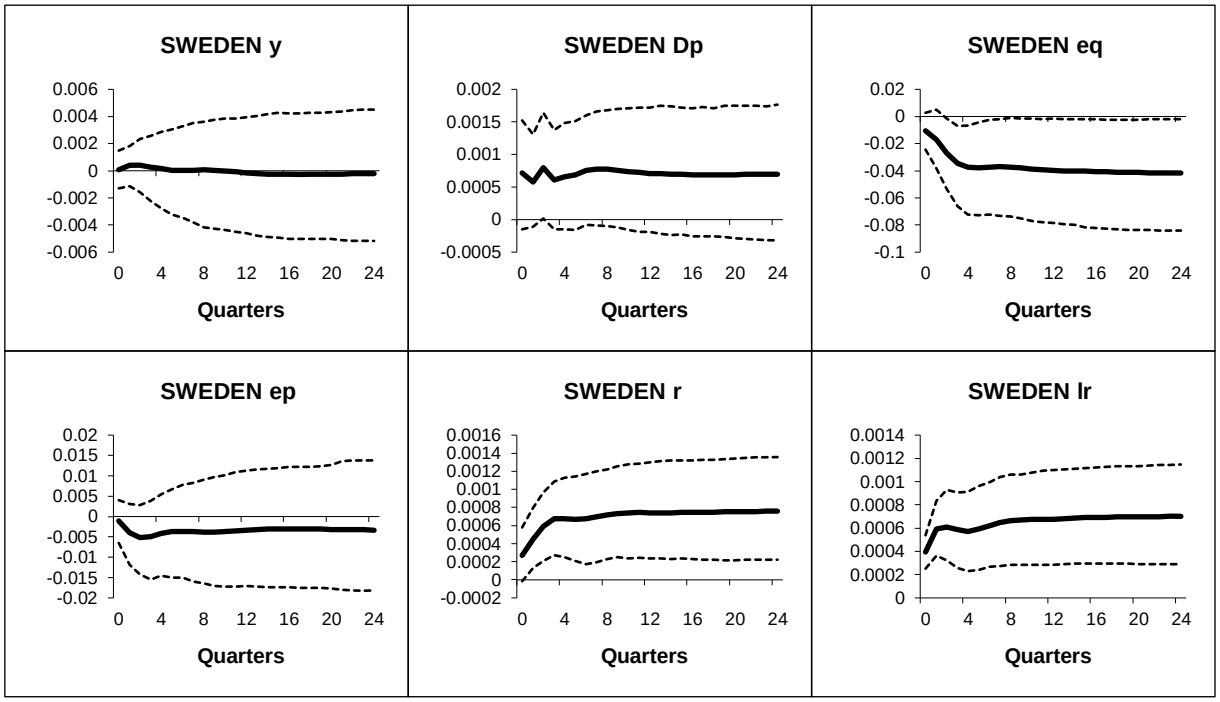

Figure B.1.41: Sweden - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

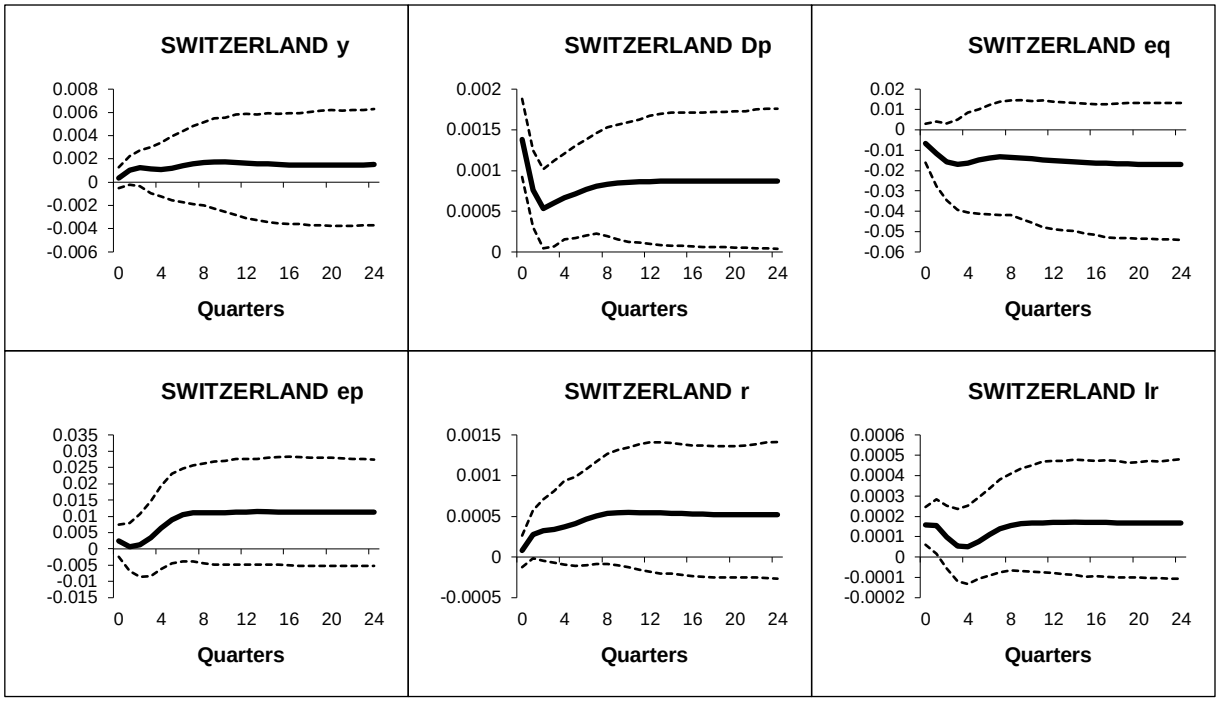

Figure B.1.42: Switzerland - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

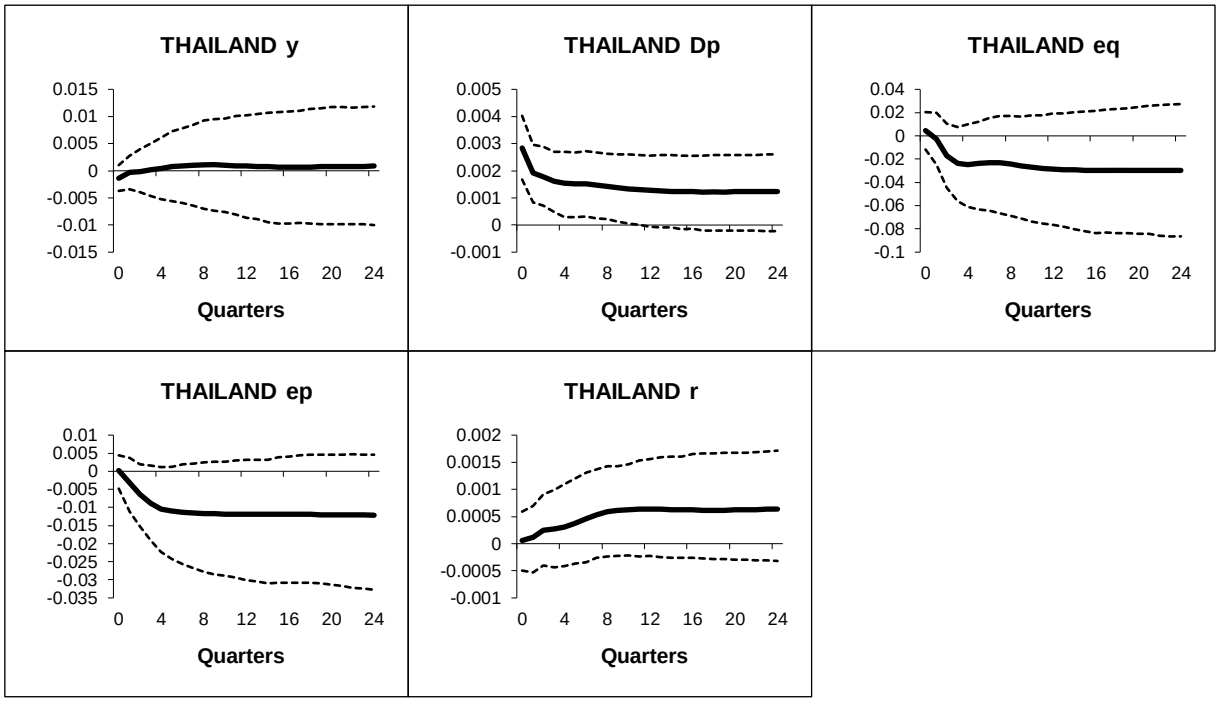

Figure B.1.43: Thailand - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

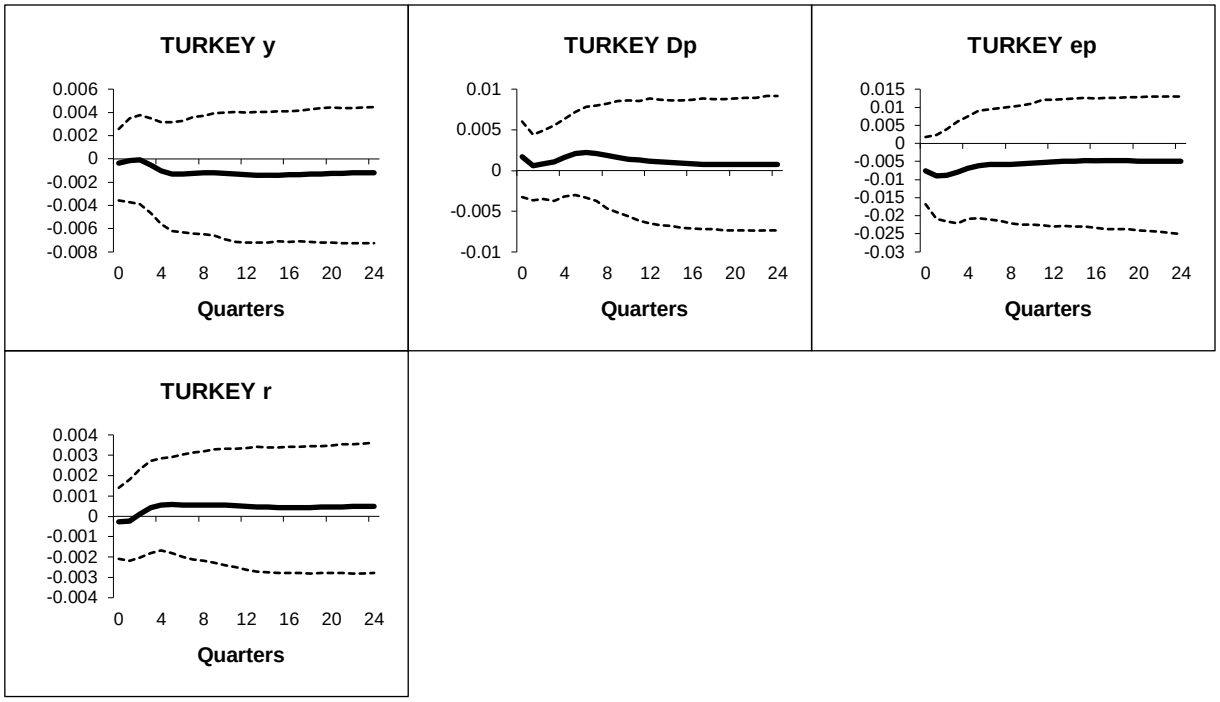

Figure B.1.44: Turkey - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

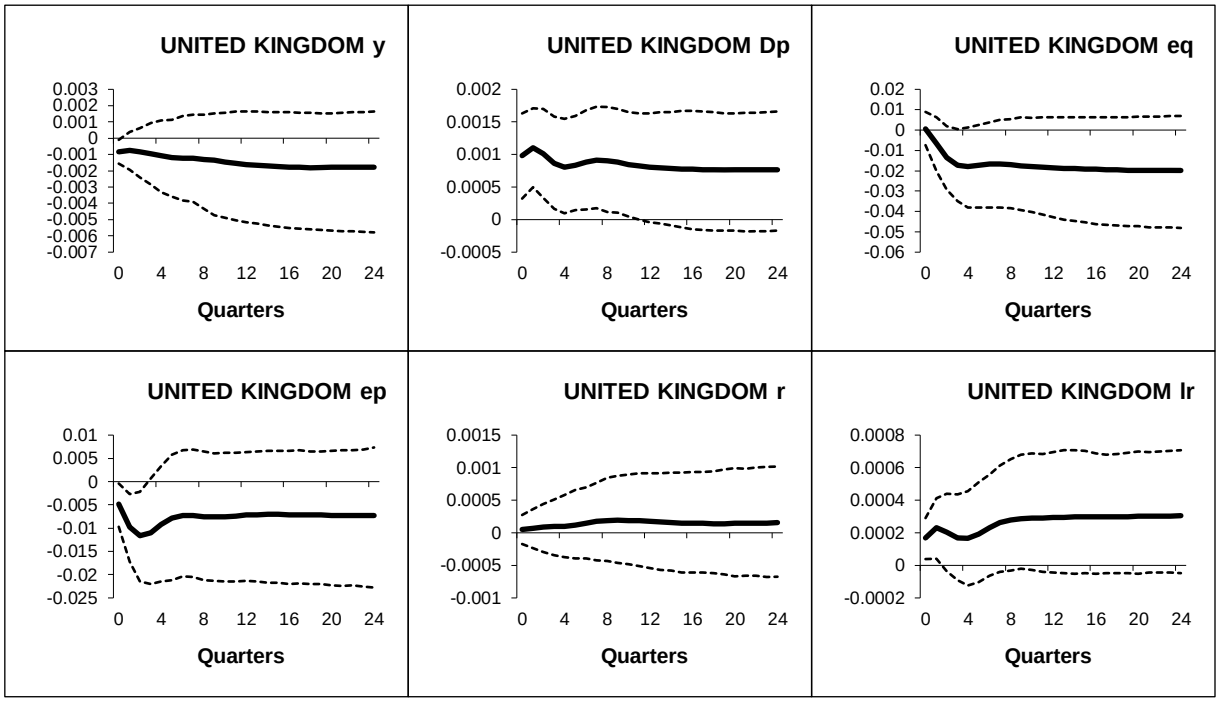

Figure B.1.45: UK - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

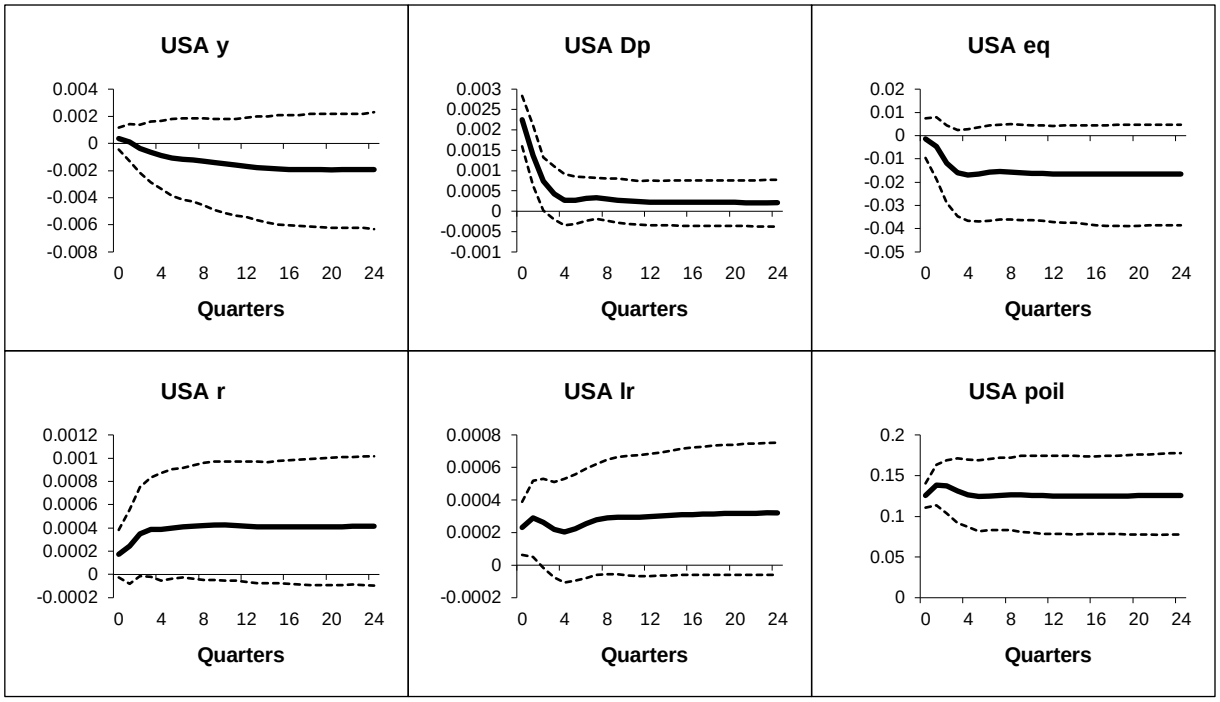

Figure B.1.46: USA - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

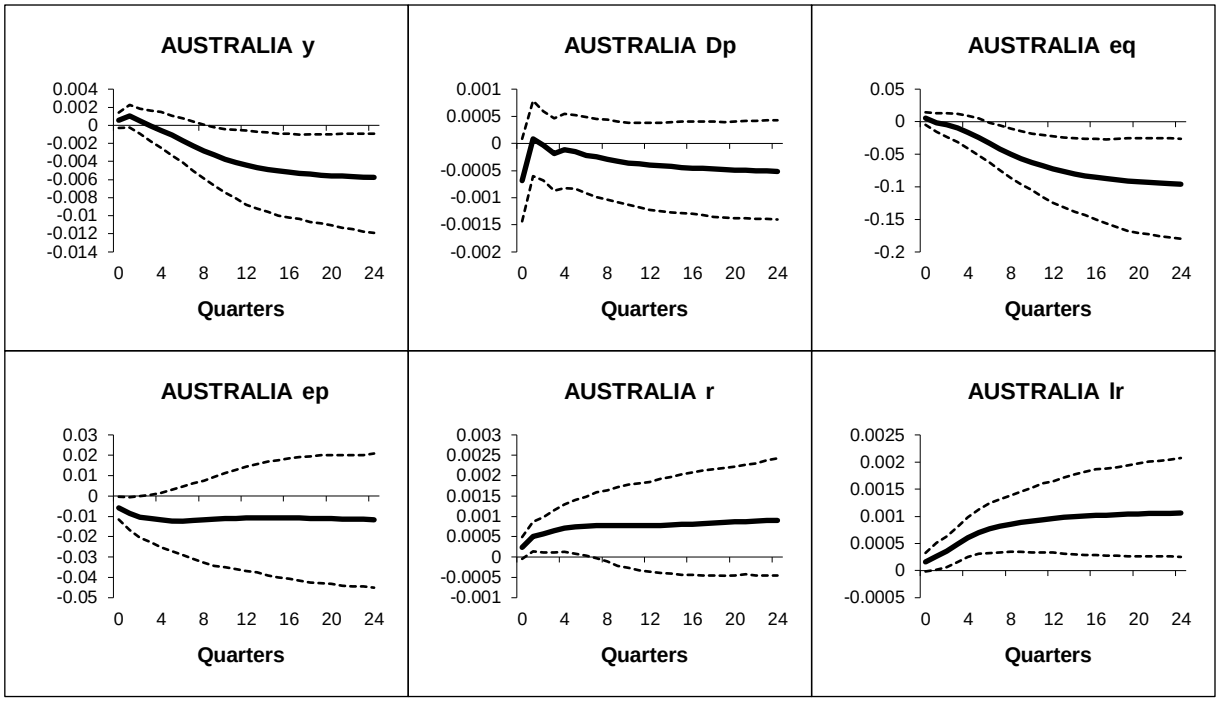

Figure B.1.47: Australia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

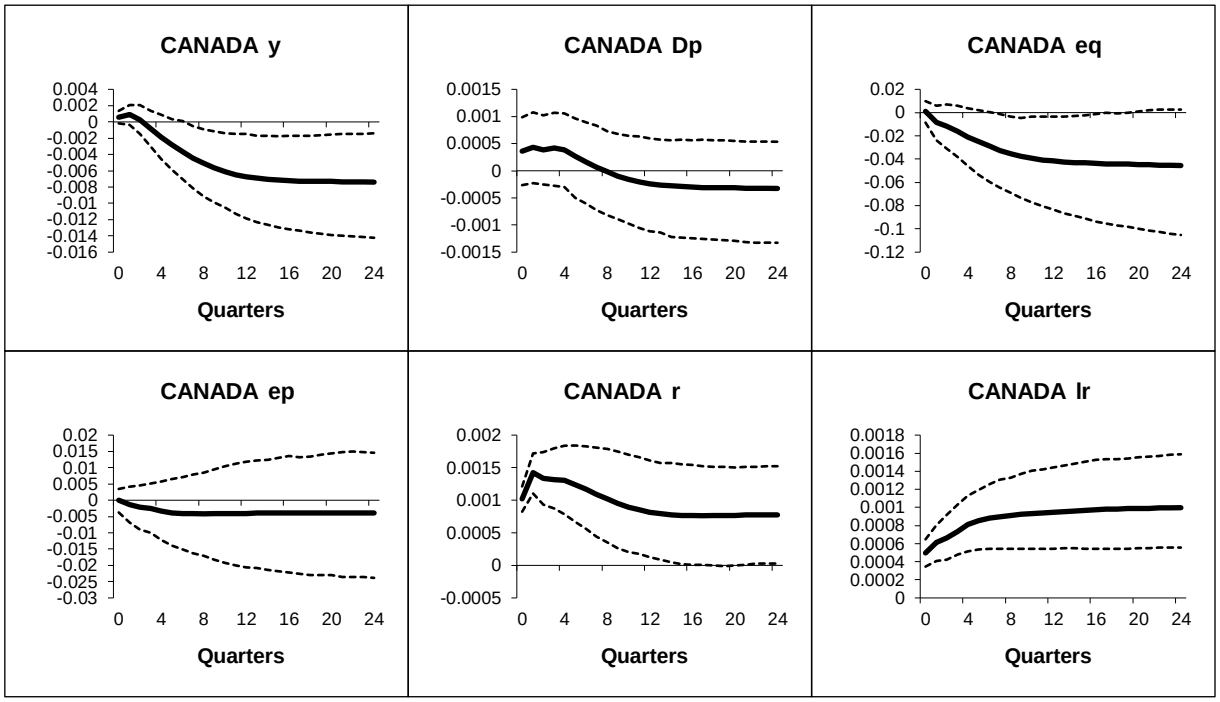

Figure B.1.48: Canada - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

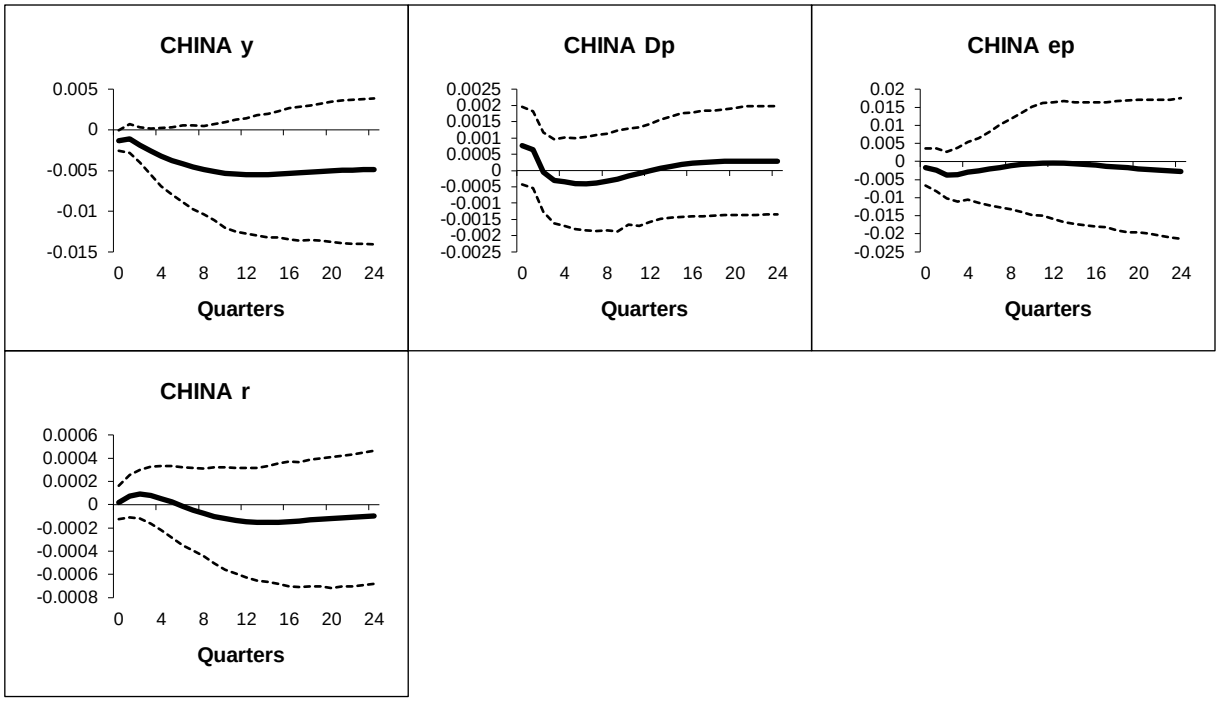

Figure B.1.49: China - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

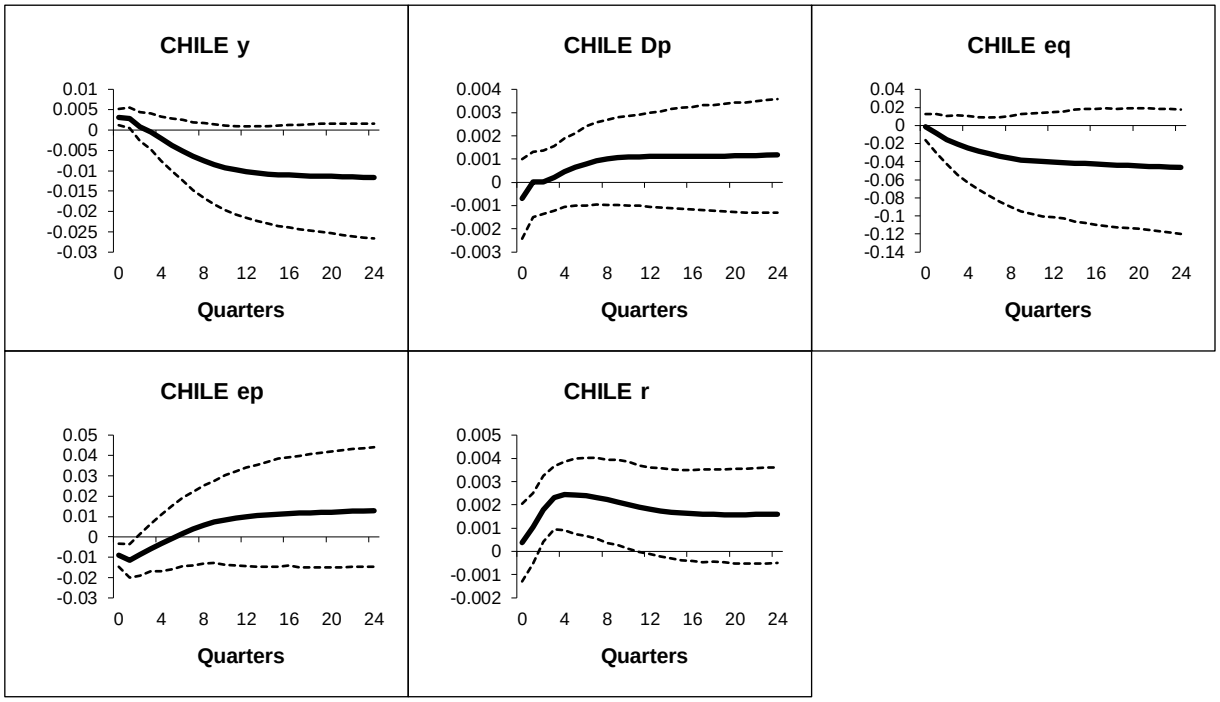

Figure B.1.50: Chile - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

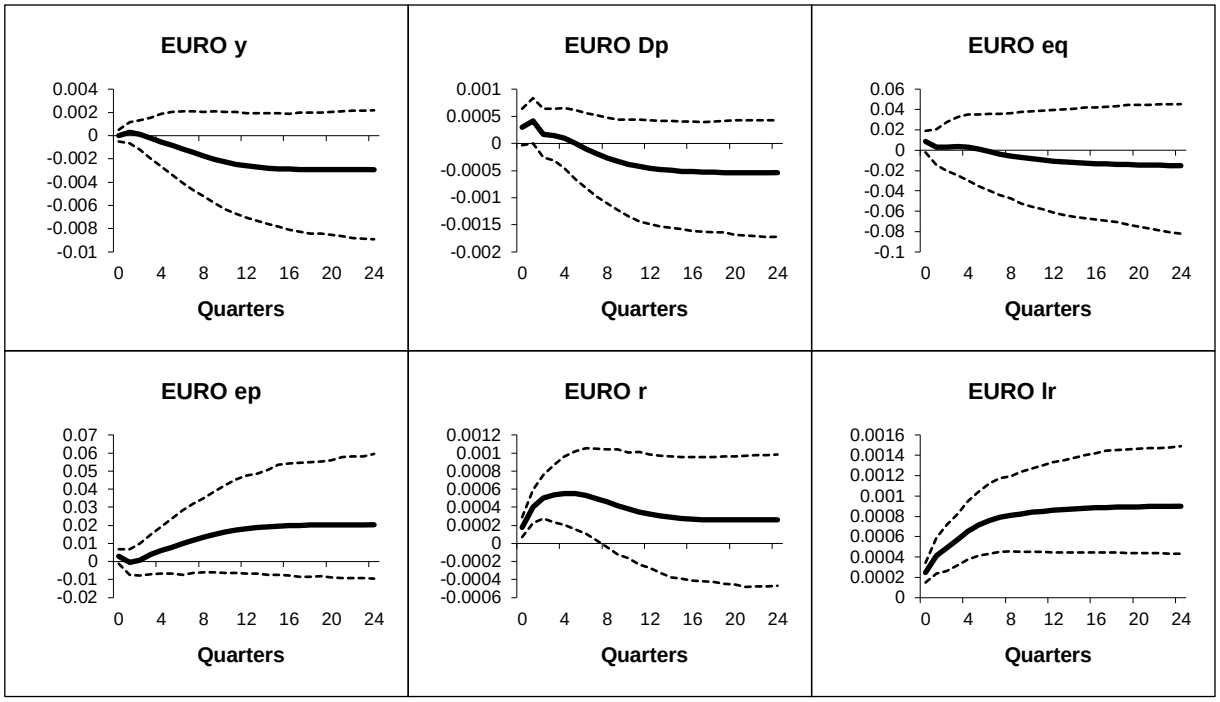

Figure B.1.51: Euro - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

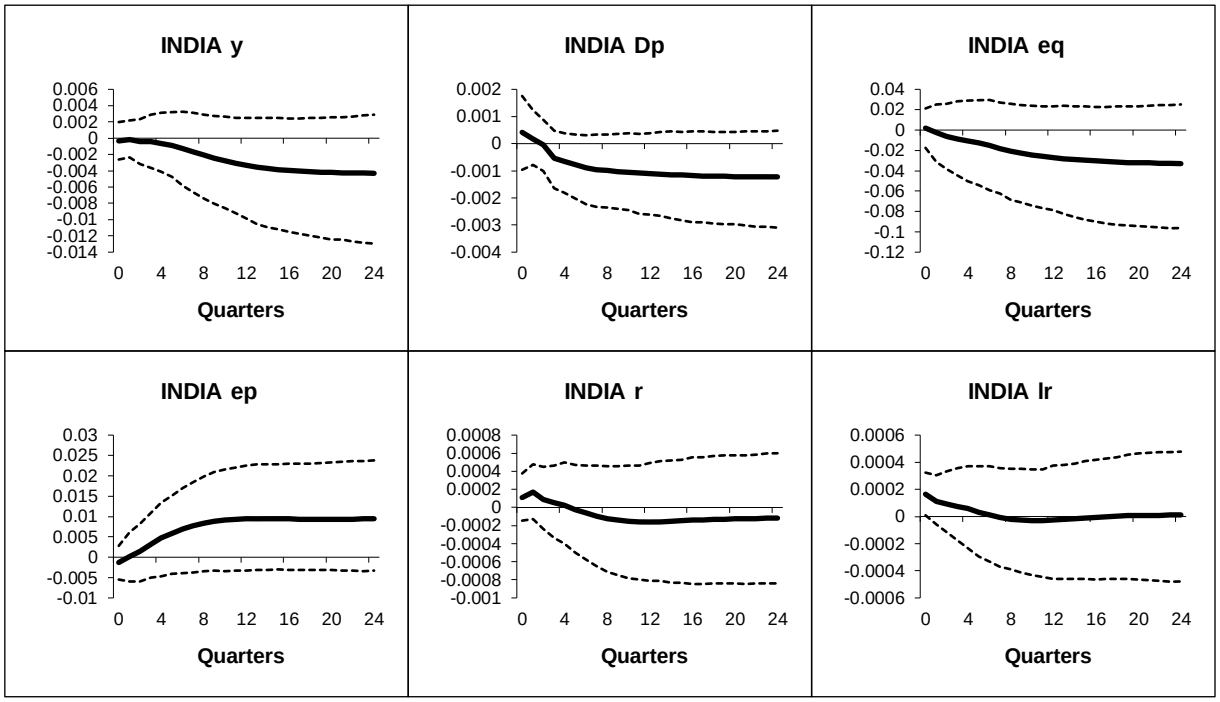

Figure B.1.52: India - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

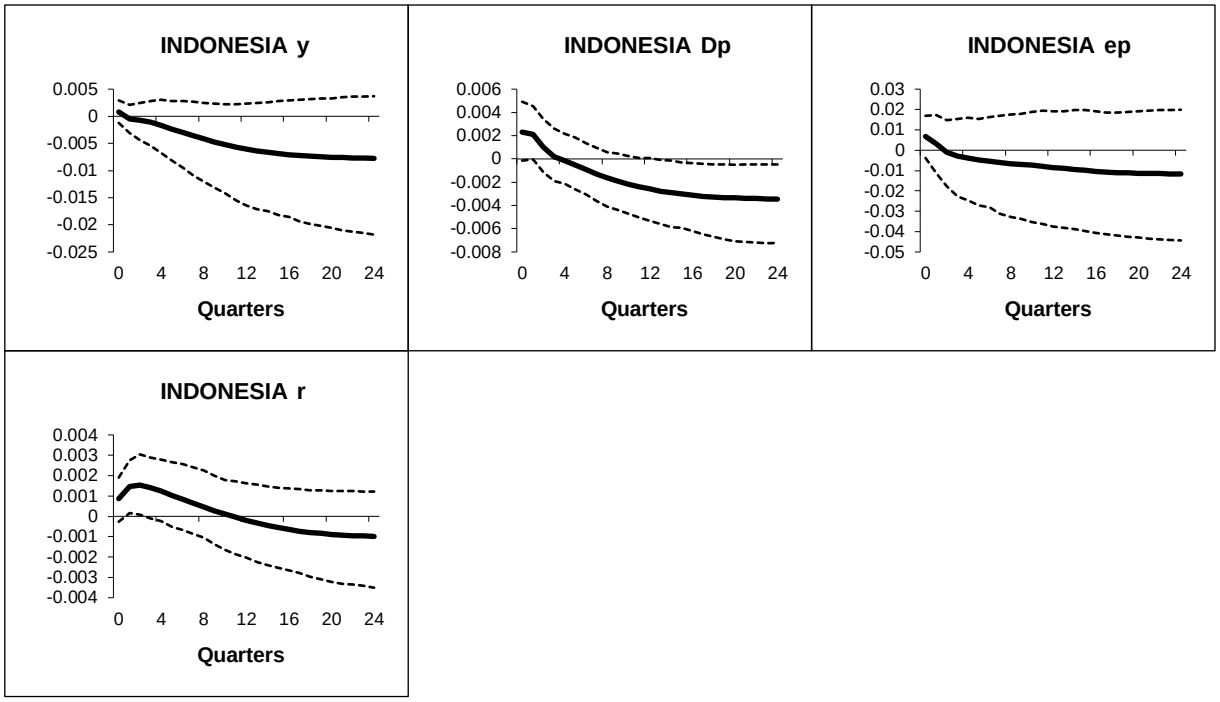

Figure B.1.53: Indonesia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

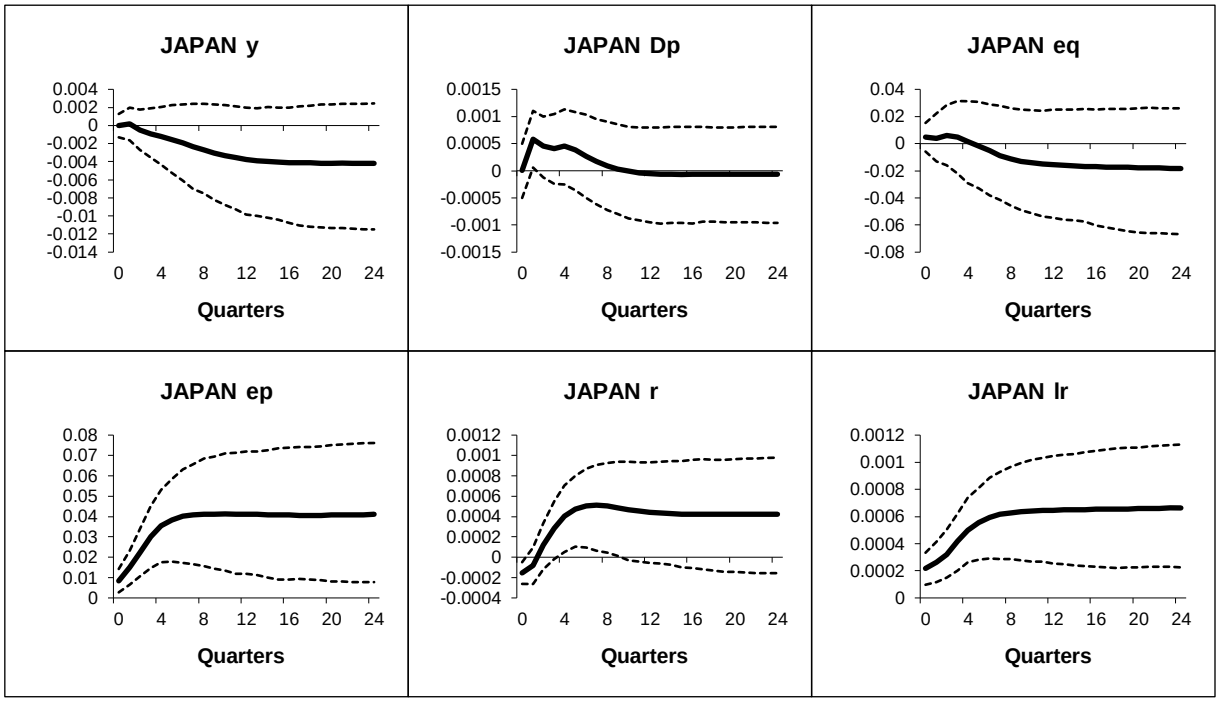

Figure B.1.54: Japan - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

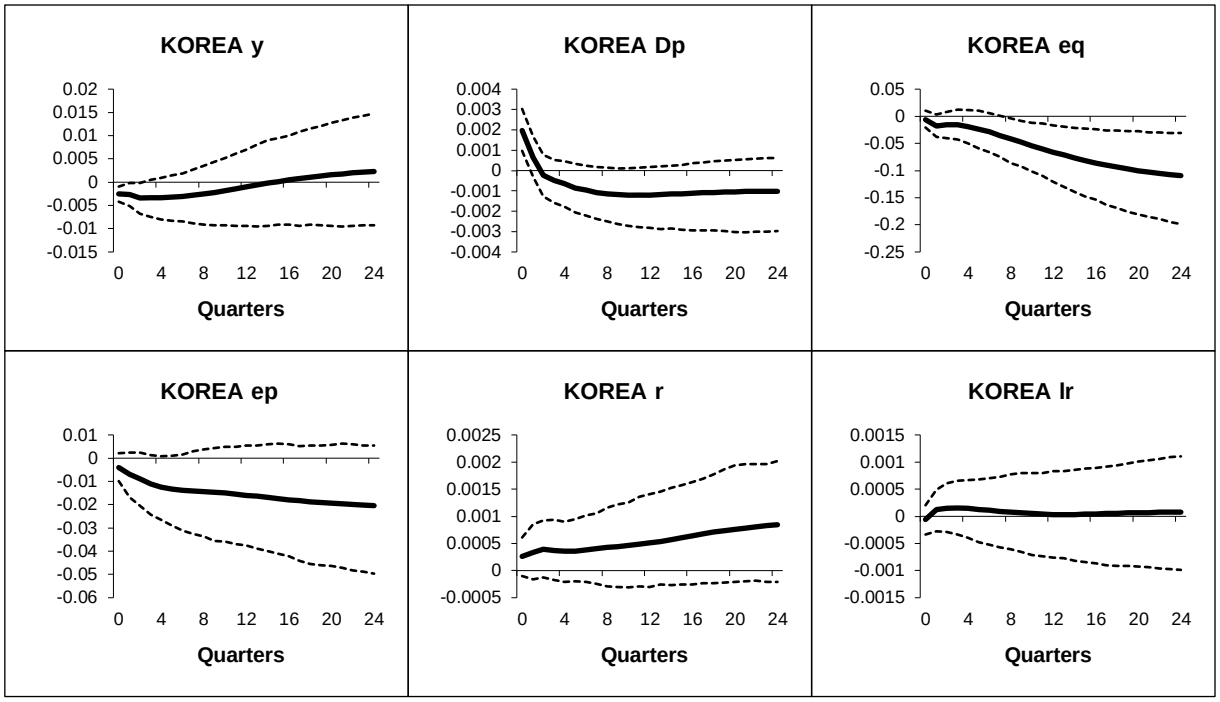

Figure B.1.55: Korea - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

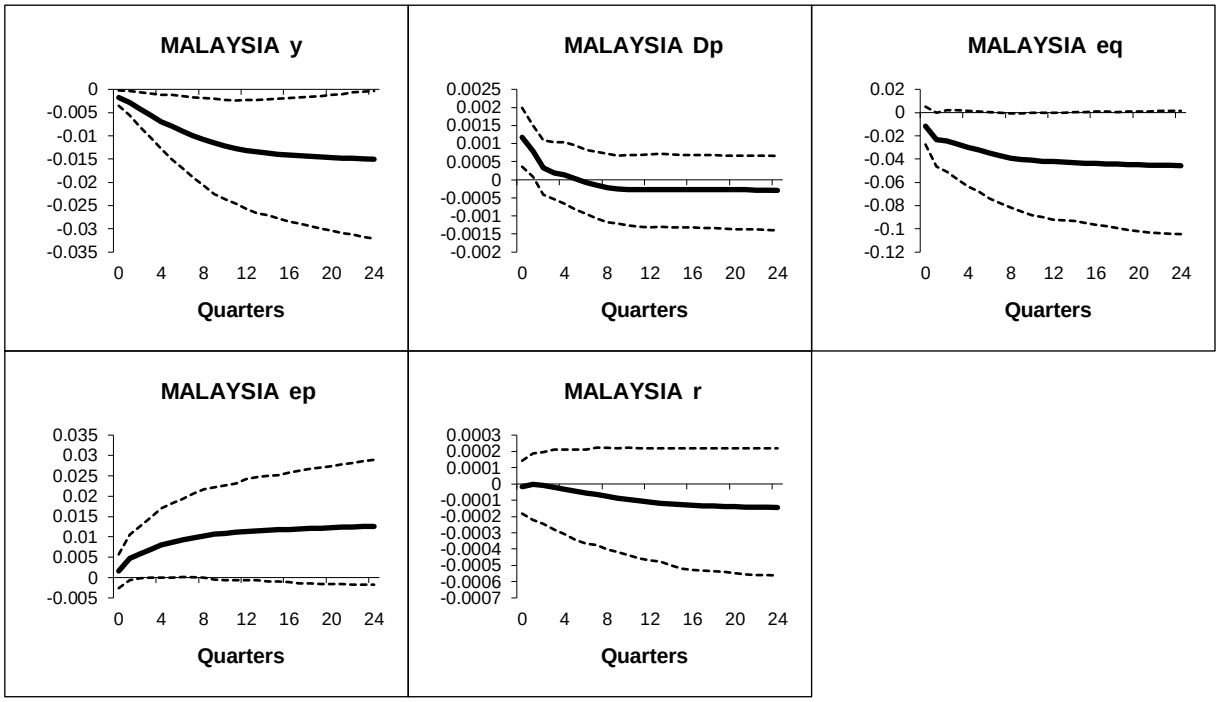

Figure B.1.56: Malaysia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

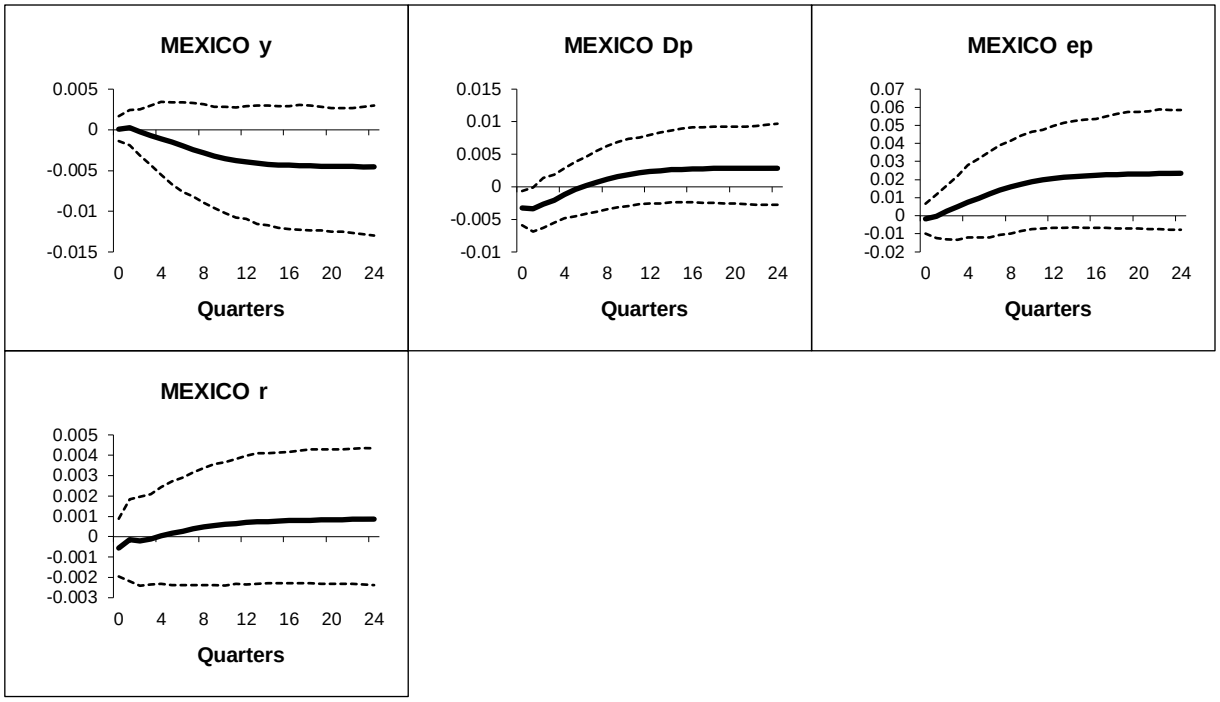

Figure B.1.57: Mexico - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

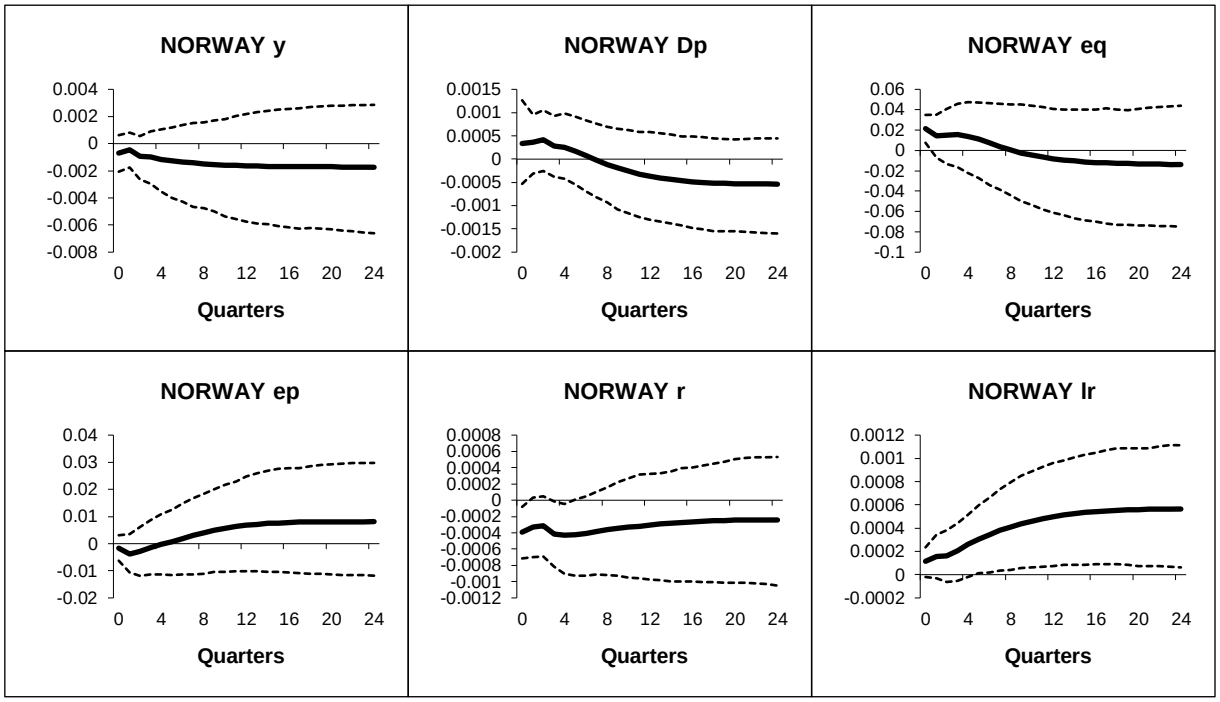

Figure B.1.58: Norway - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

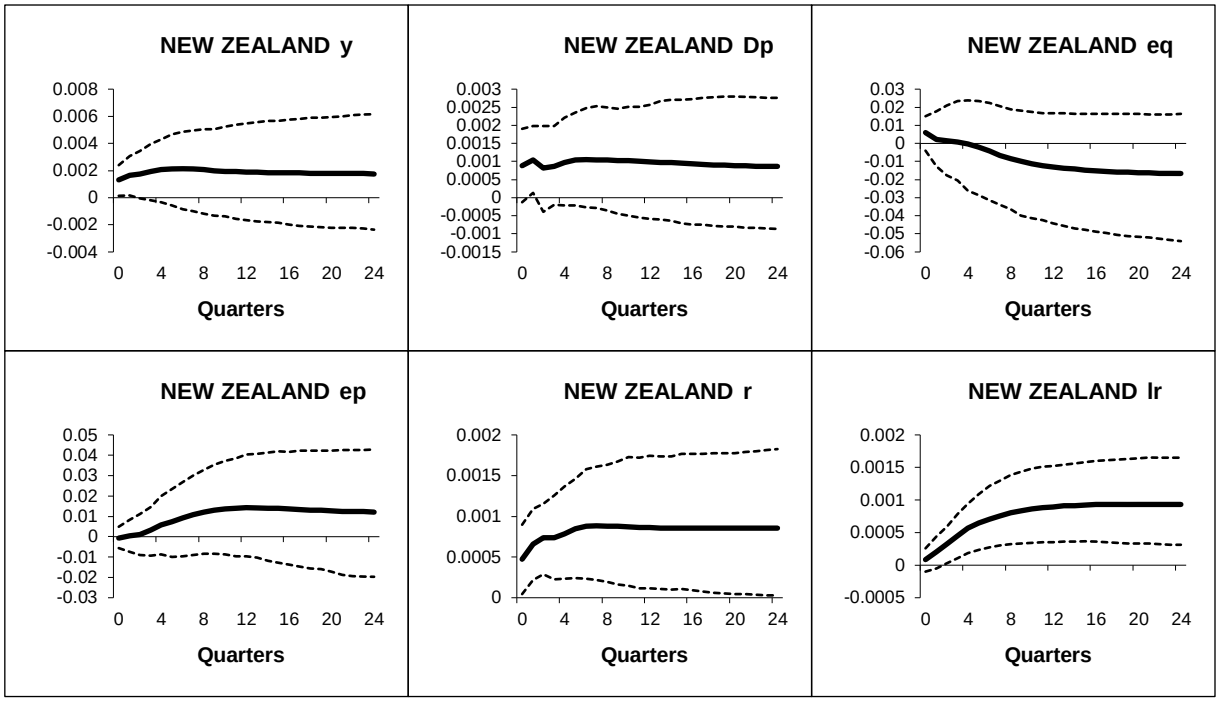

Figure B.1.59: New Zealand - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

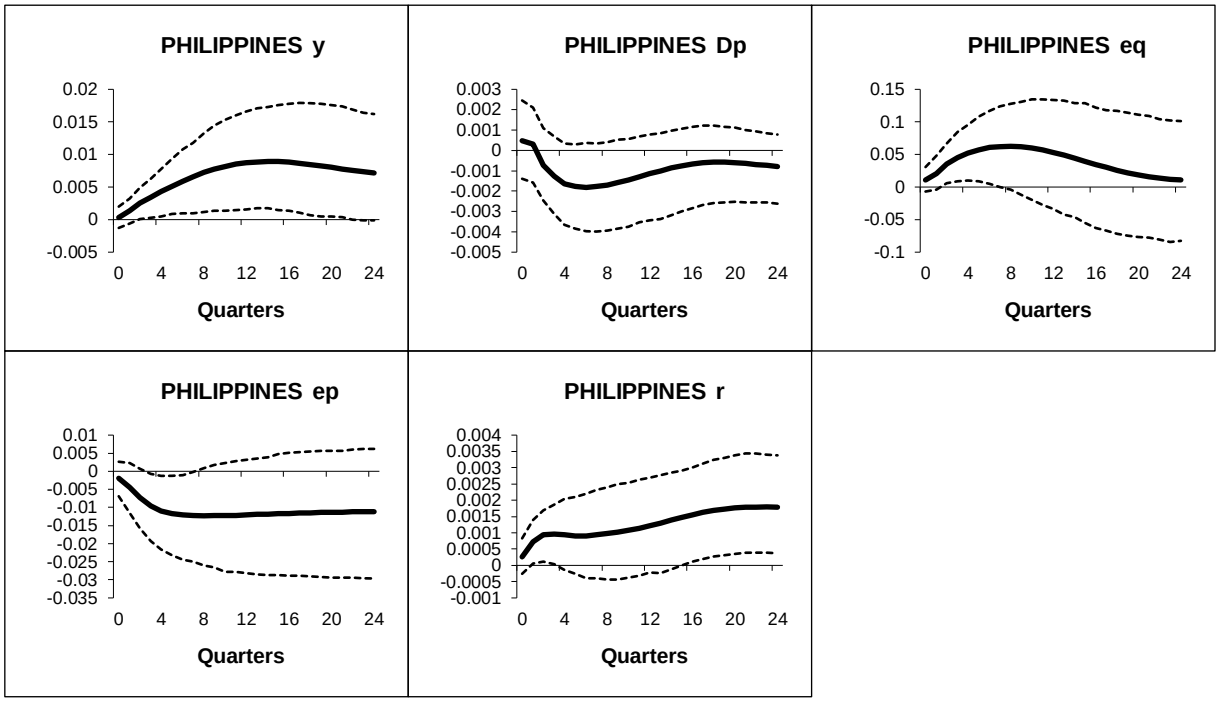

Figure B.1.60: Philippines - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

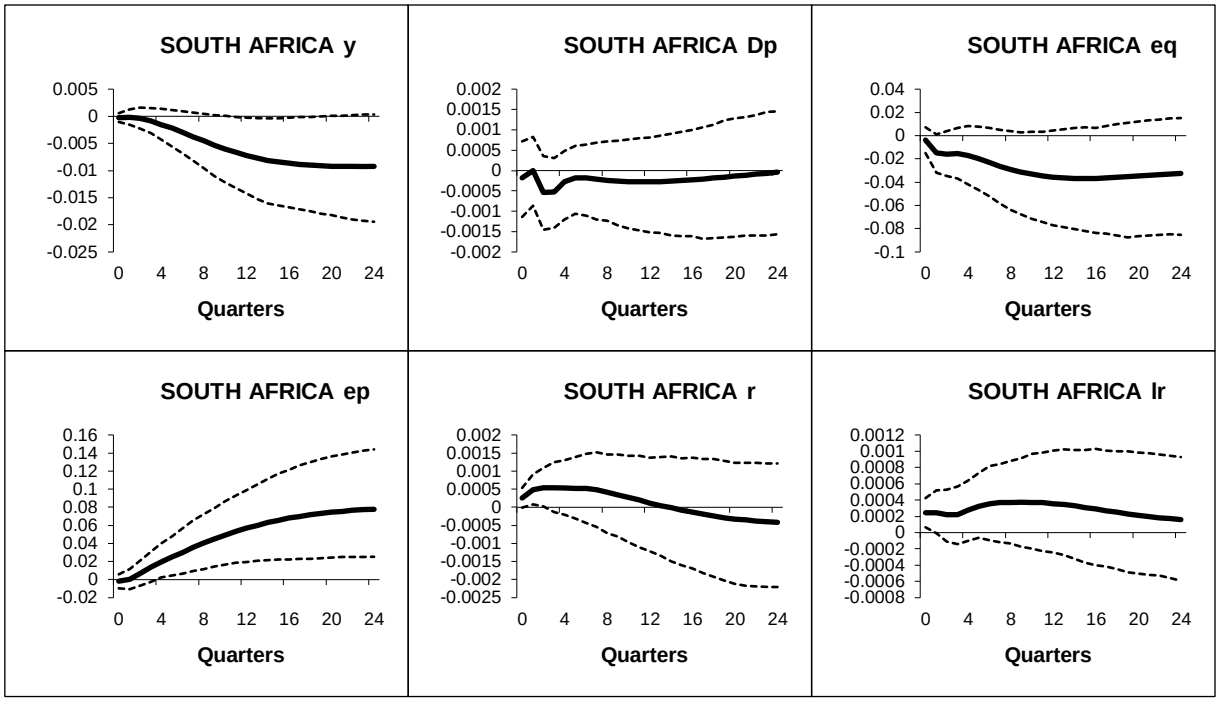

Figure B.1.61: South Africa - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

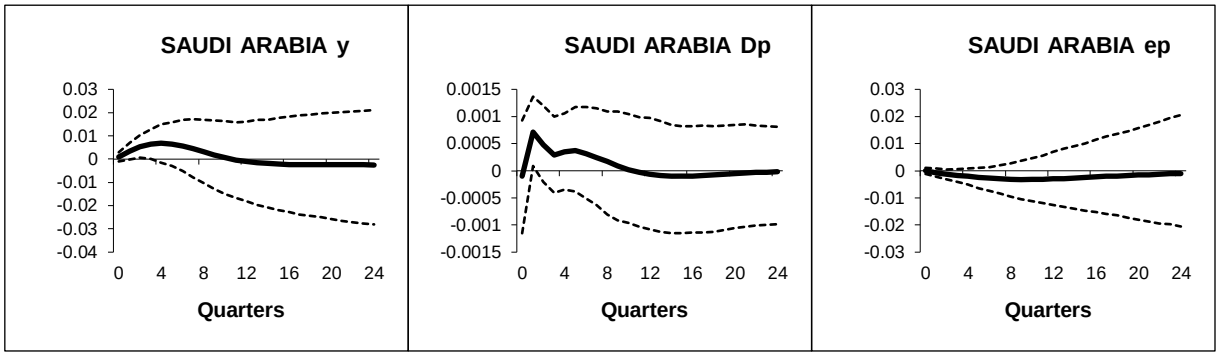

Figure B.1.62: Saudi Arabia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

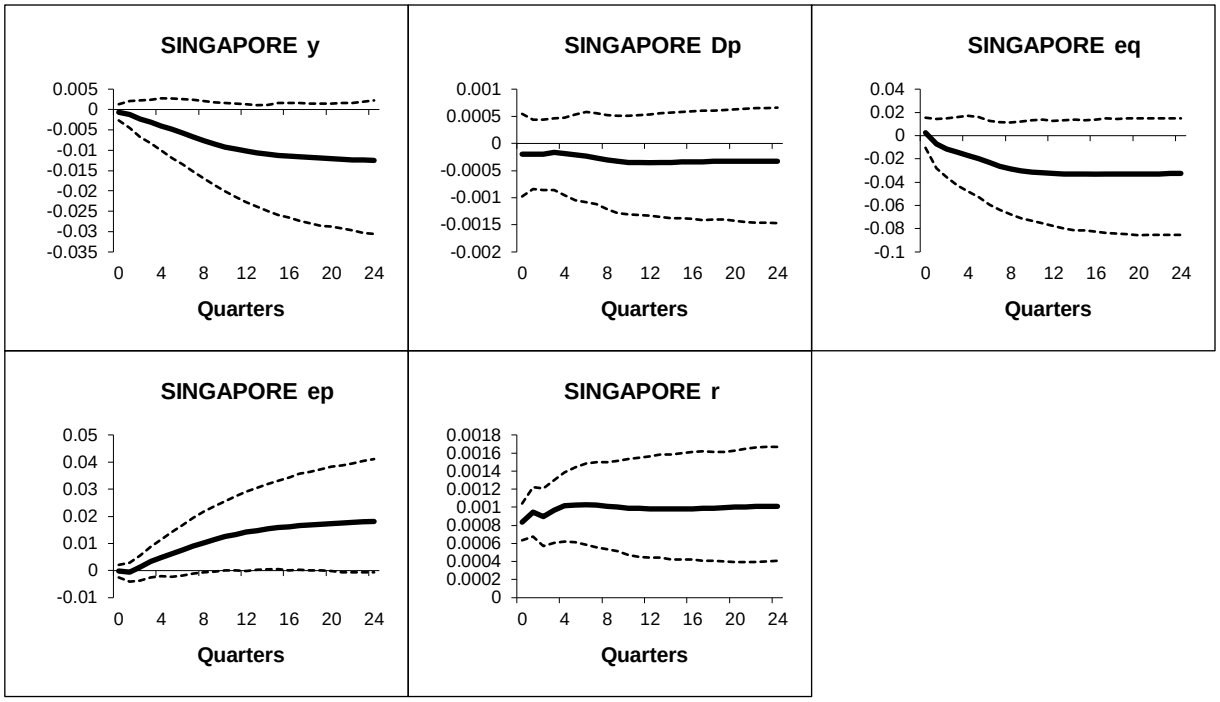

Figure B.1.63: Singapore - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

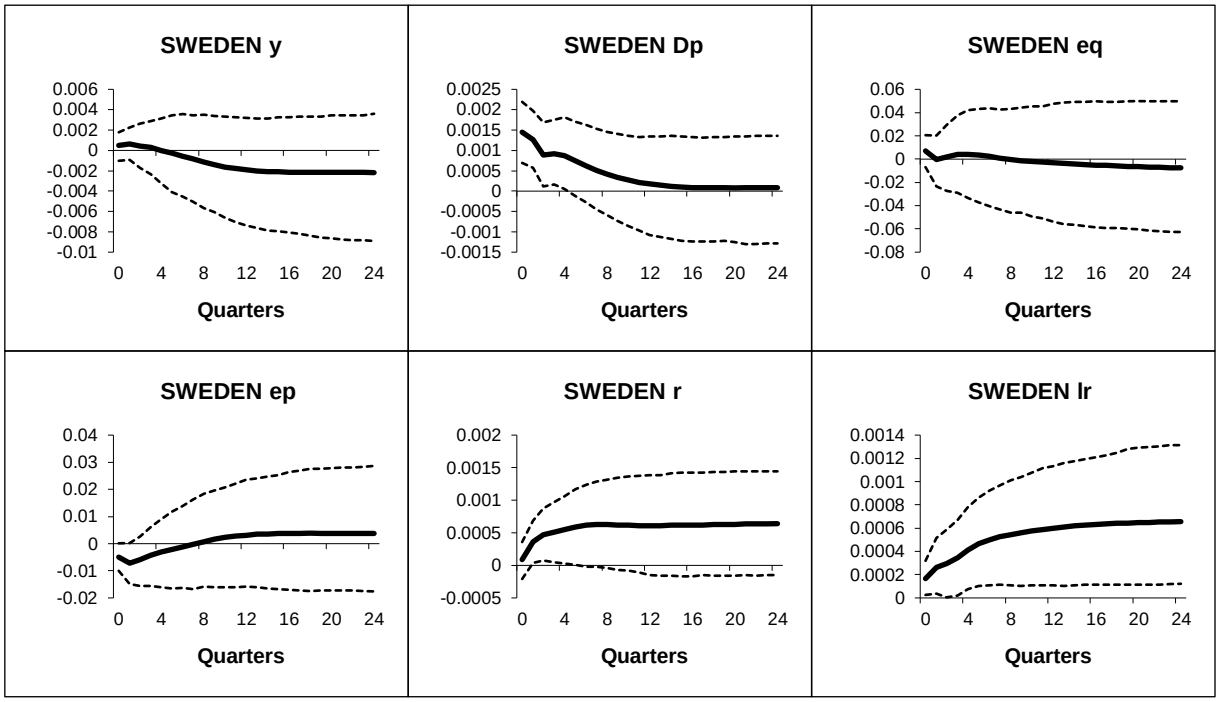

Figure B.1.64: Sweden - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

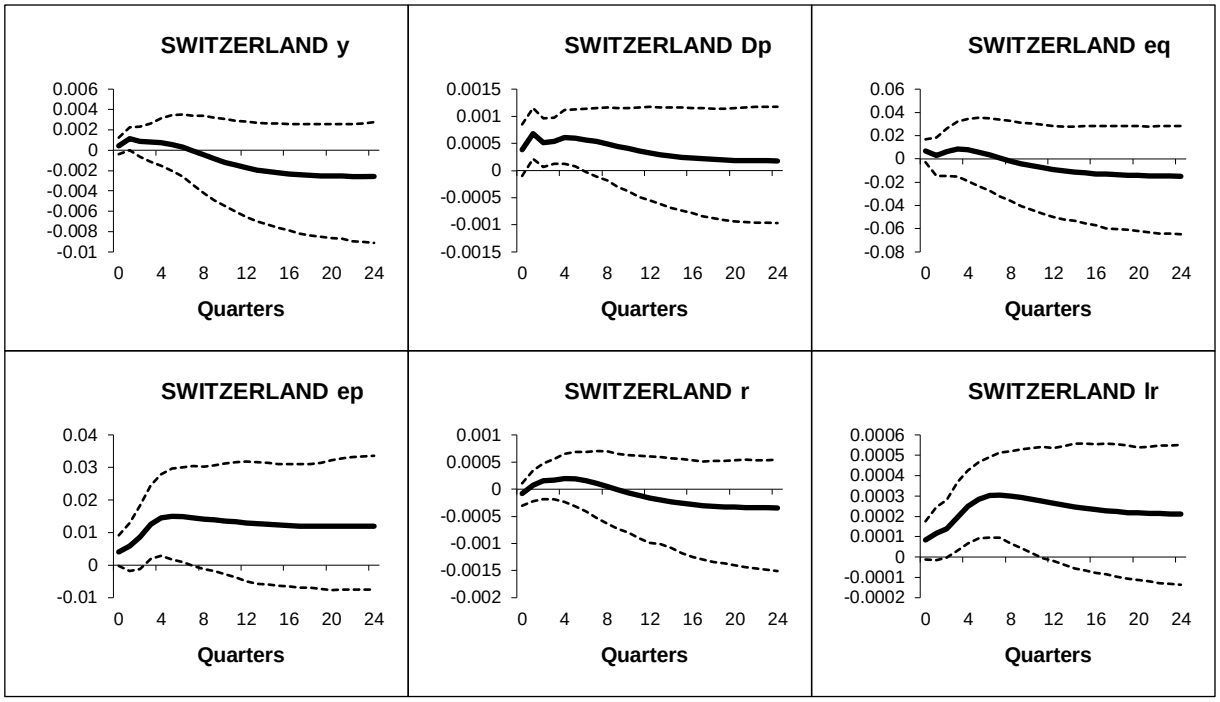

Figure B.1.65: Switzerland - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

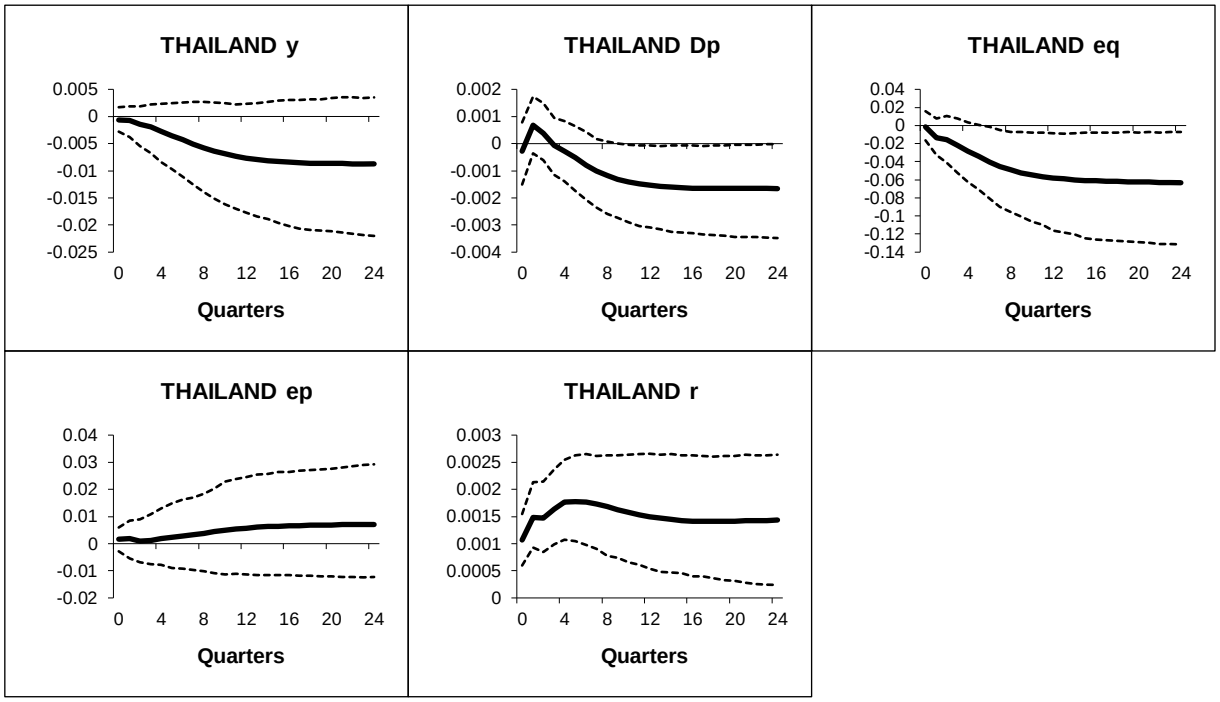

Figure B.1.66: Thailand - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

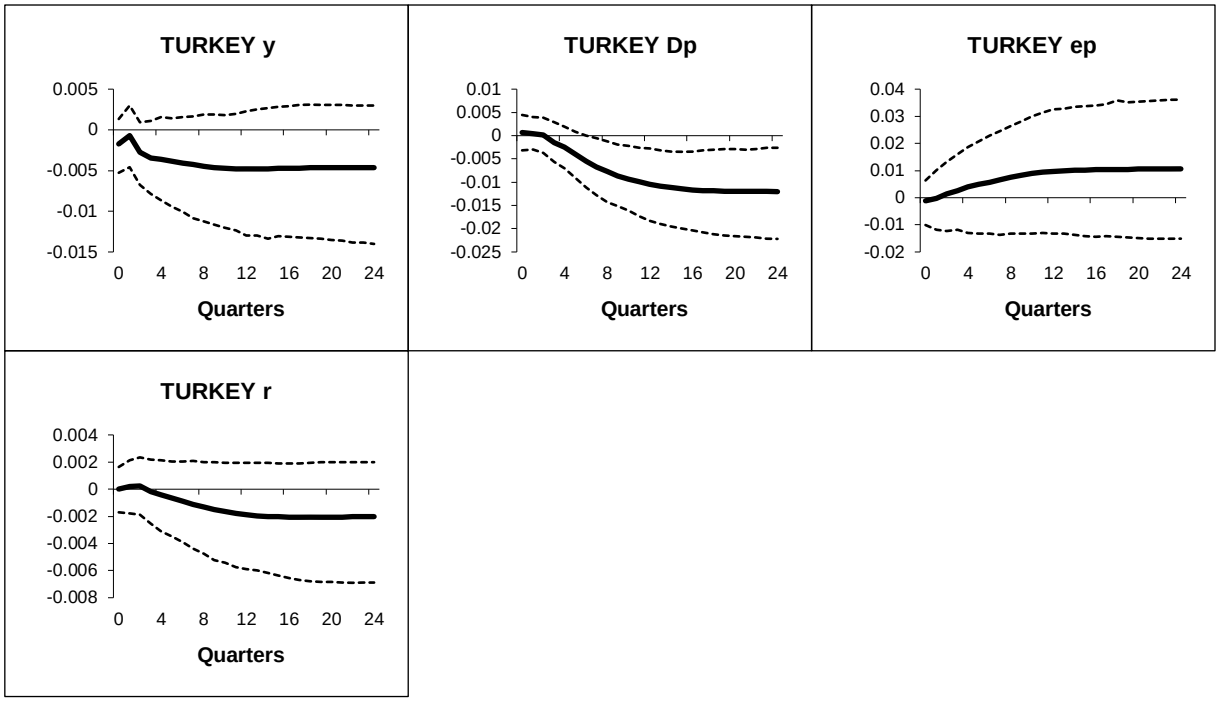

Figure B.1.67: Turkey - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

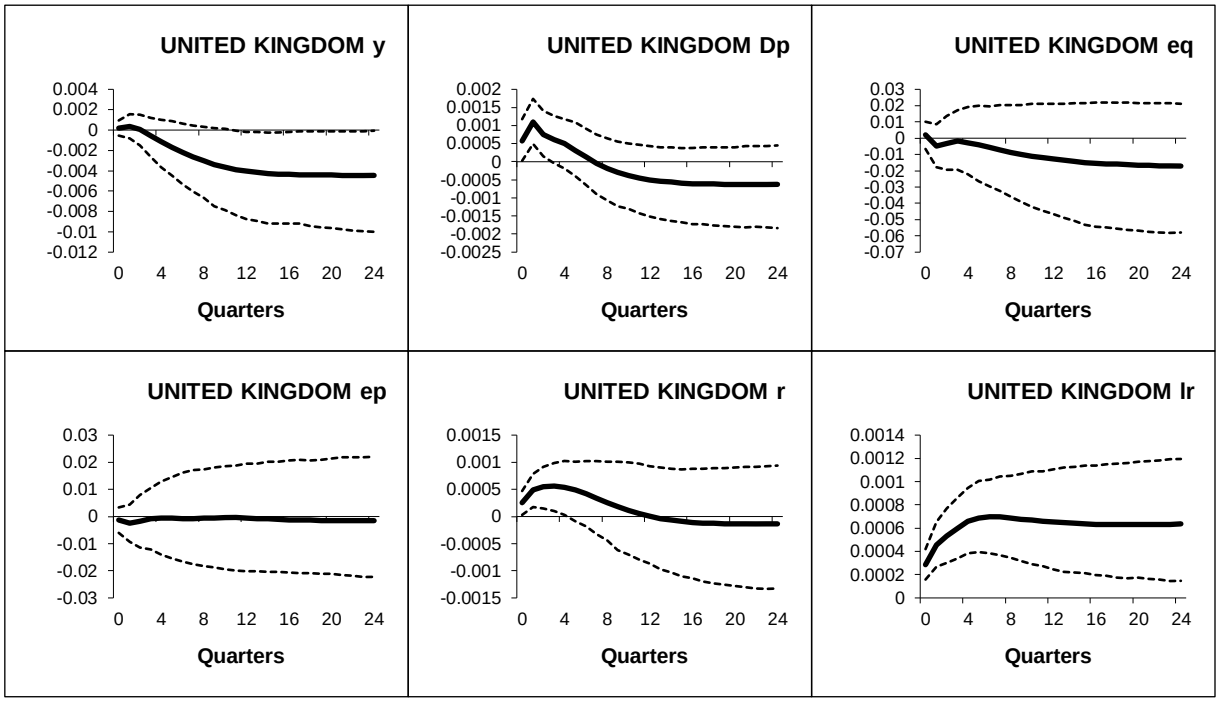

Figure B.1.68: UK - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

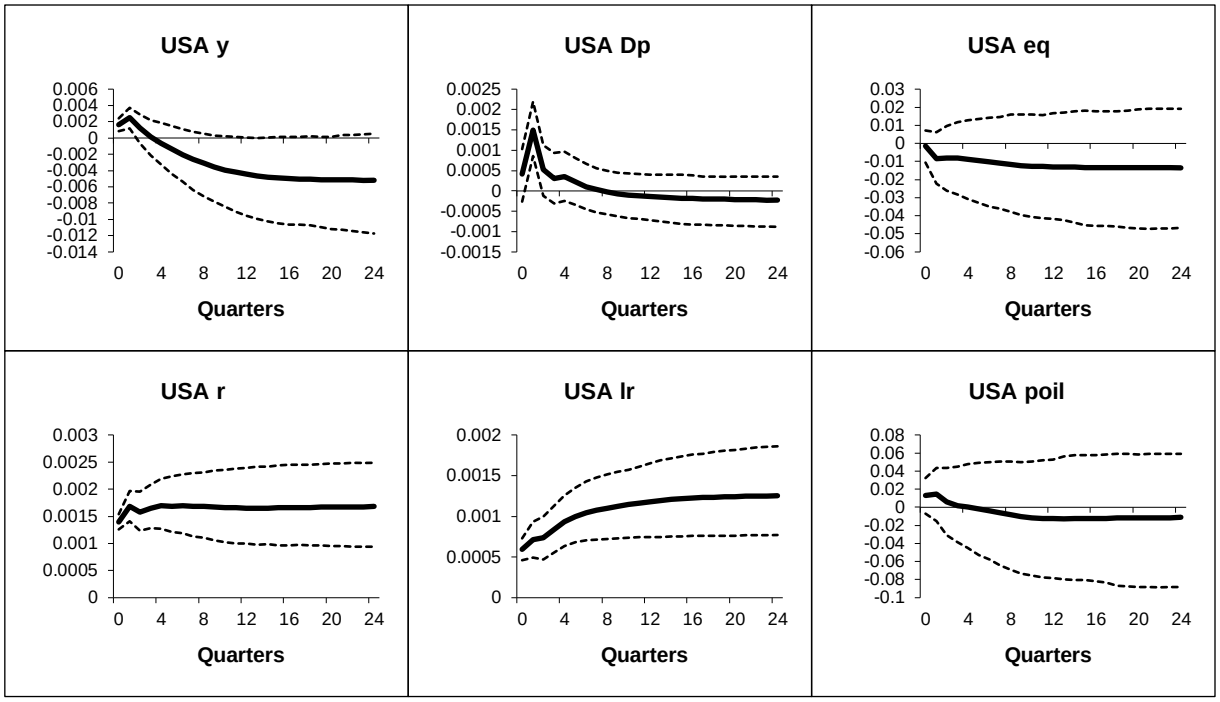

Figure B.1.69: USA - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

APPENDIX B.1

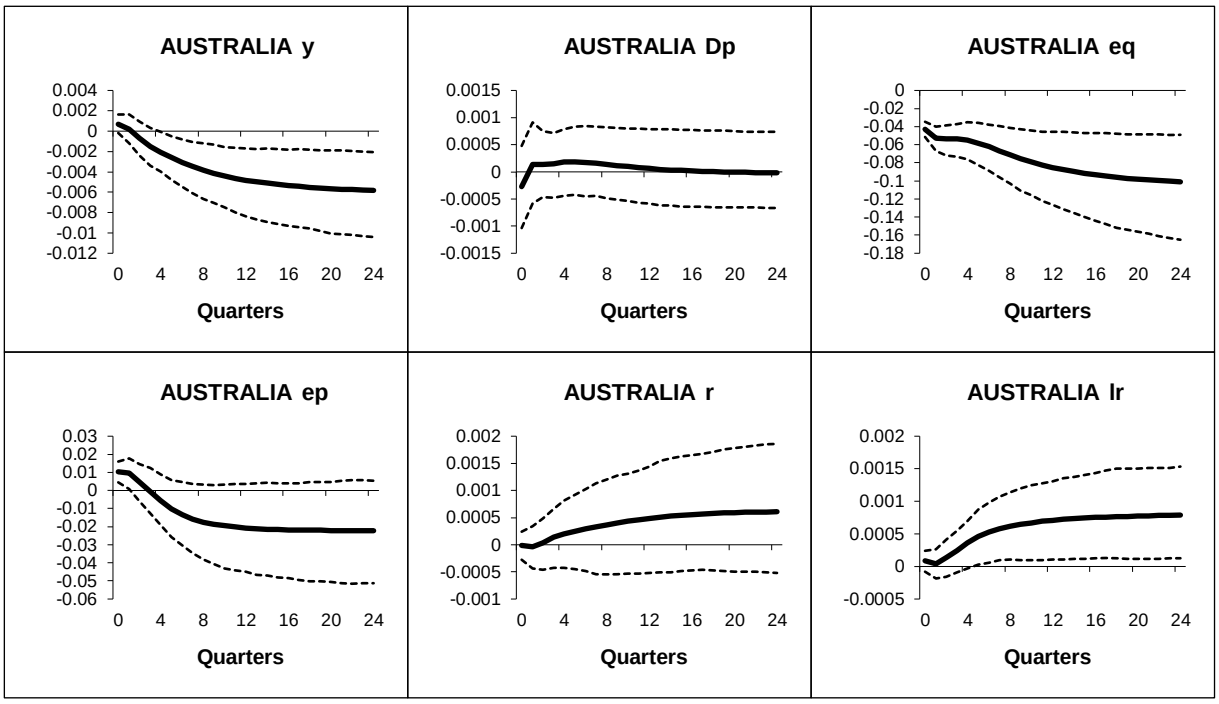

Figure B.1.1: Australia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

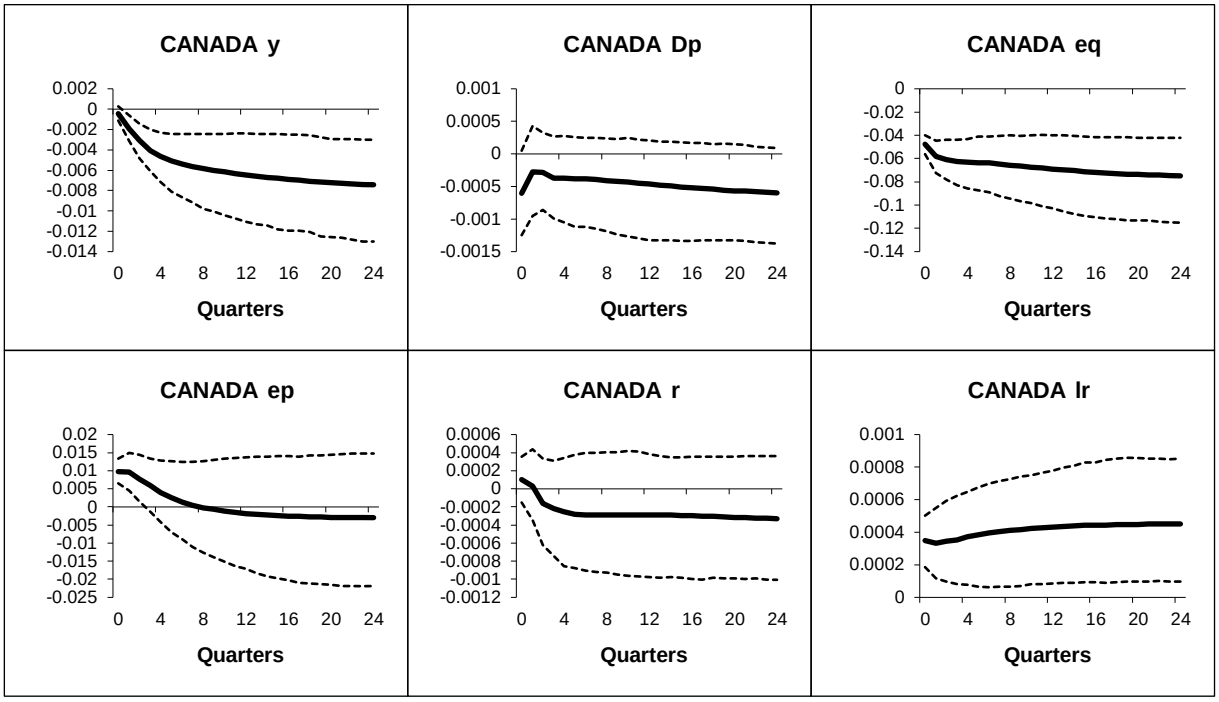

Figure B.1.2: Canada - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

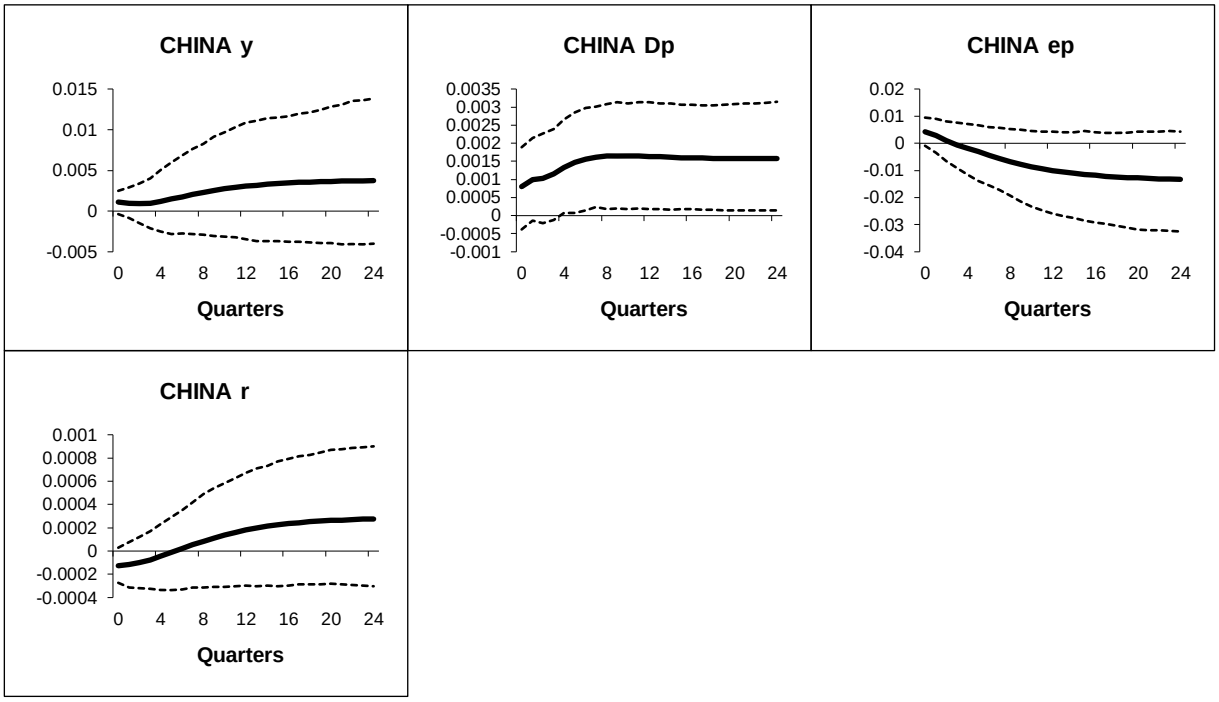

Figure B.1.3: China - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

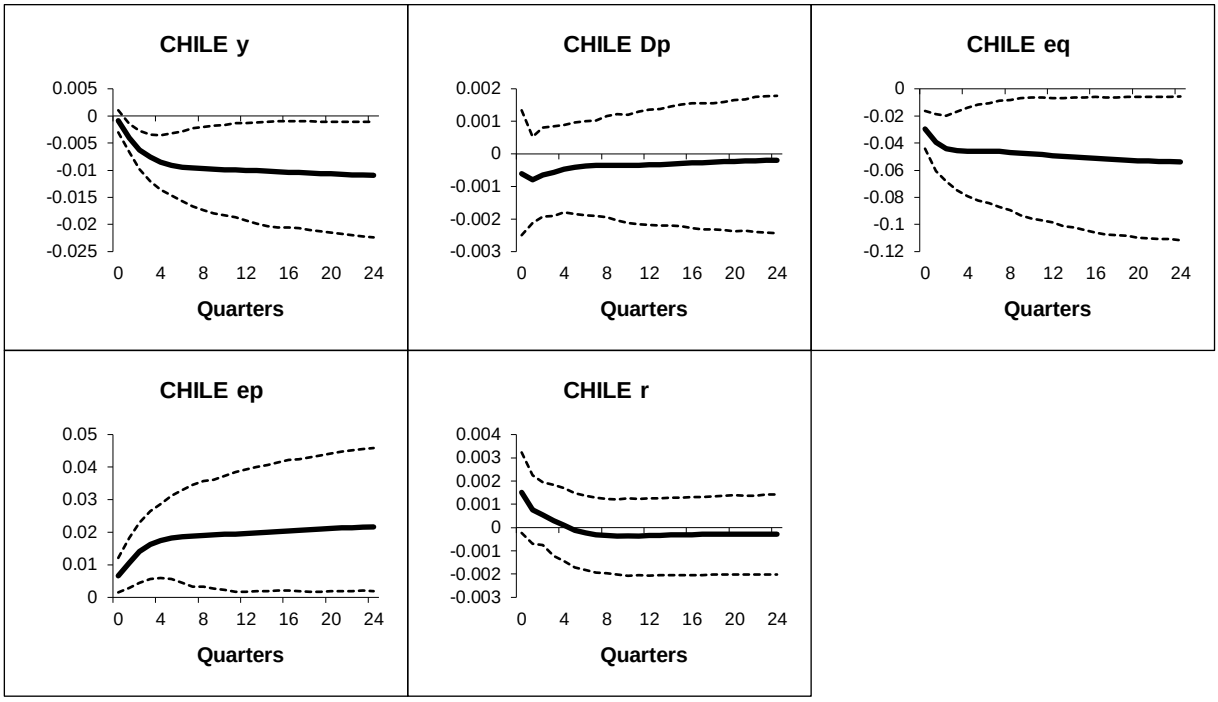

Figure B.1.4: Chile - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

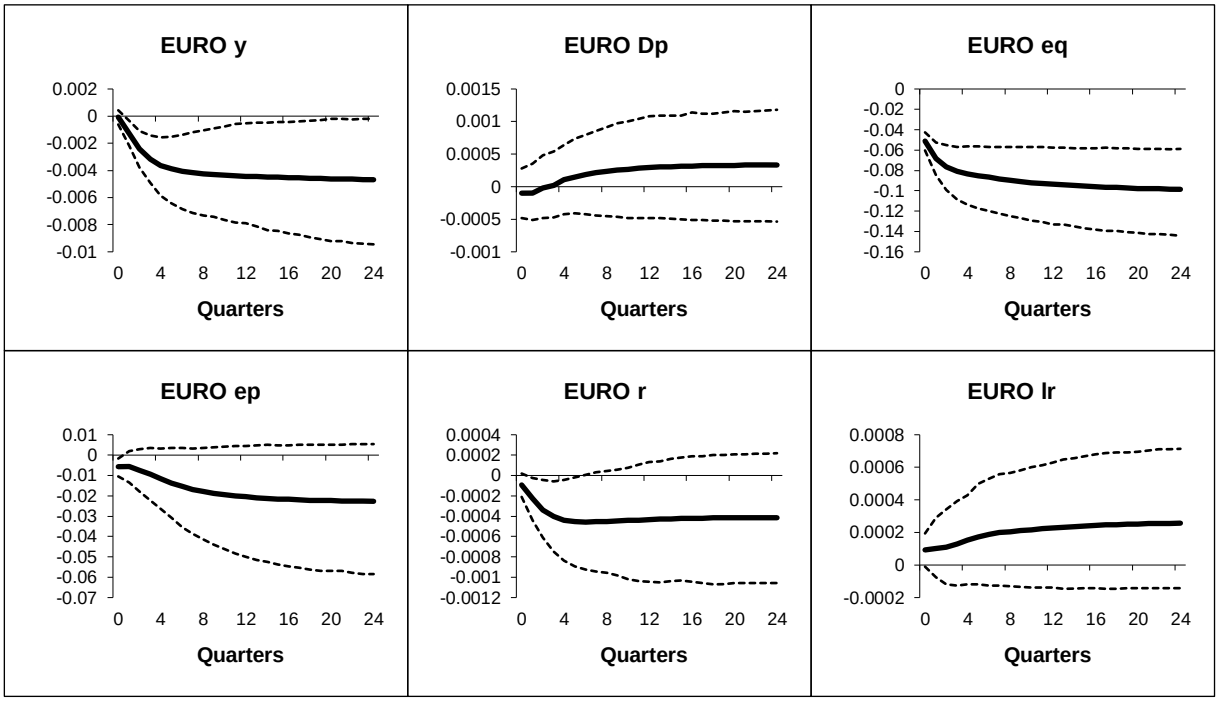

Figure B.1.5: Euro - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

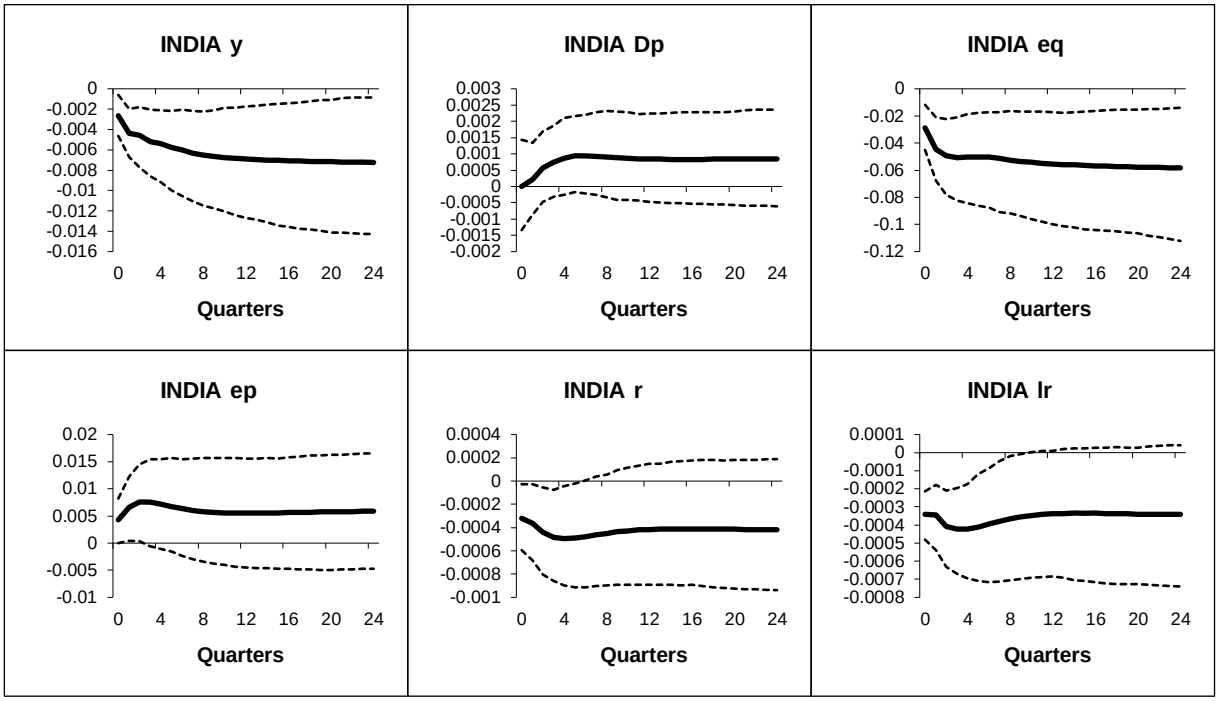

Figure B.1.6: India - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

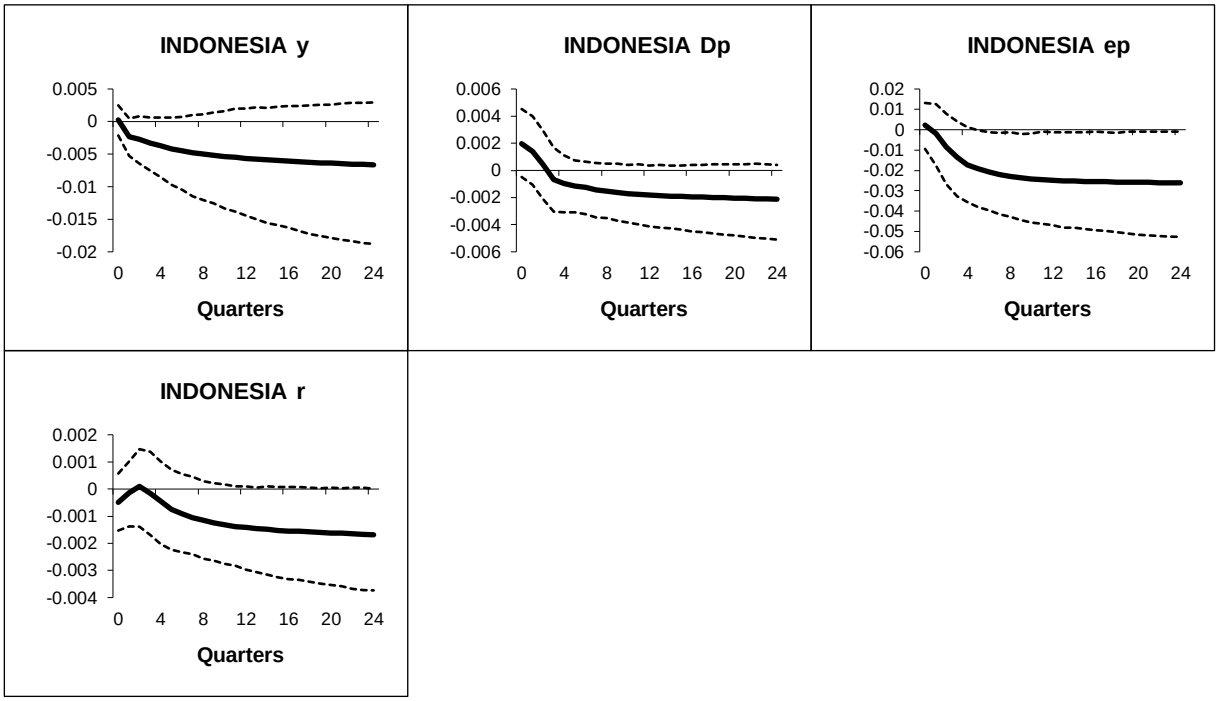

Figure B.1.7: Indonesia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

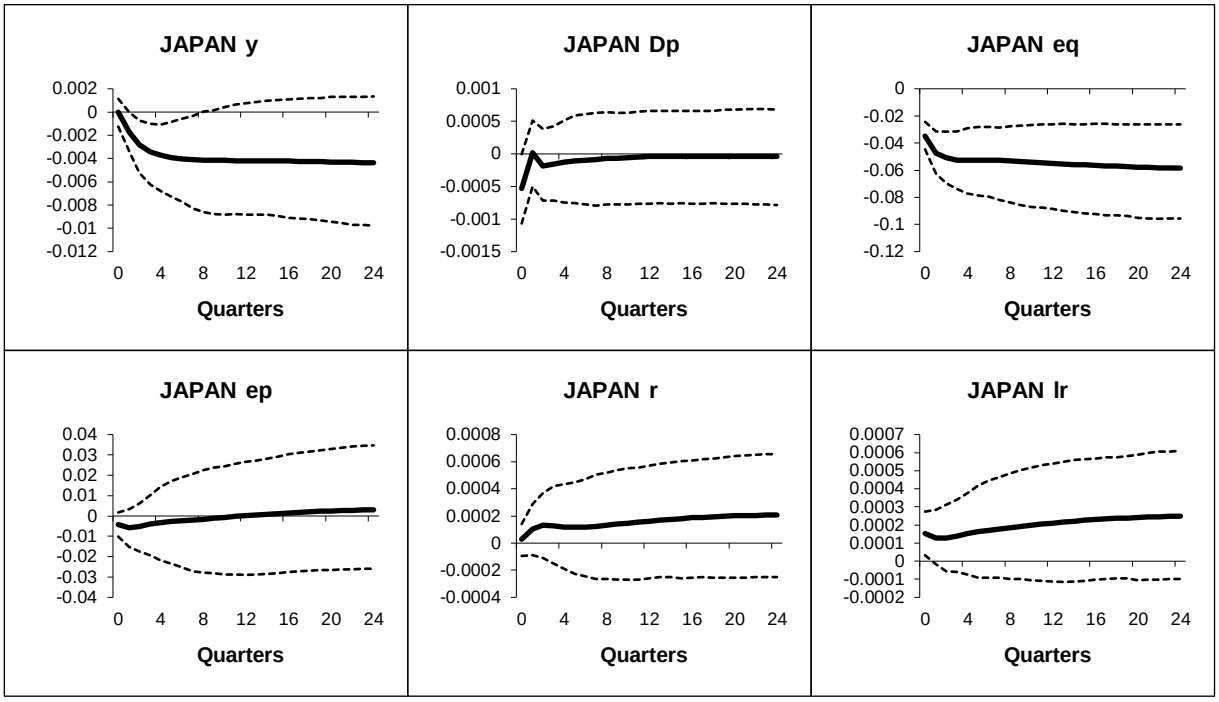

Figure B.1.8: Japan - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

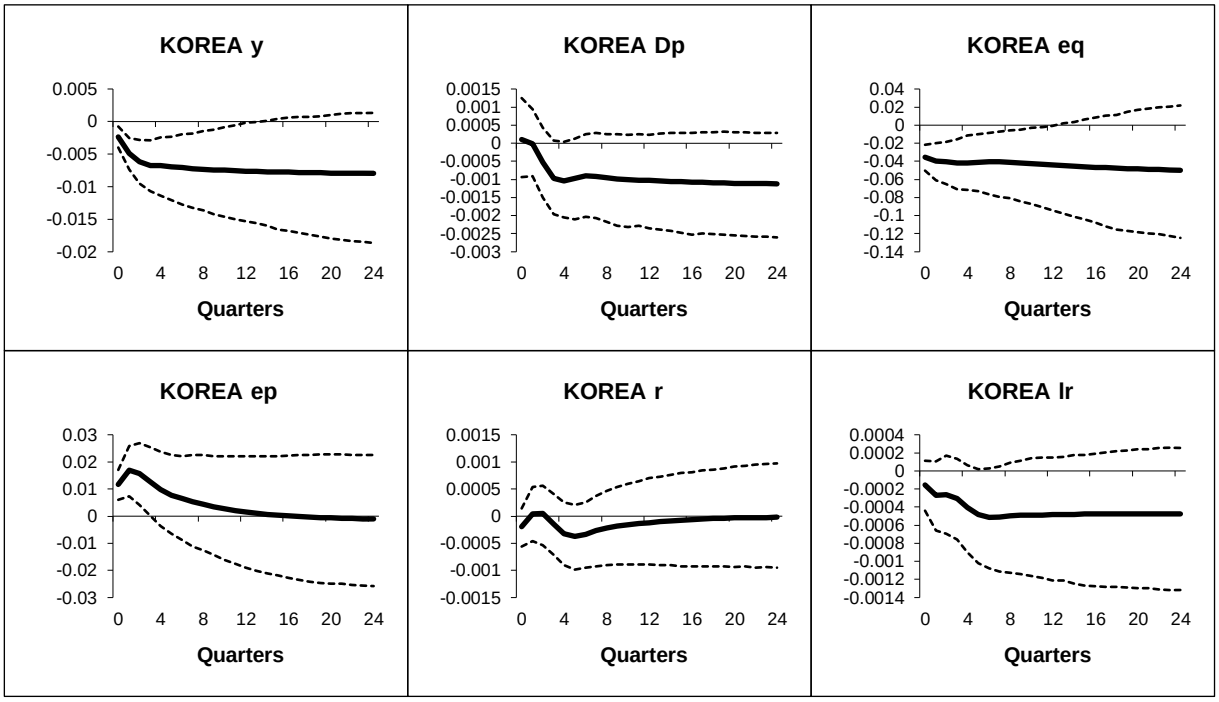

Figure B.1.9: Korea - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

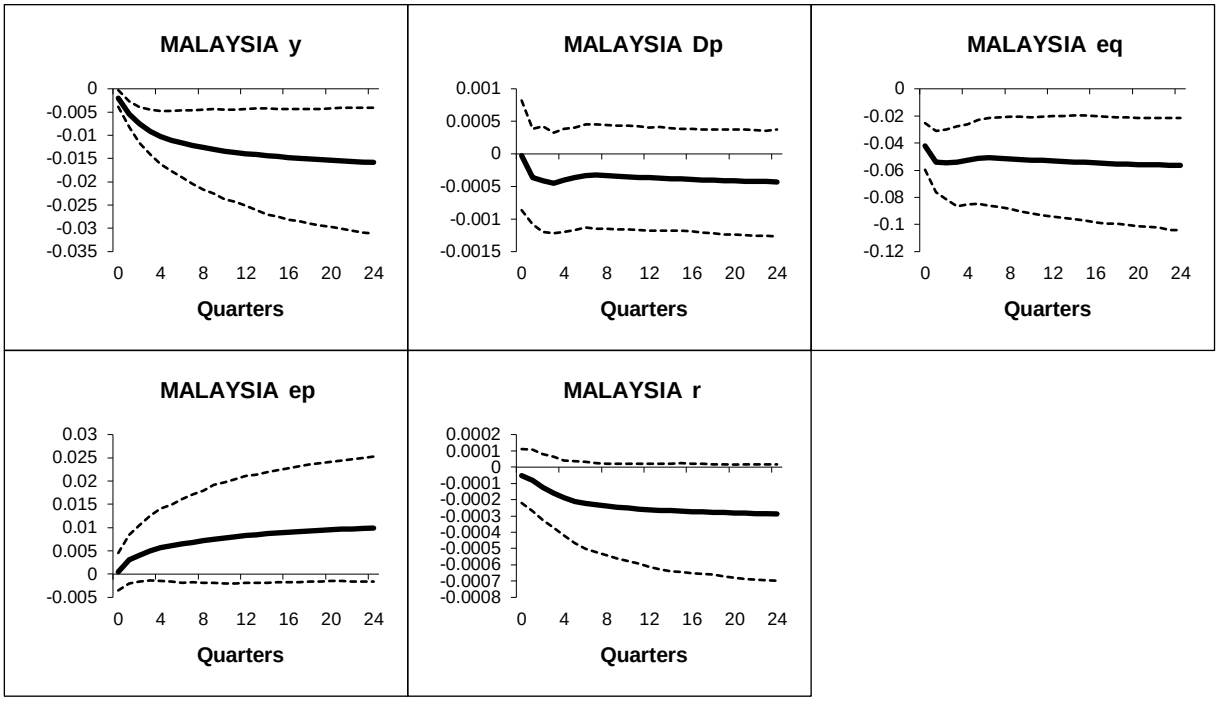

Figure B.1.10: Malaysia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

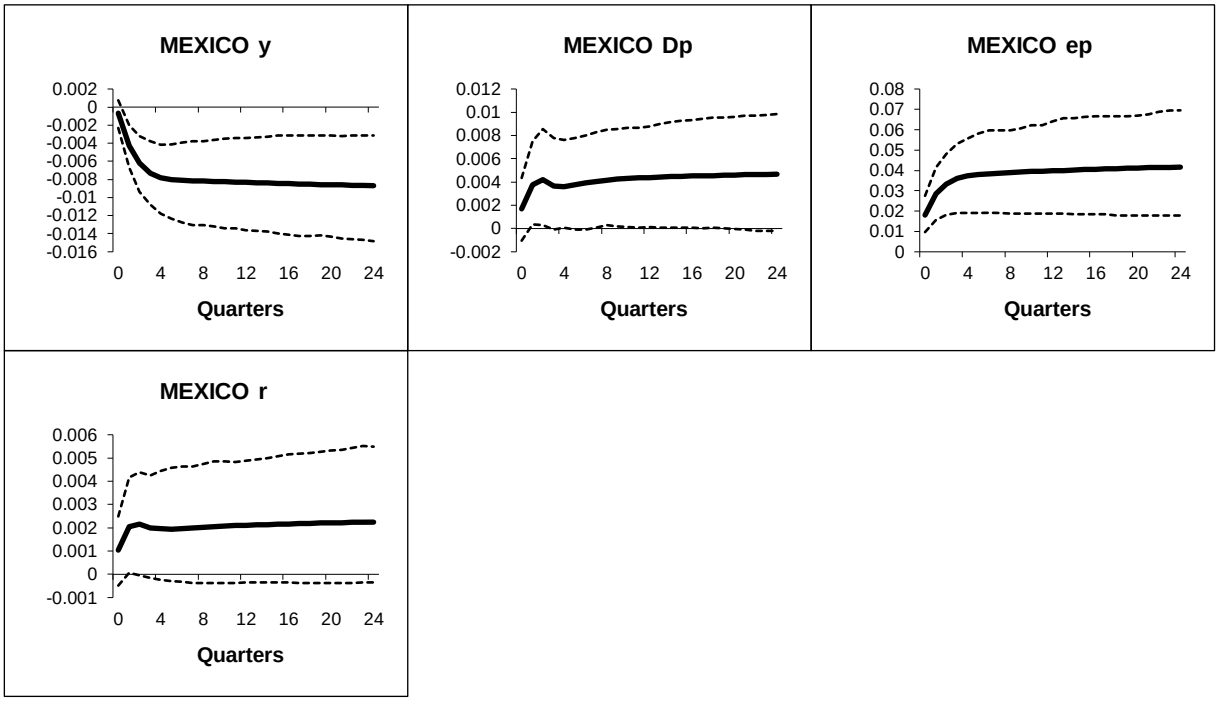

Figure B.1.11: Mexico - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

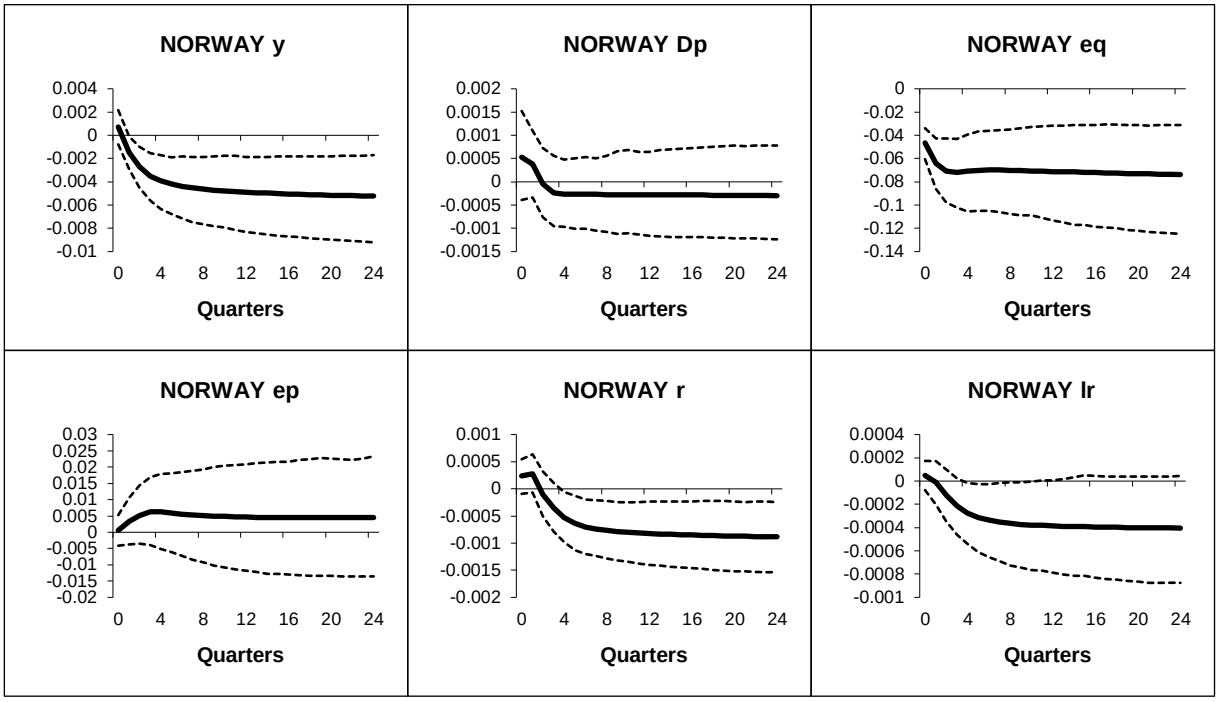

Figure B.1.12: Norway - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

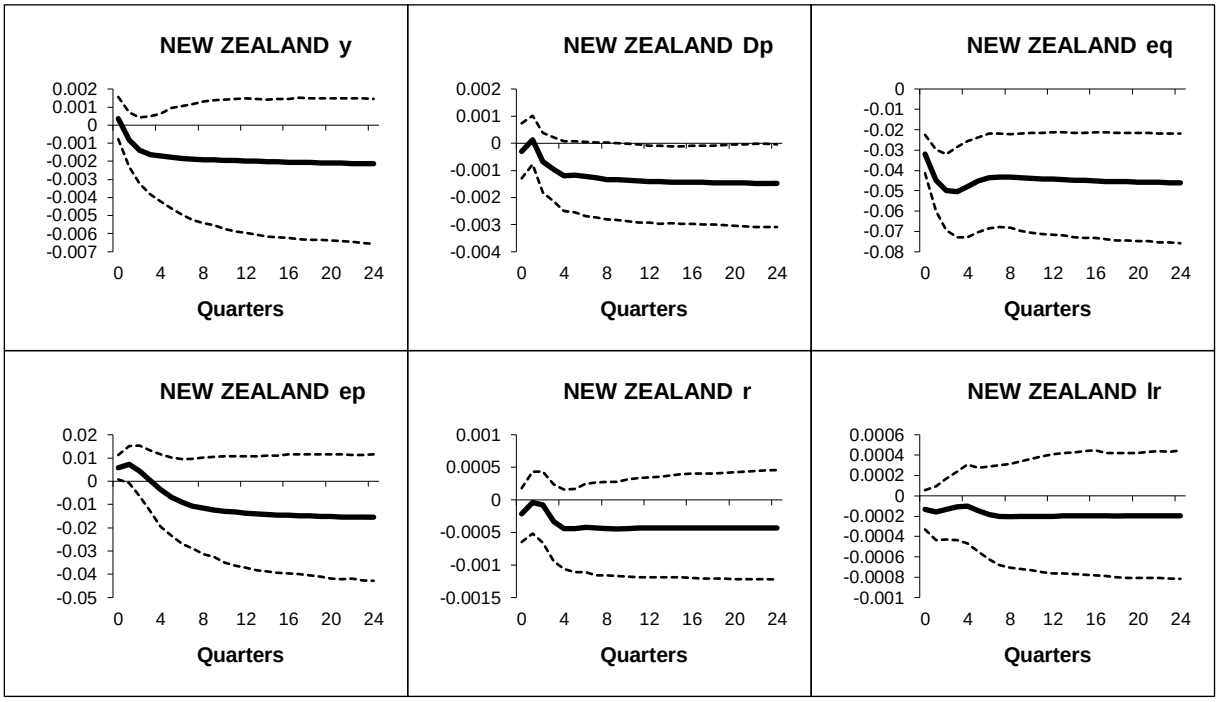

Figure B.1.13: New Zealand - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

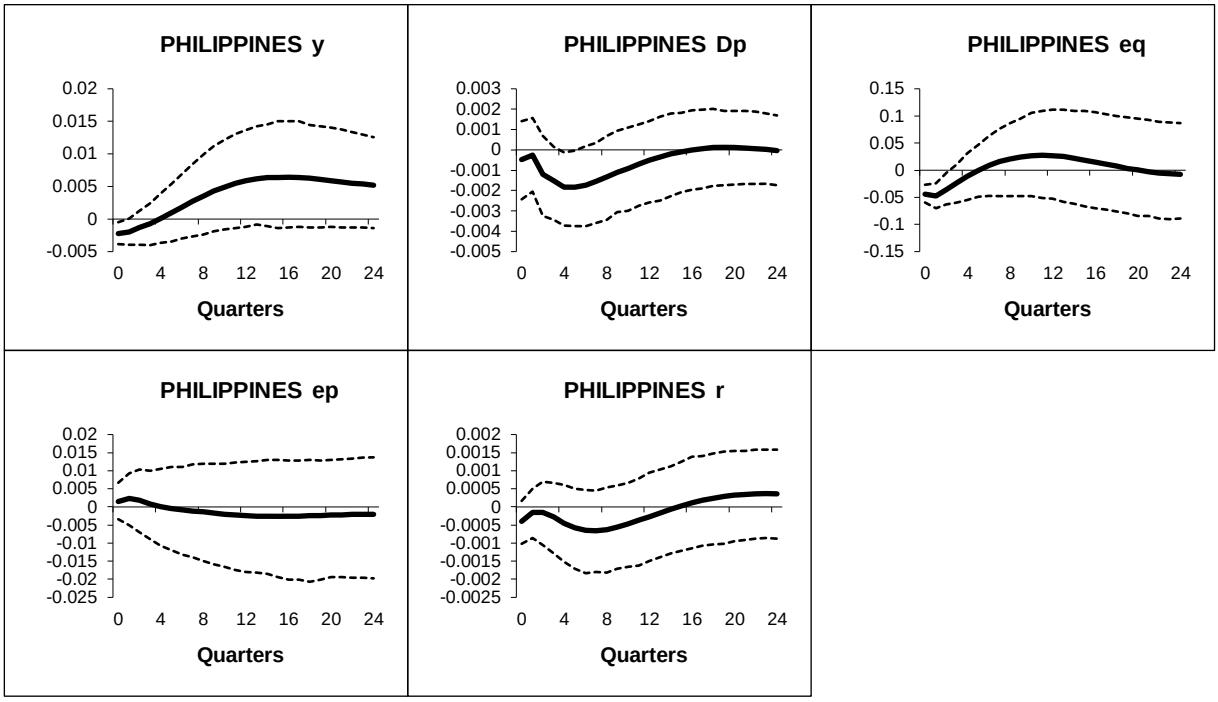

Figure B.1.14: Philippines - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

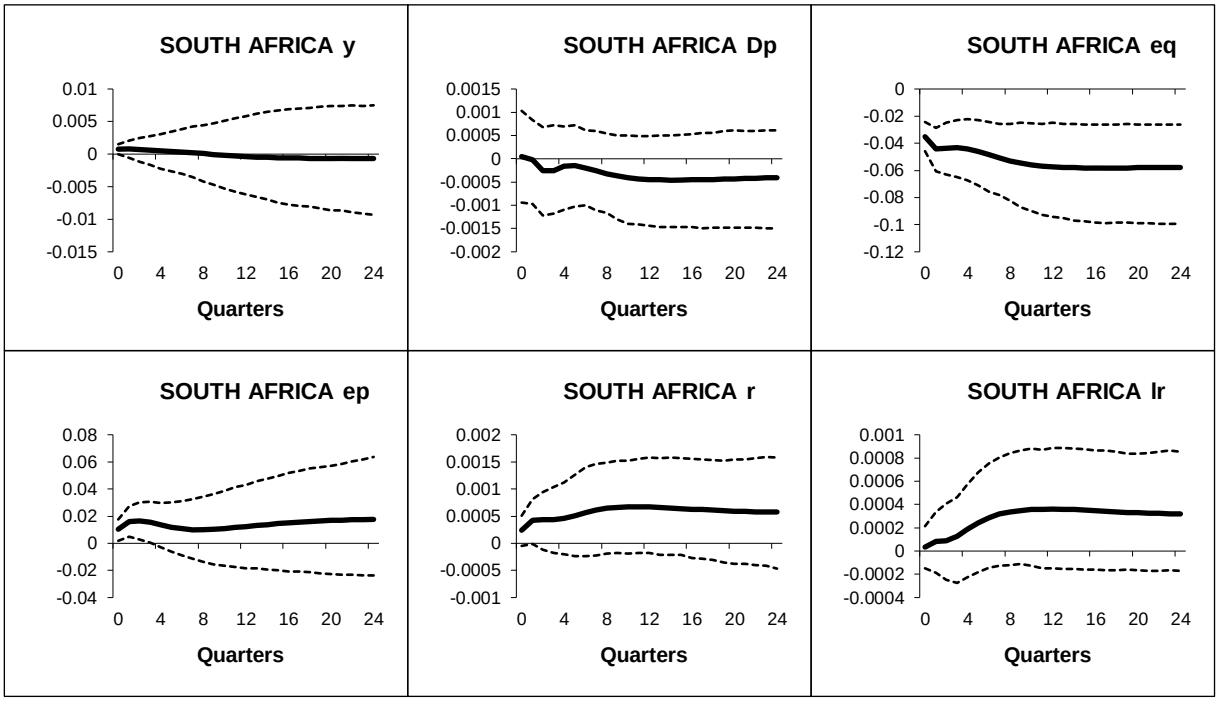

Figure B.1.15: South Africa - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

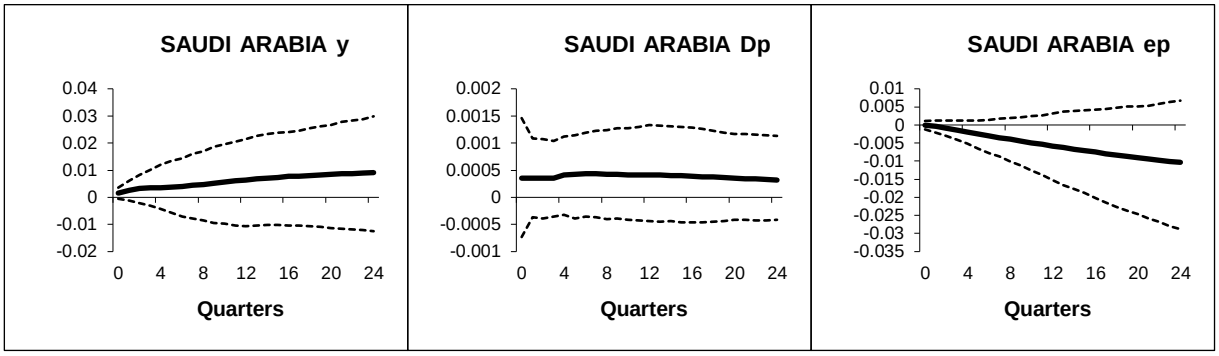

Figure B.1.16: Saudi Arabia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

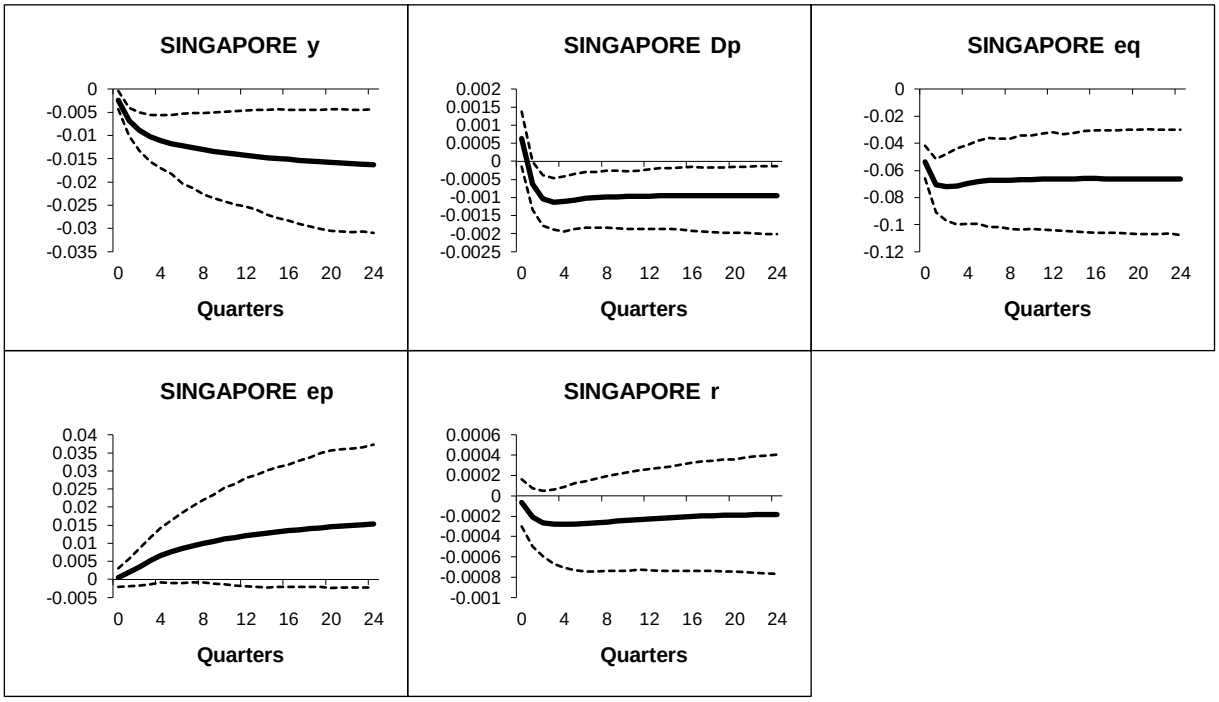

Figure B.1.17: Singapore - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

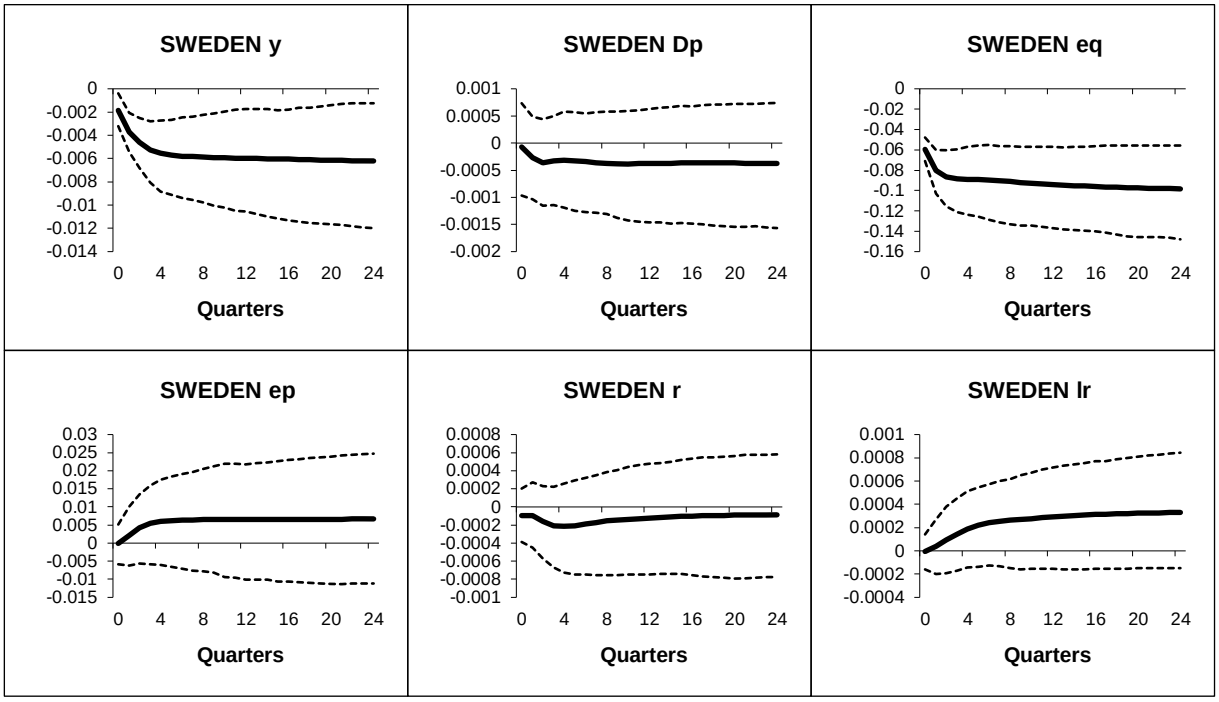

Figure B.1.18: Sweden - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

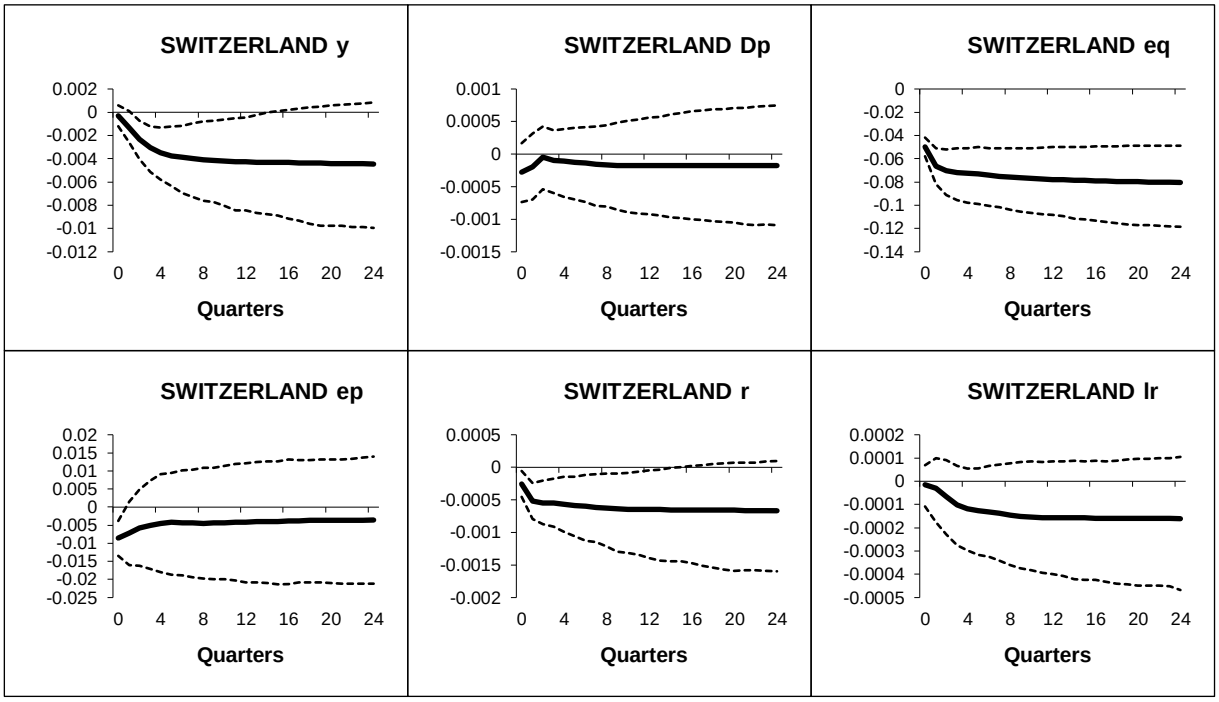

Figure B.1.19: Switzerland - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

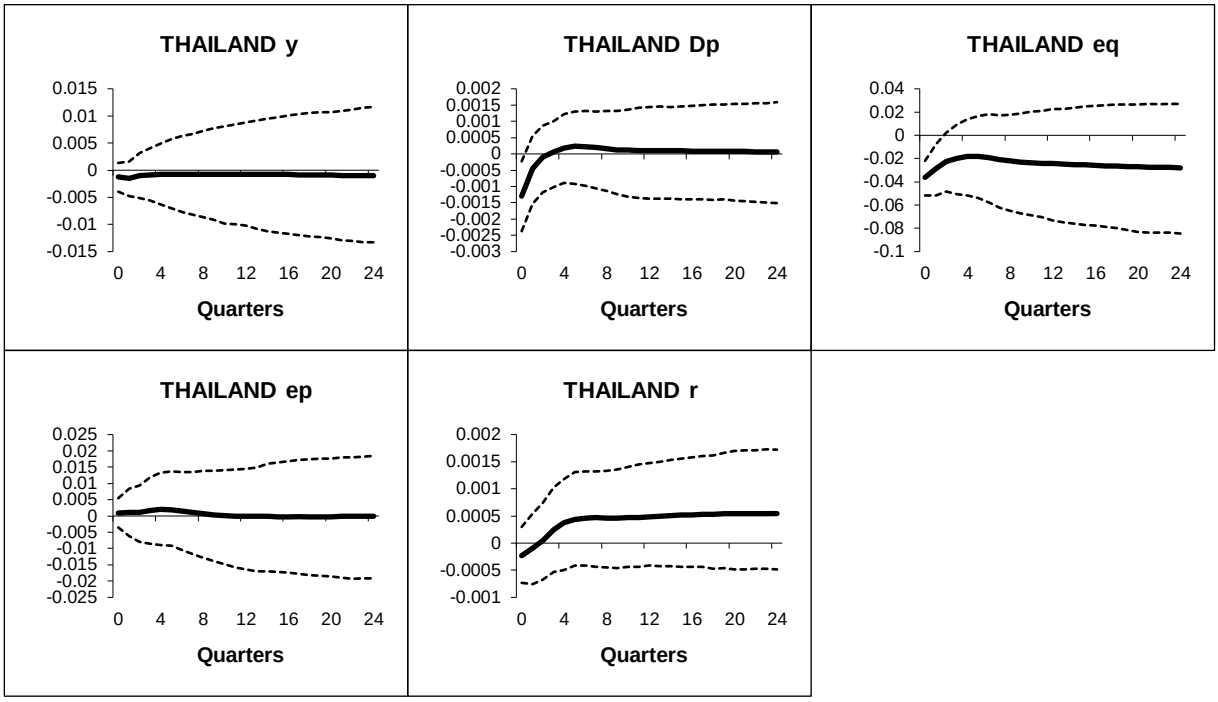

Figure B.1.20: Thailand - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

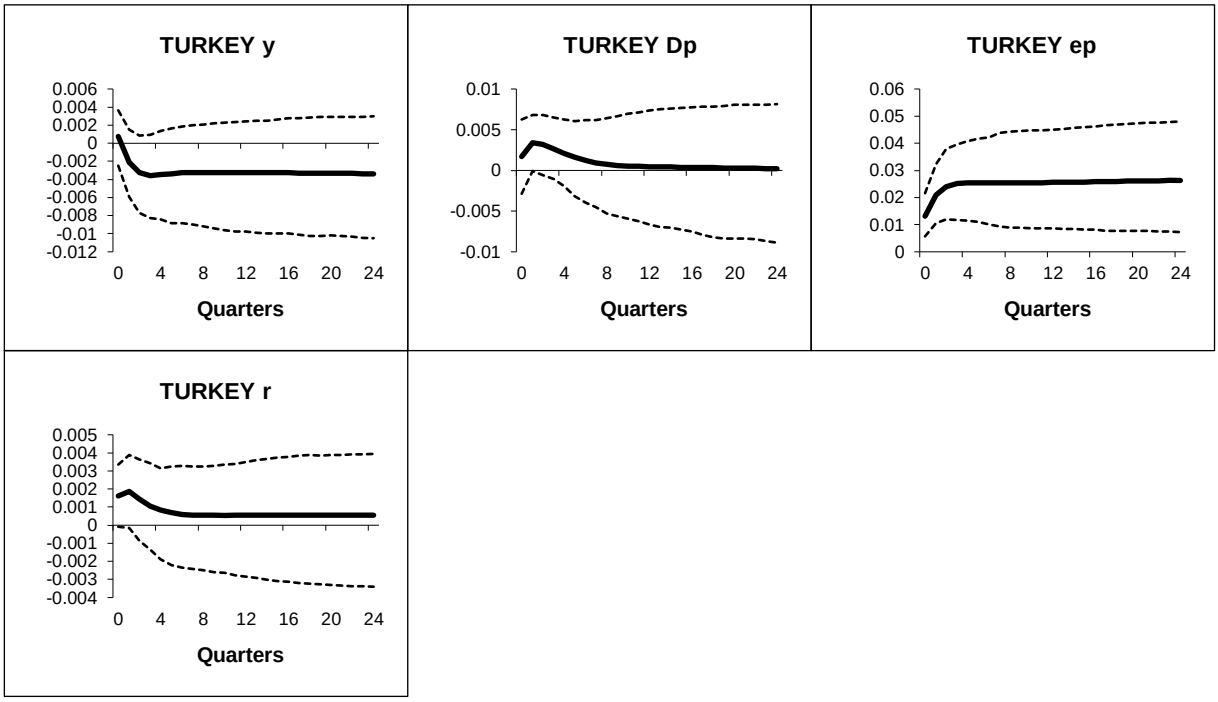

Figure B.1.21: Turkey - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

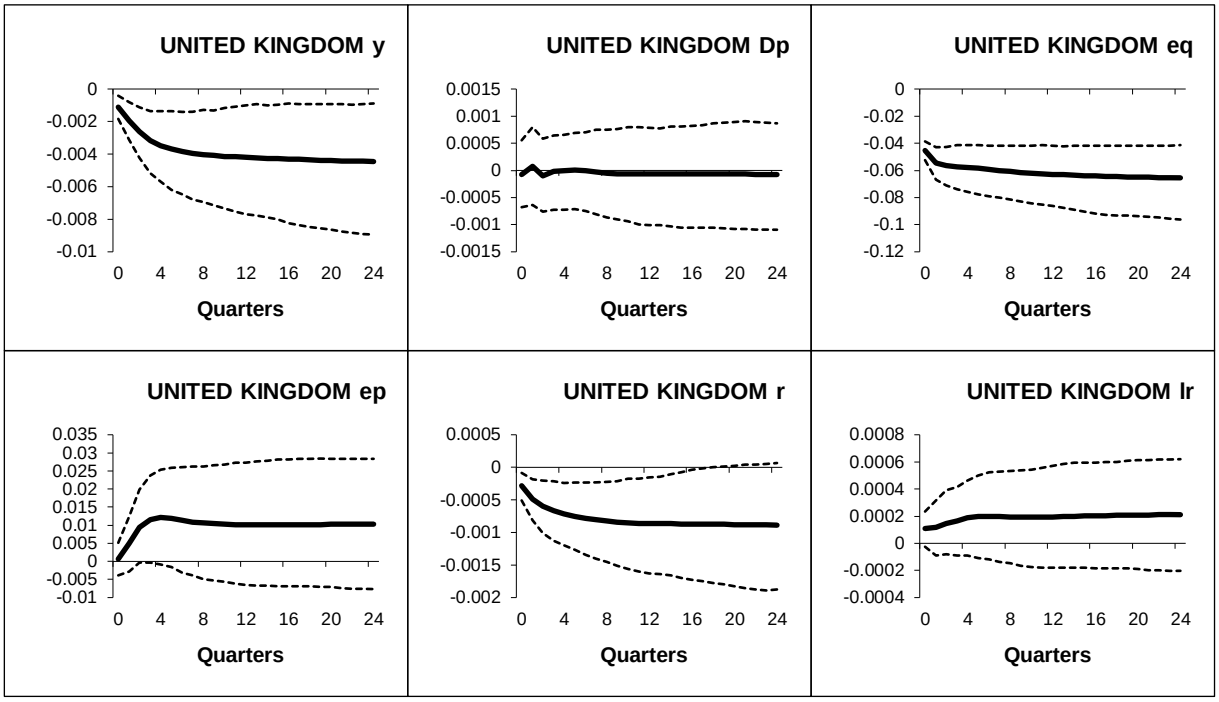

Figure B.1.22: UK - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

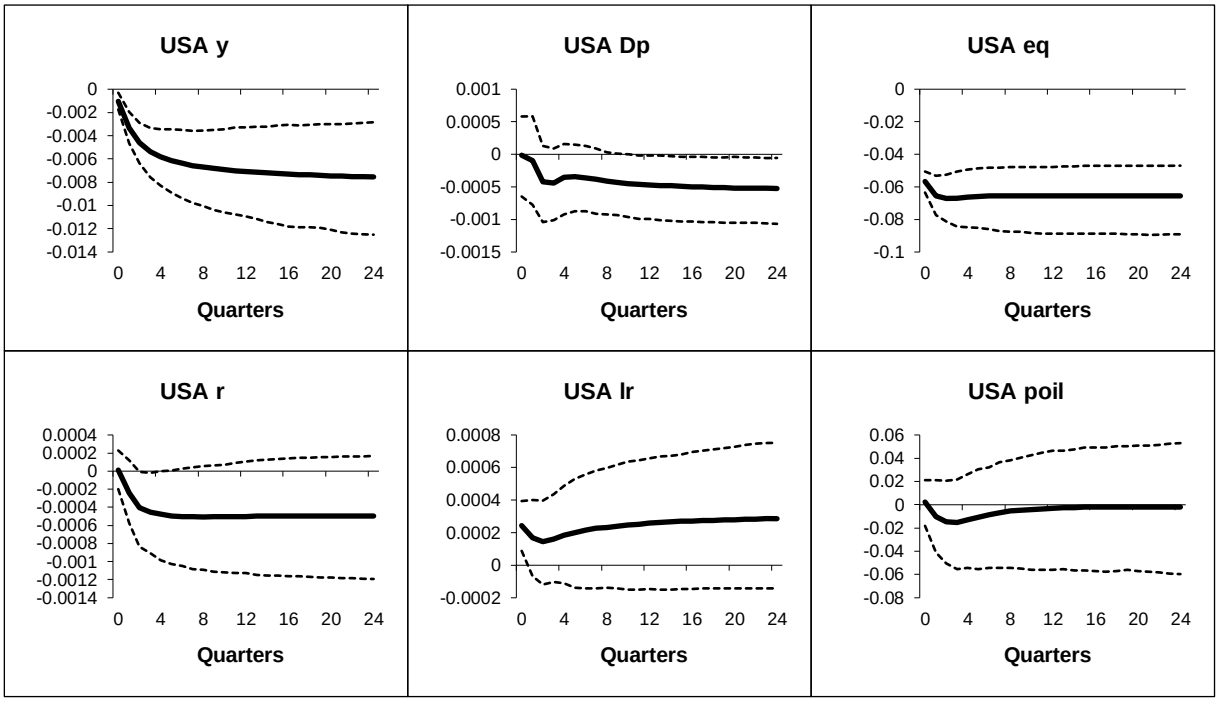

Figure B.1.23: USA - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

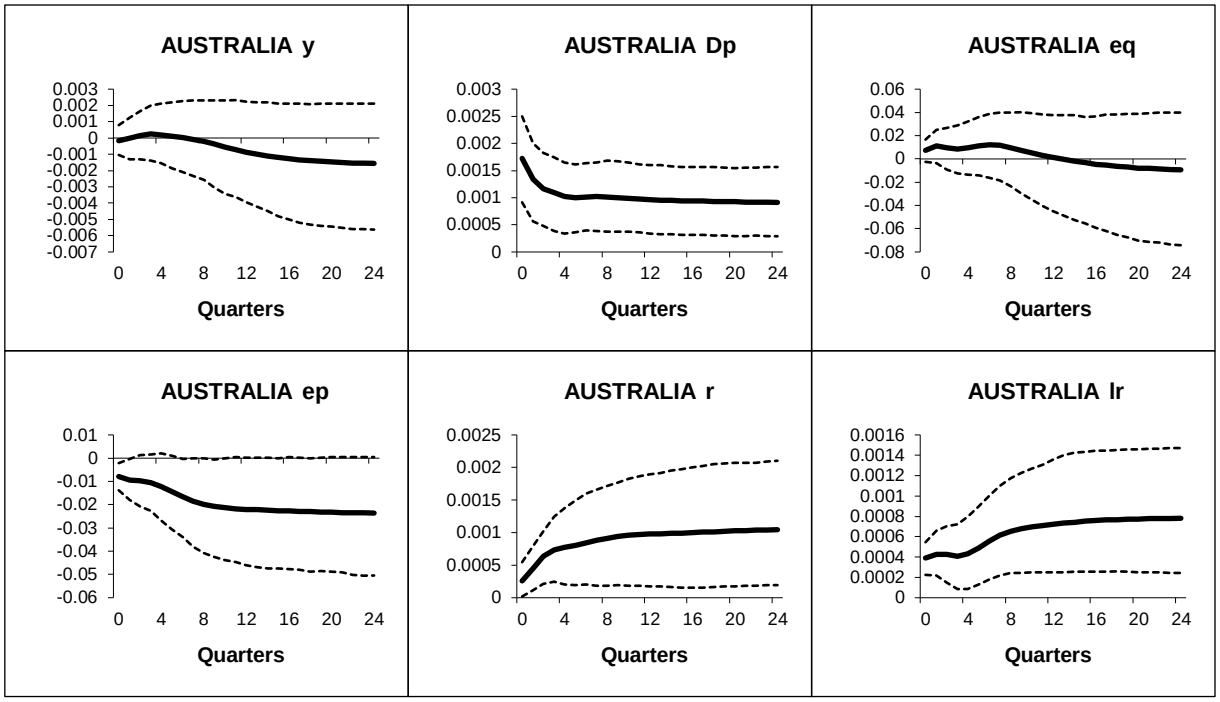

Figure B.1.24: Australia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

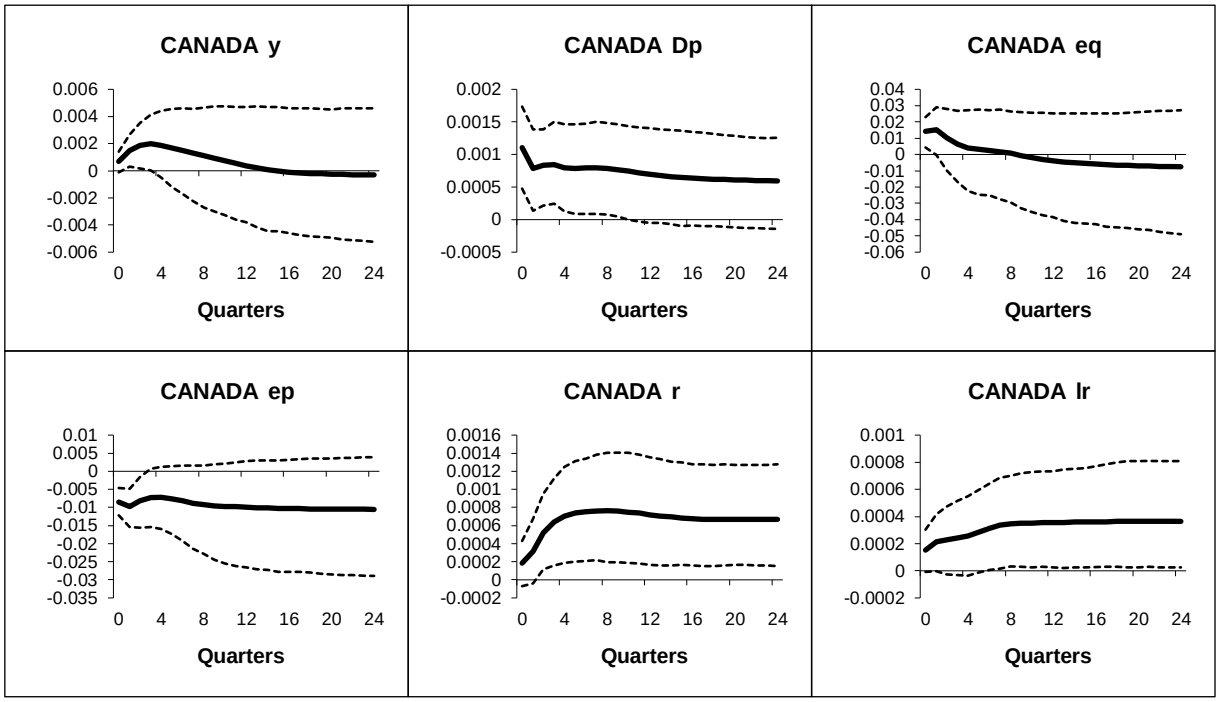

Figure B.1.25: Canada - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

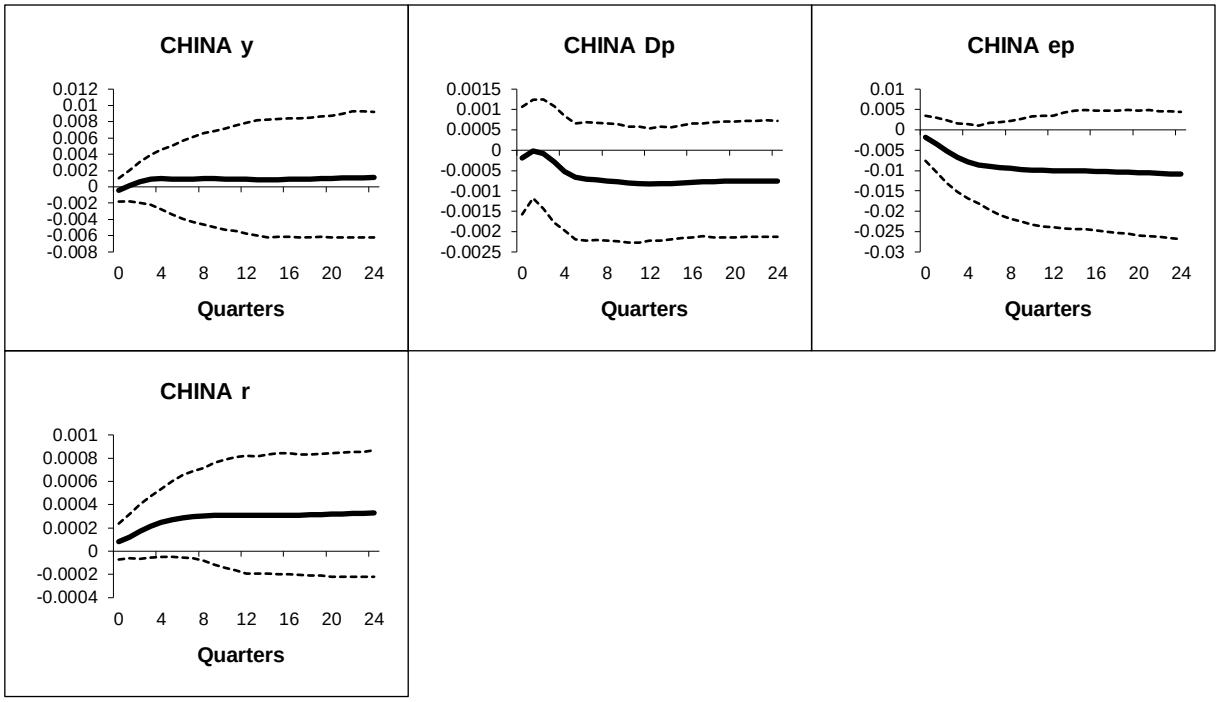

Figure B.1.26: China - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

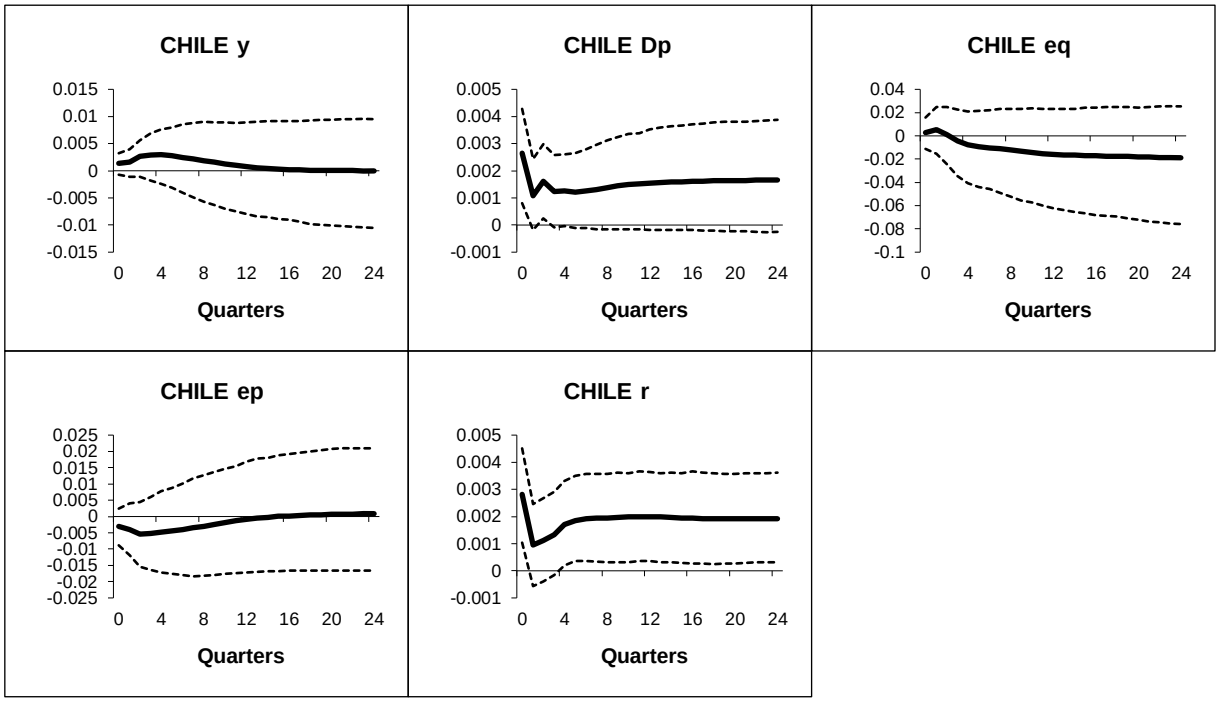

Figure B.1.27: Chile - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

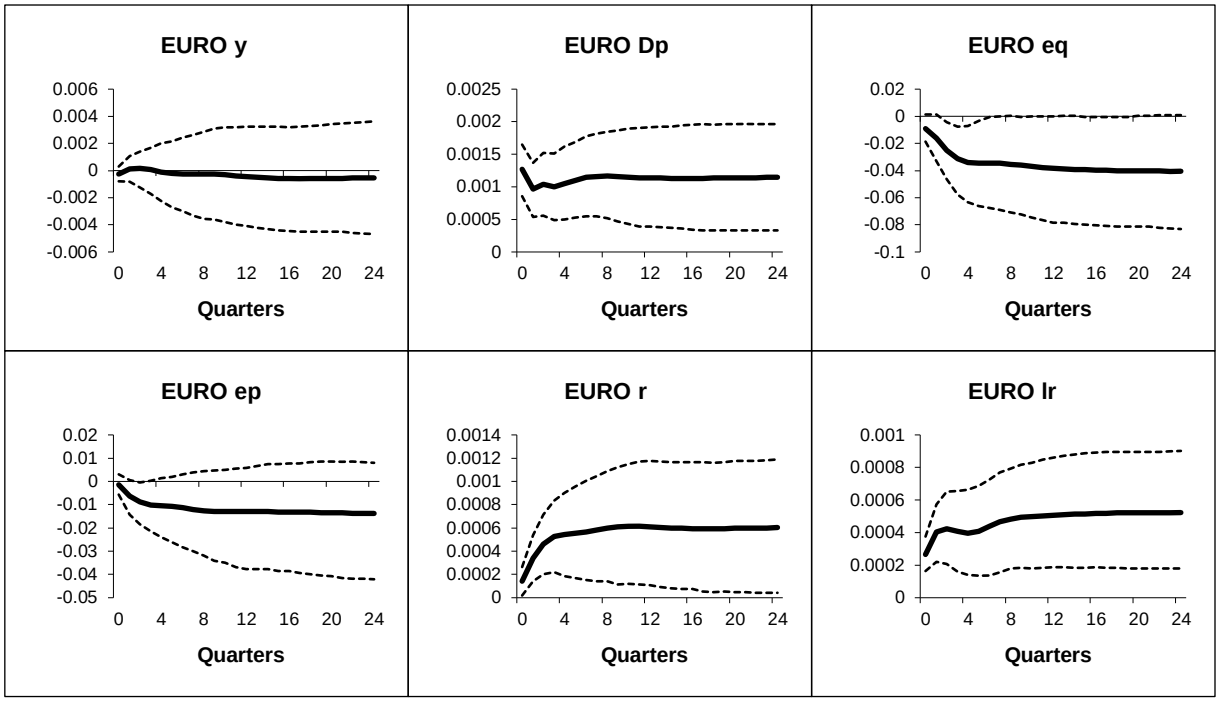

Figure B.1.28: Euro - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

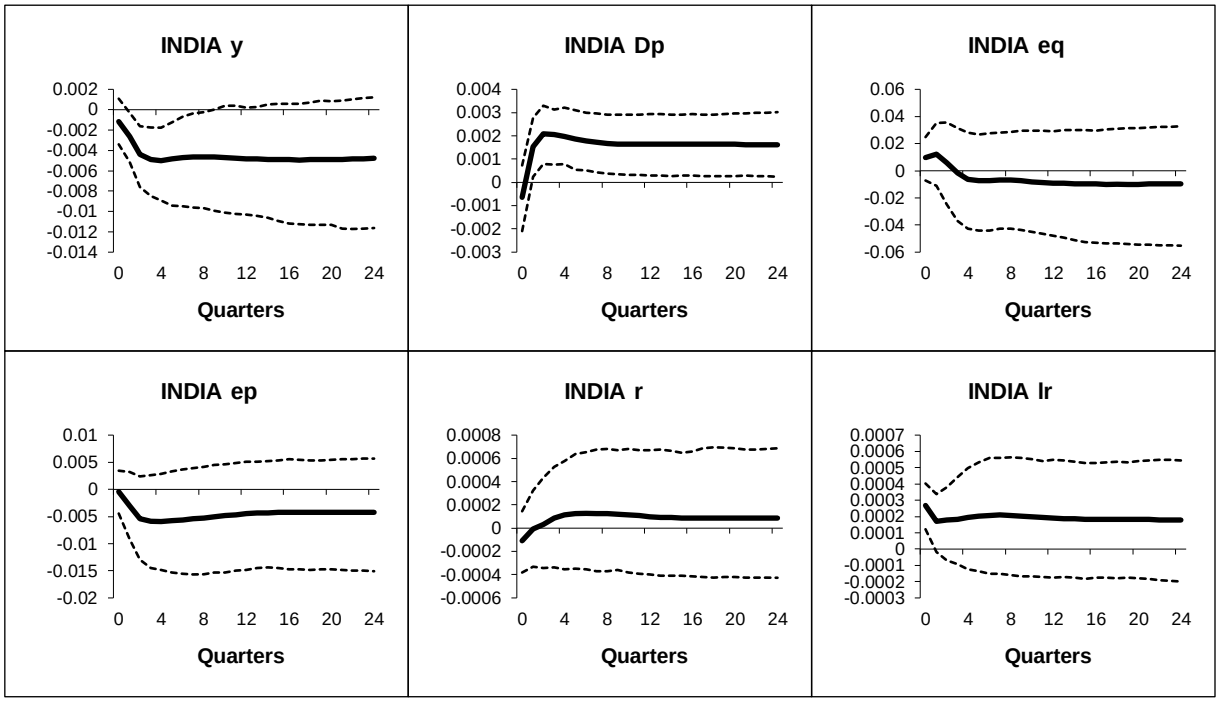

Figure B.1.29: India - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

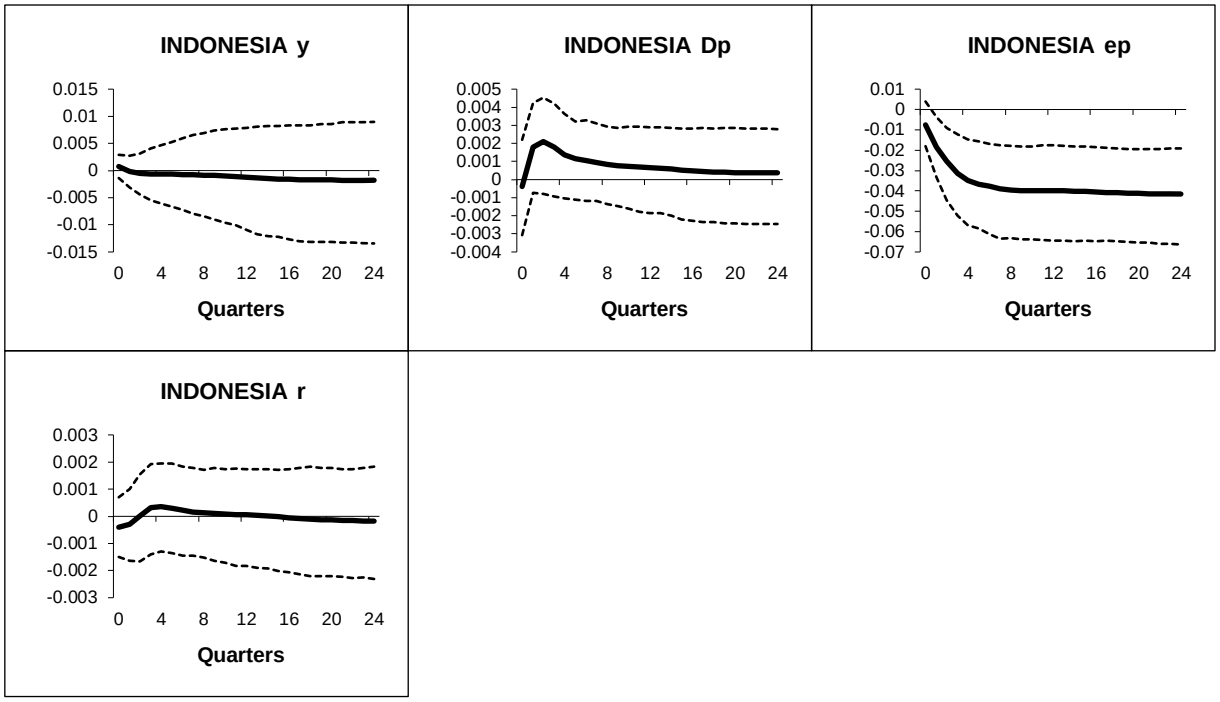

Figure B.1.30: Indonesia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

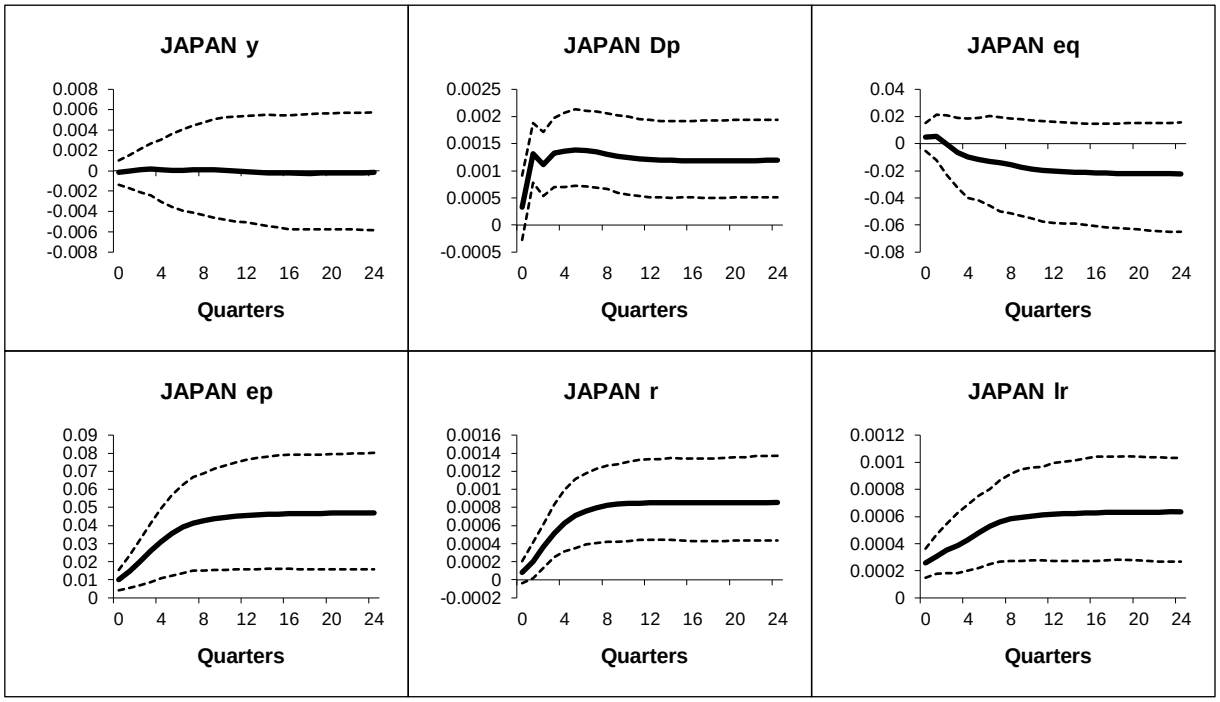

Figure B.1.31: Japan - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

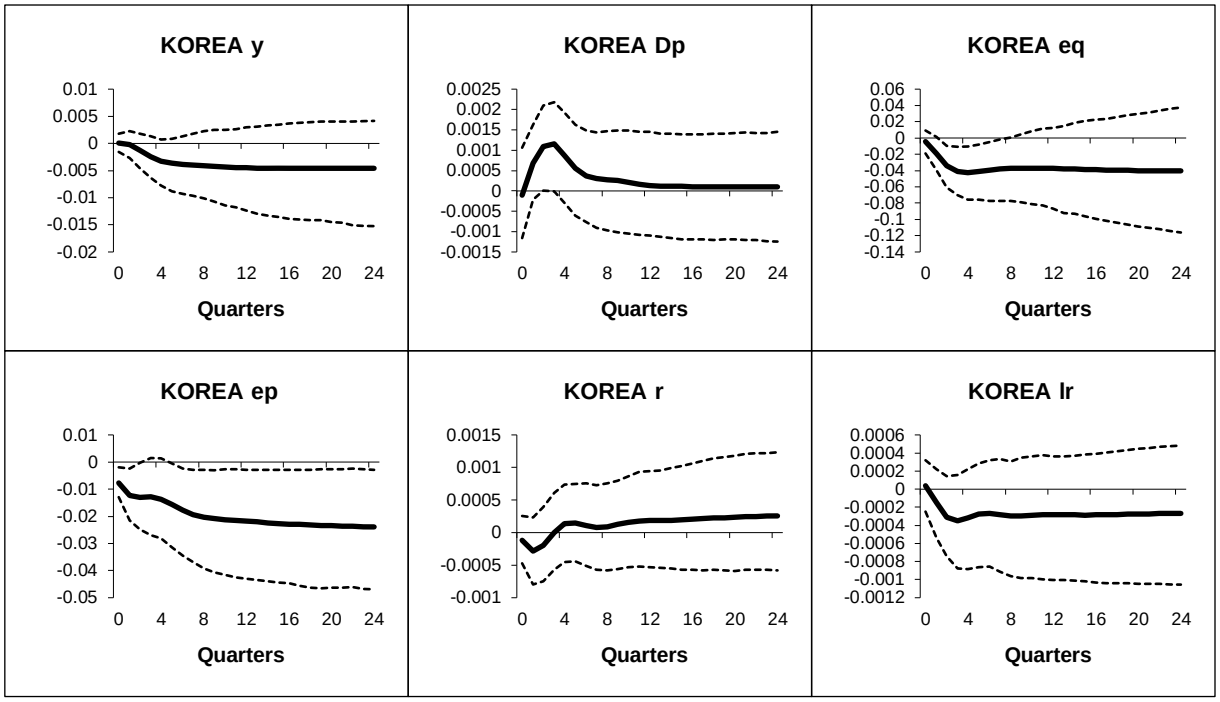

Figure B.1.32: Korea - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

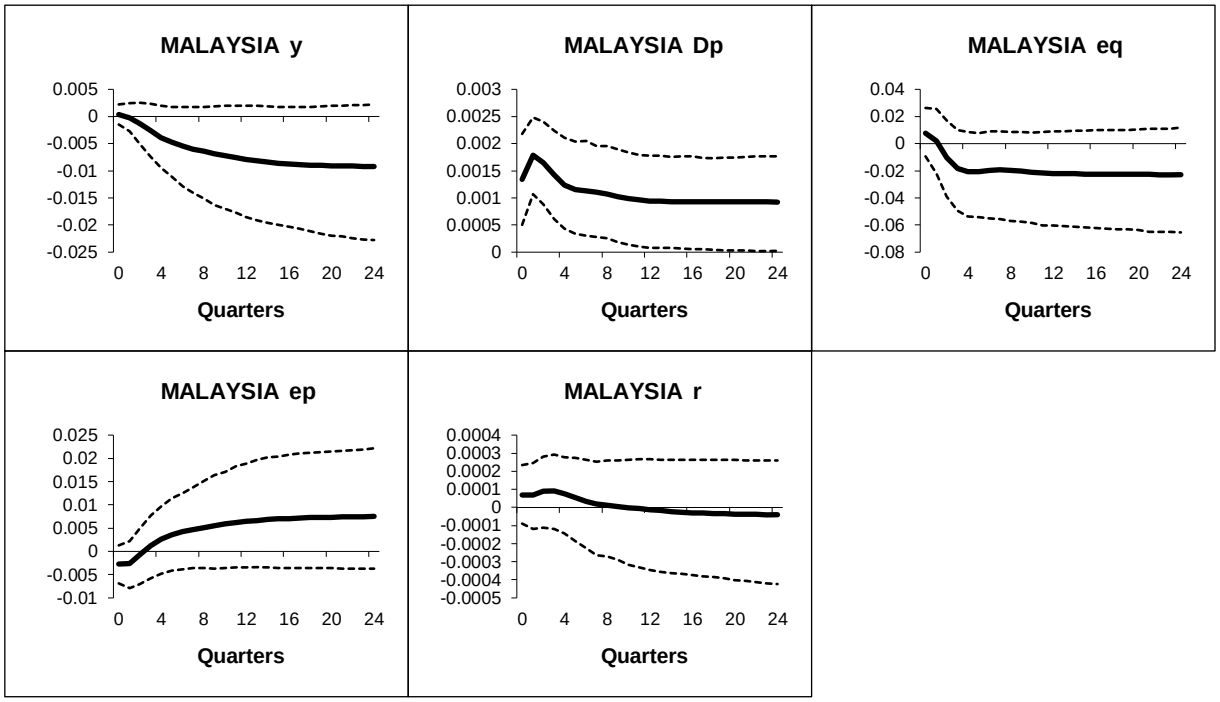

Figure B.1.33: Malaysia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

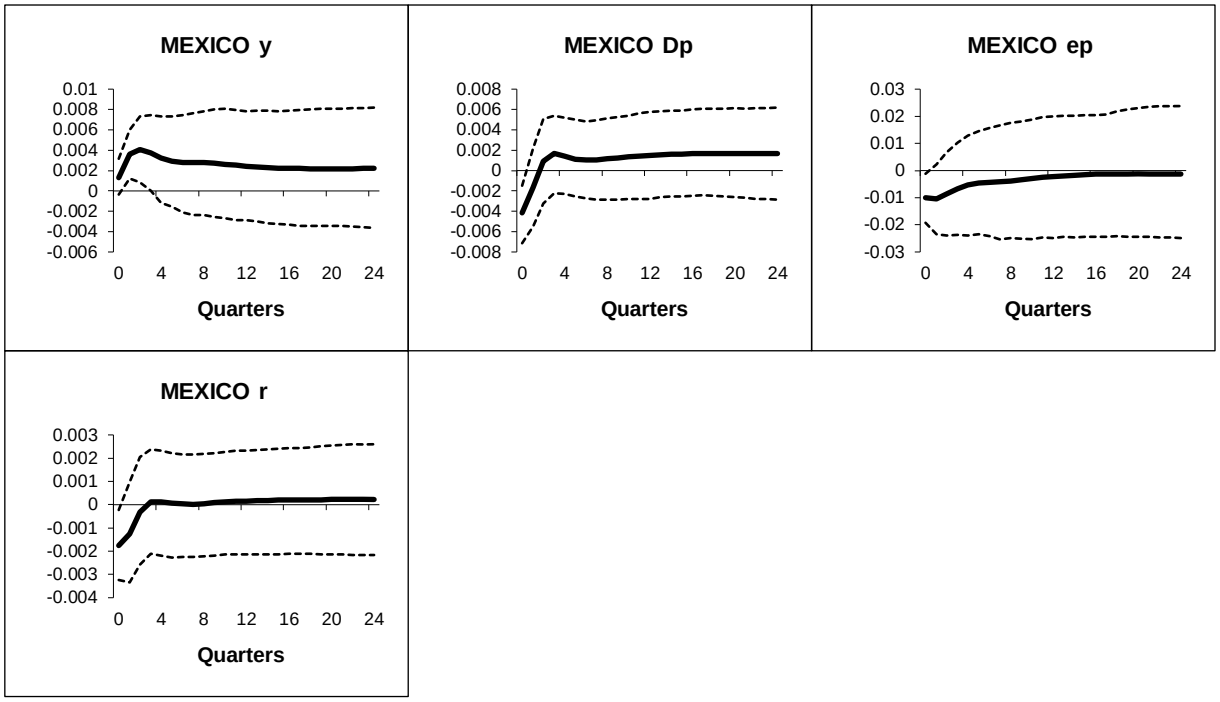

Figure B.1.34: Mexico - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

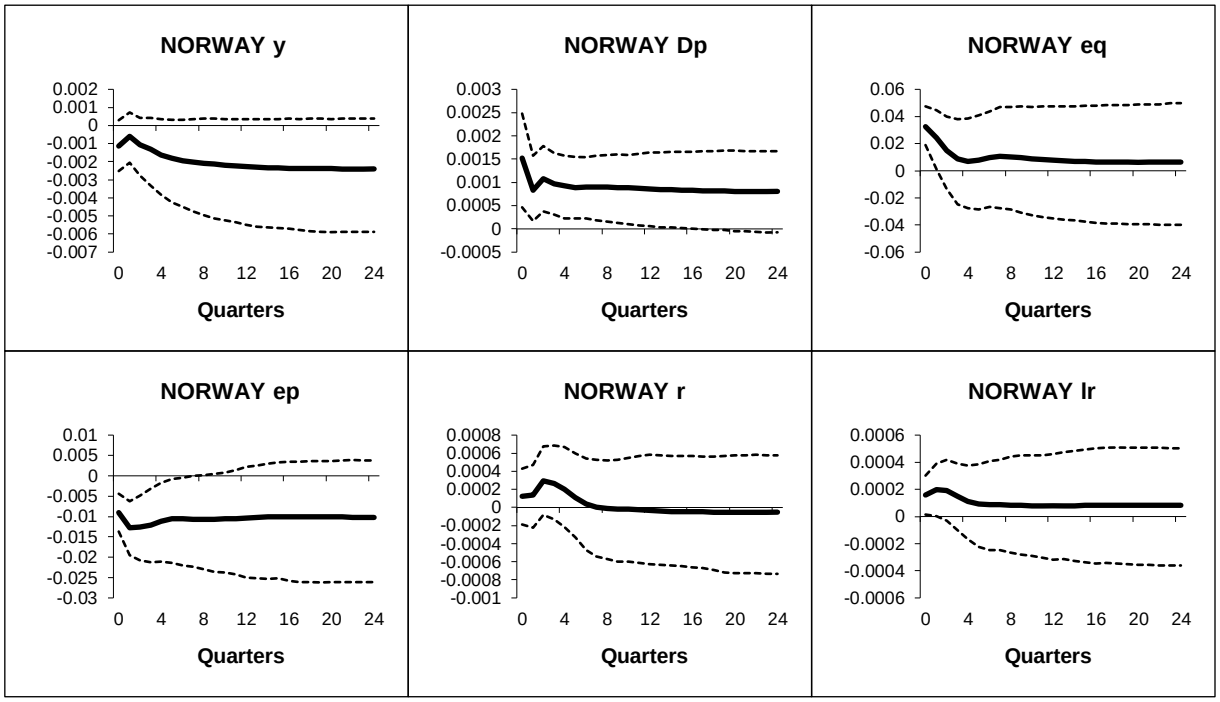

Figure B.1.35: Norway - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

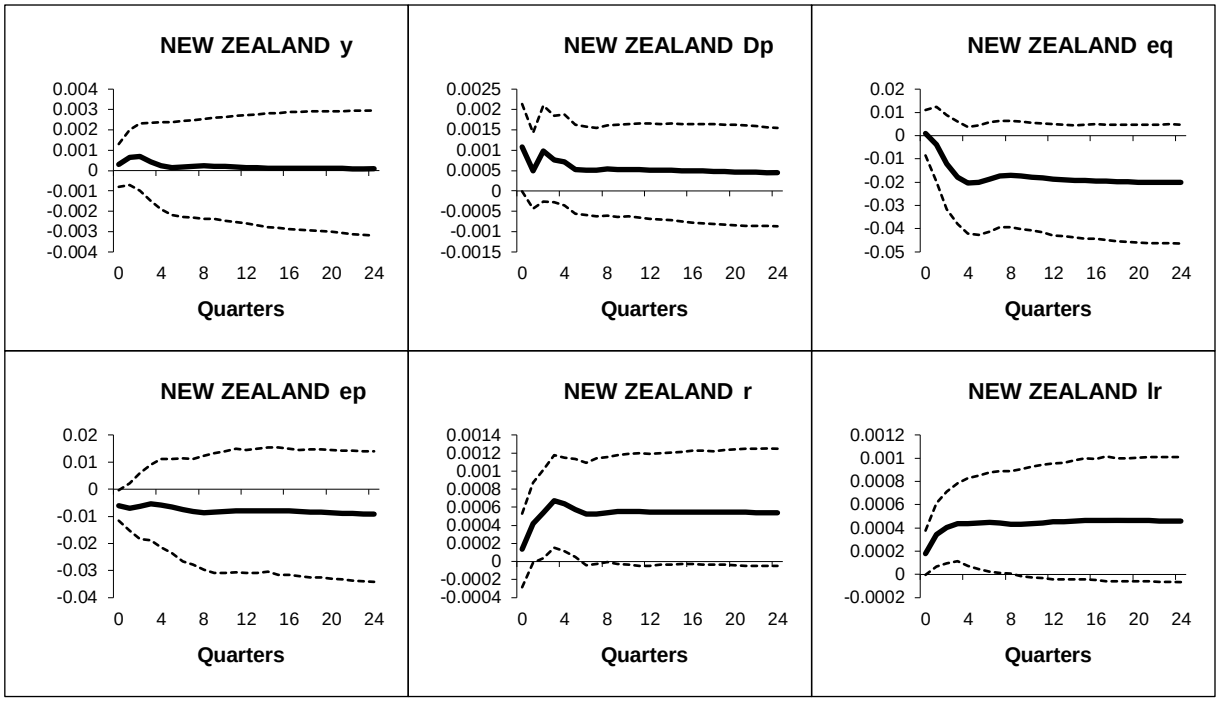

Figure B.1.36: New Zealand - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

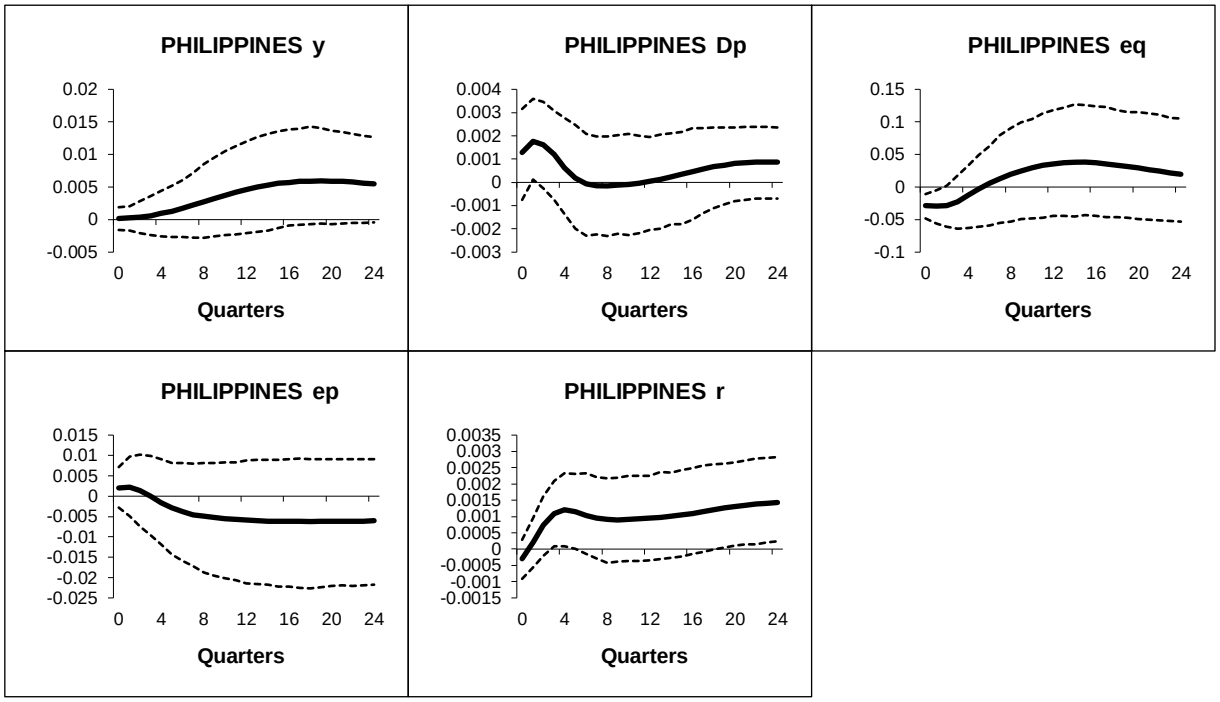

Figure B.1.37: Philippines - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

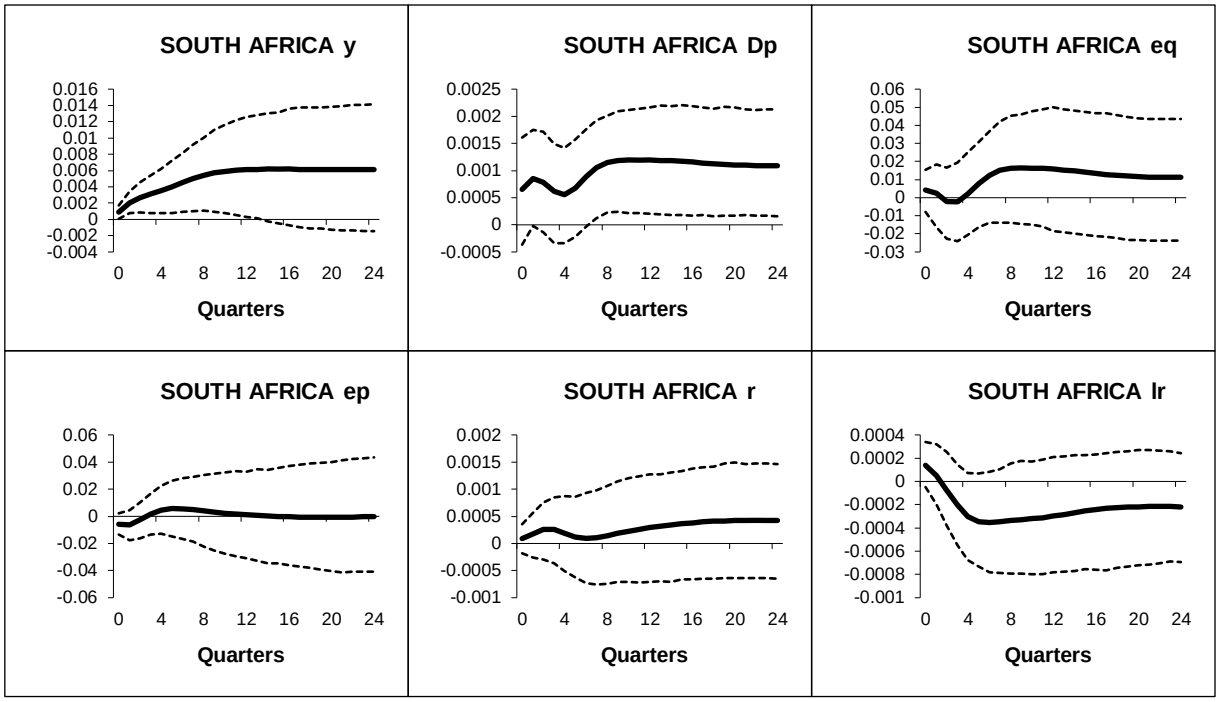

Figure B.1.38: South Africa - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

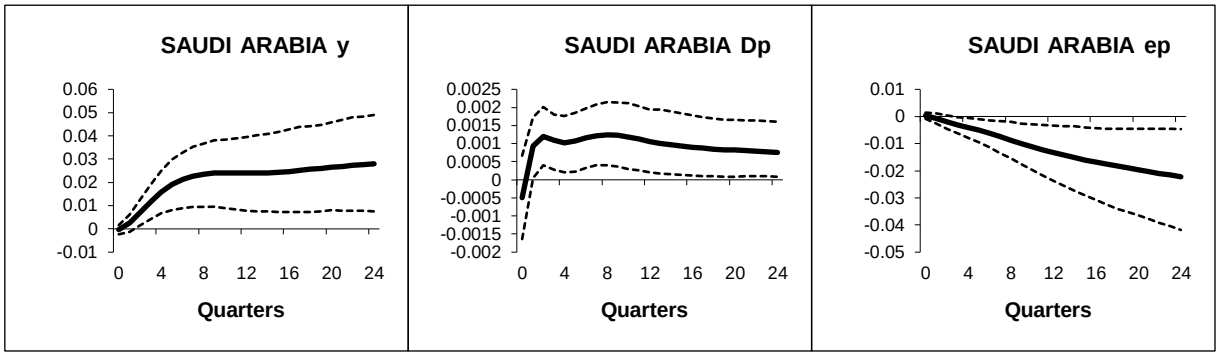

Figure B.1.39: Saudi Arabia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

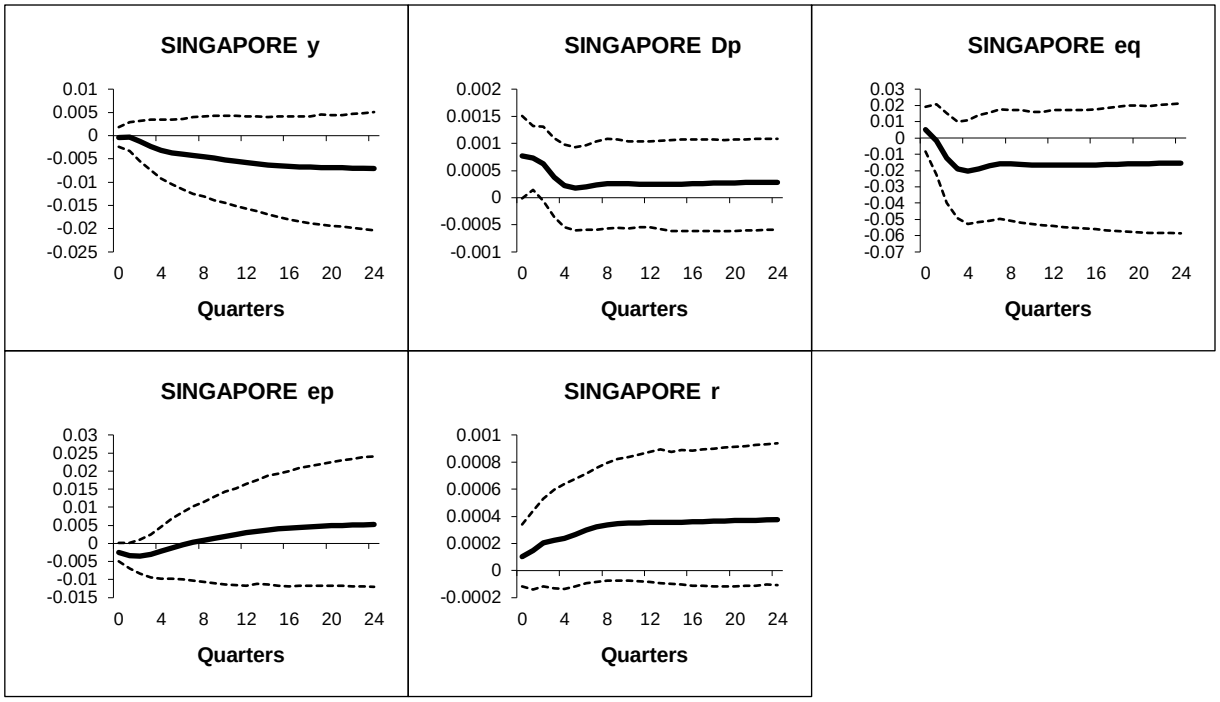

Figure B.1.40: Singapore - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

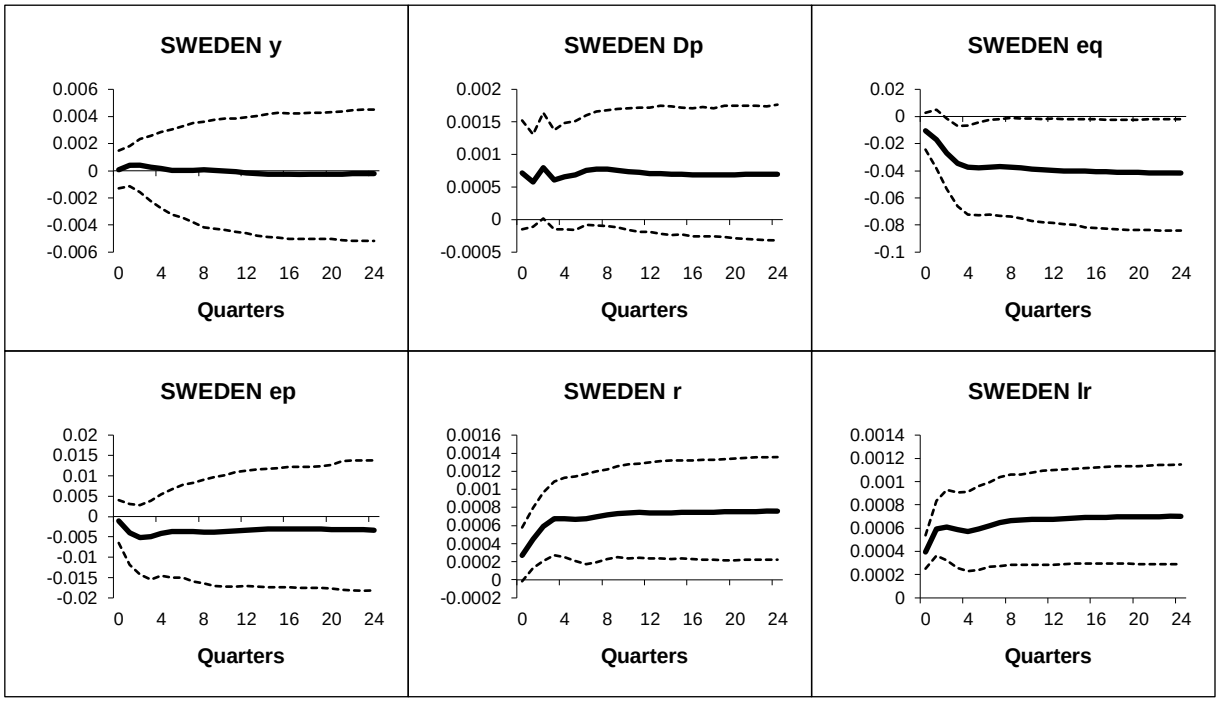

Figure B.1.41: Sweden - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

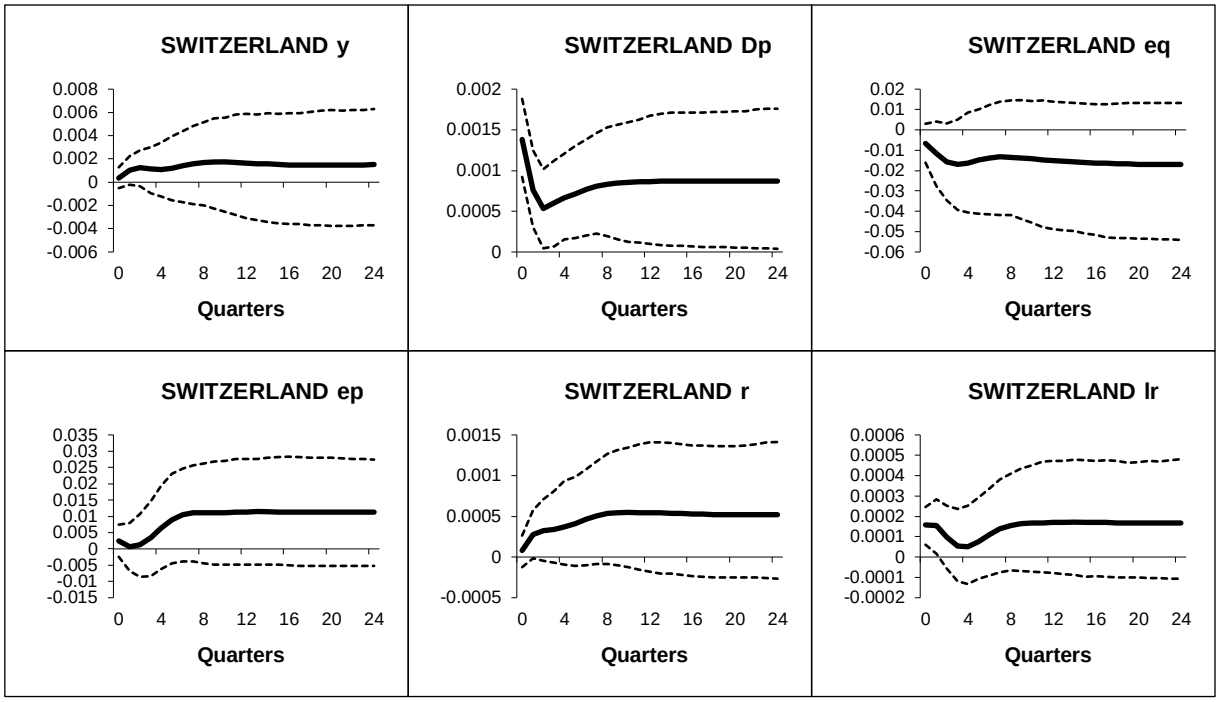

Figure B.1.42: Switzerland - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

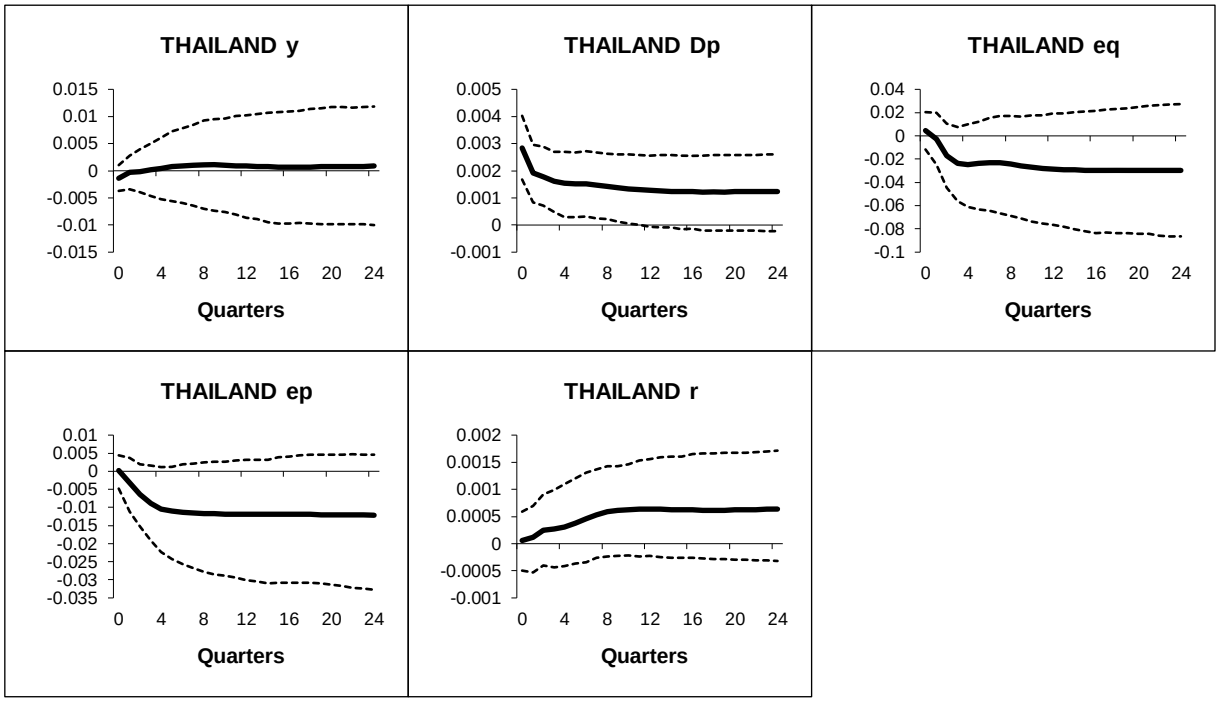

Figure B.1.43: Thailand - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

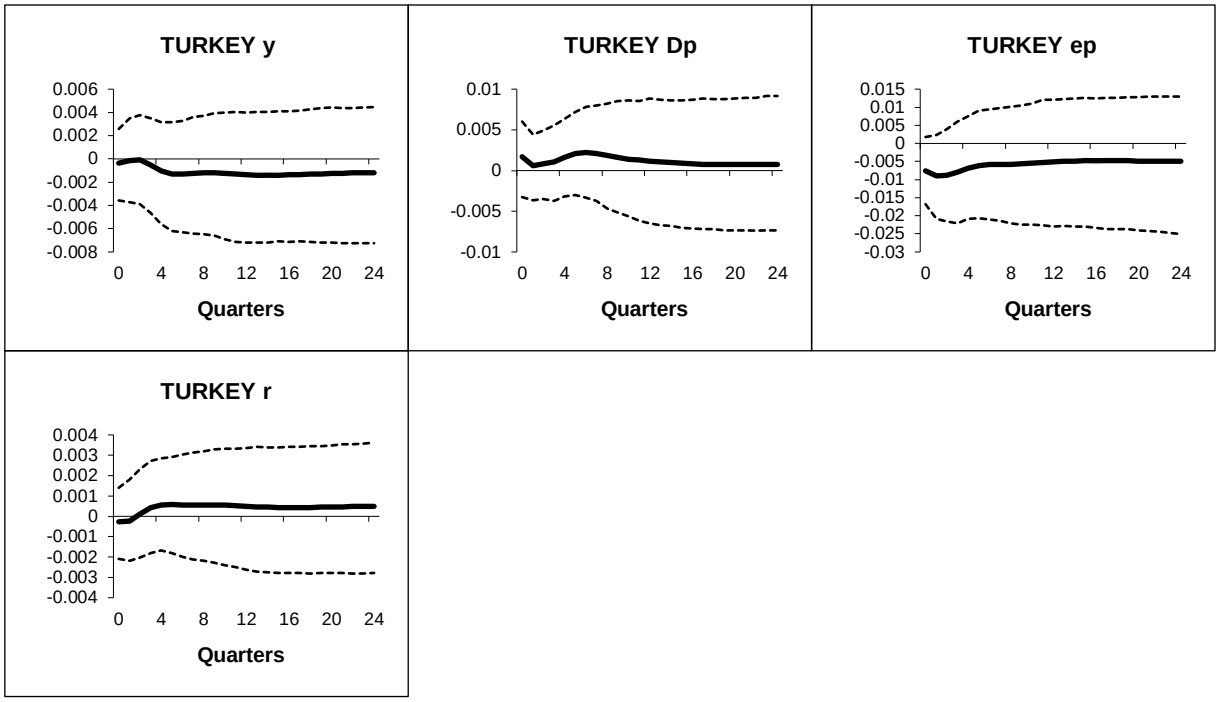

Figure B.1.44: Turkey - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

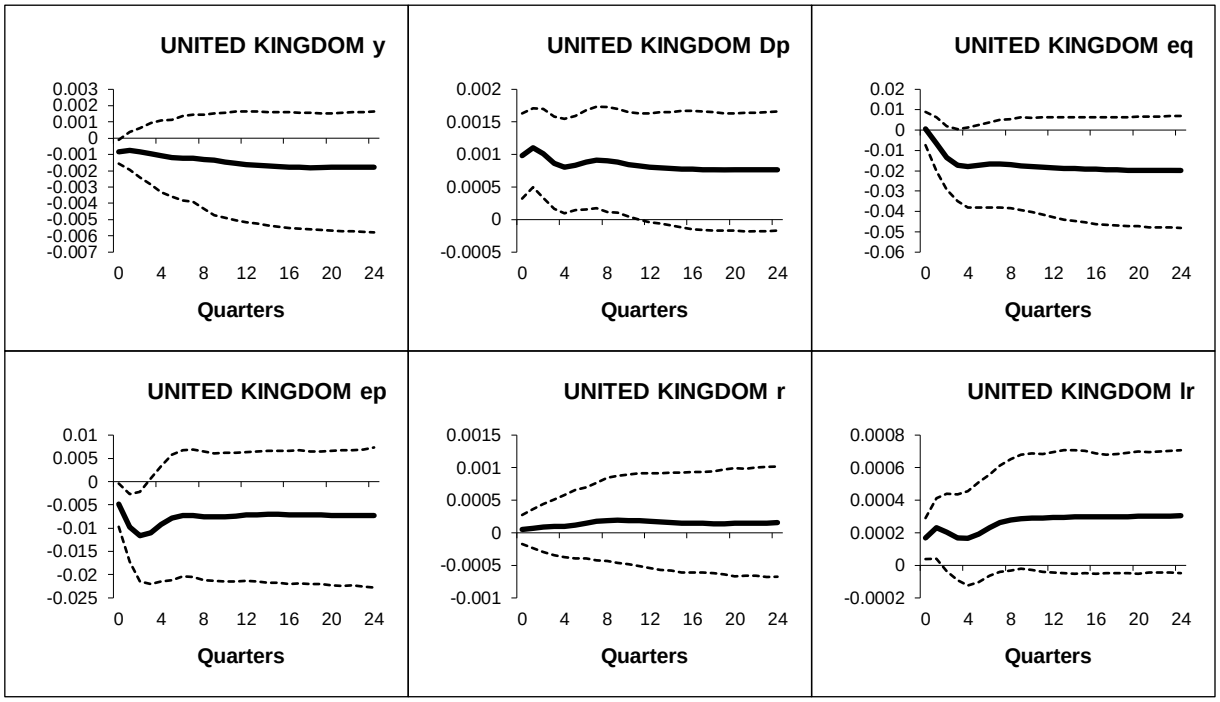

Figure B.1.45: UK - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

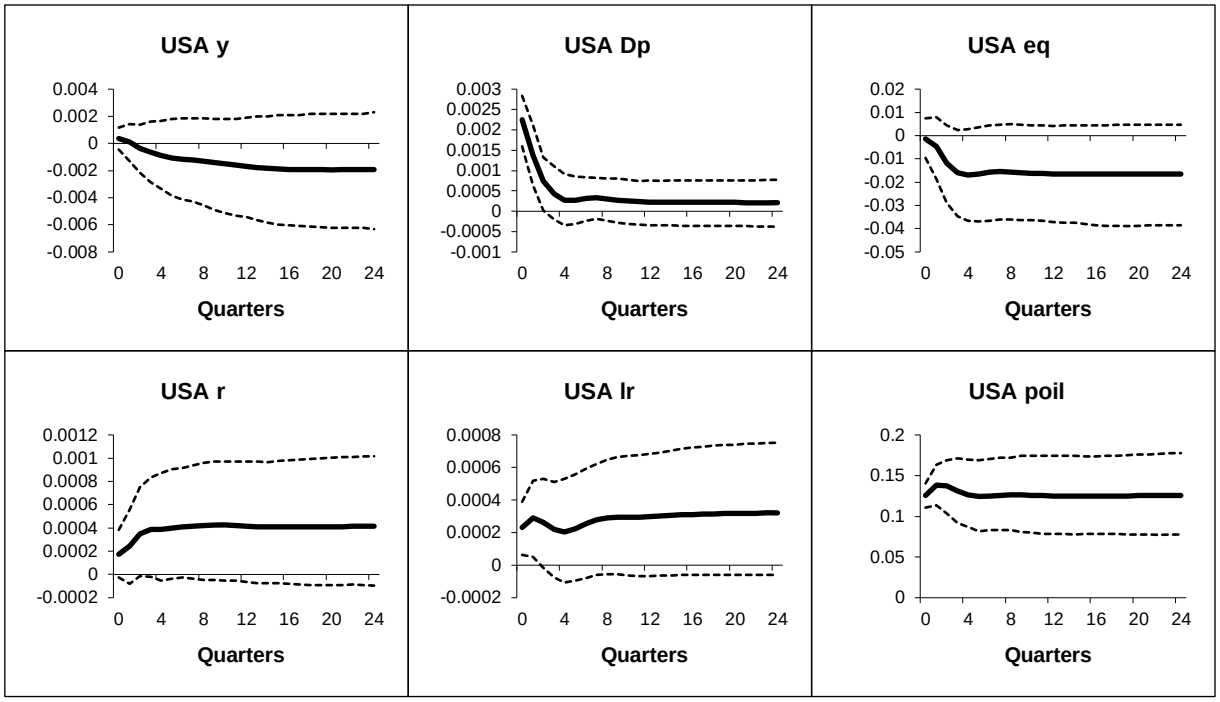

Figure B.1.46: USA - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

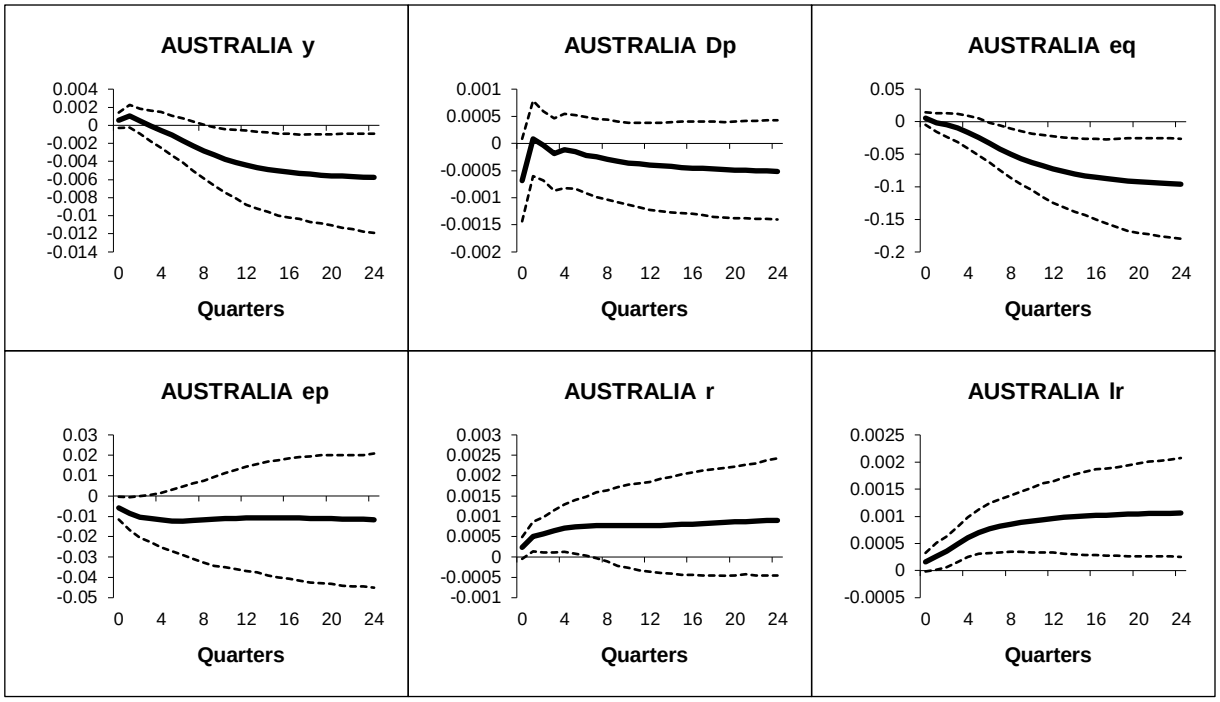

Figure B.1.47: Australia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

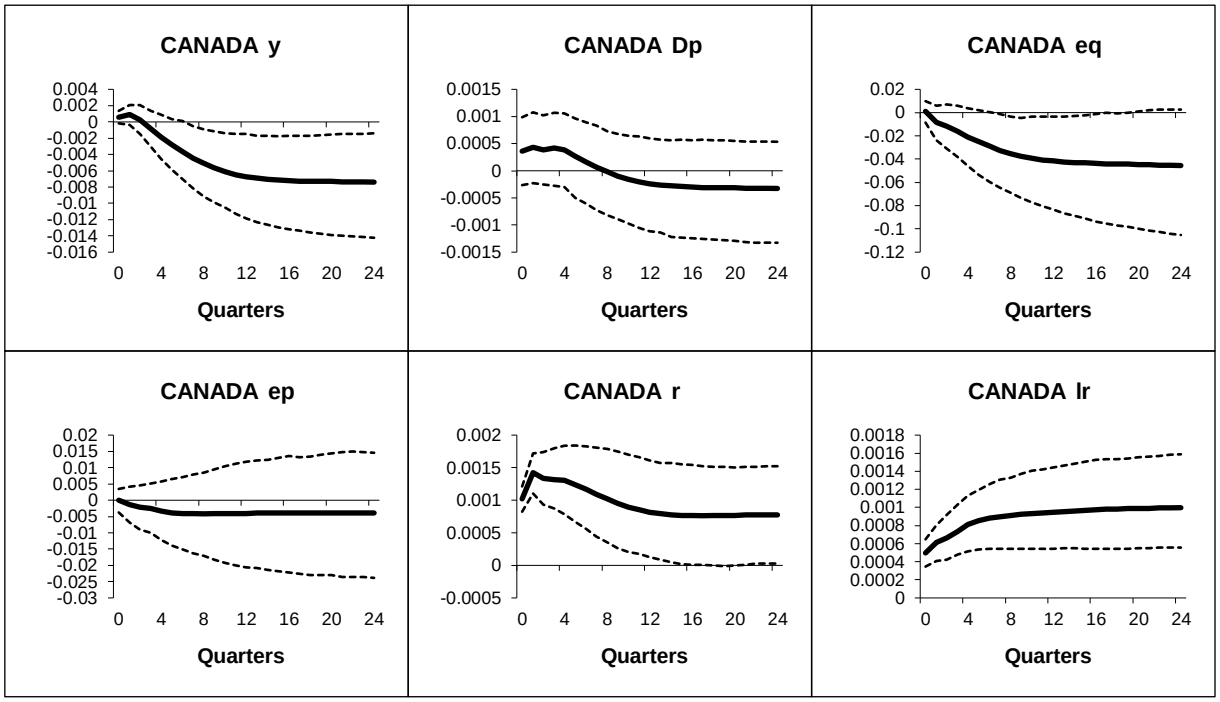

Figure B.1.48: Canada - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

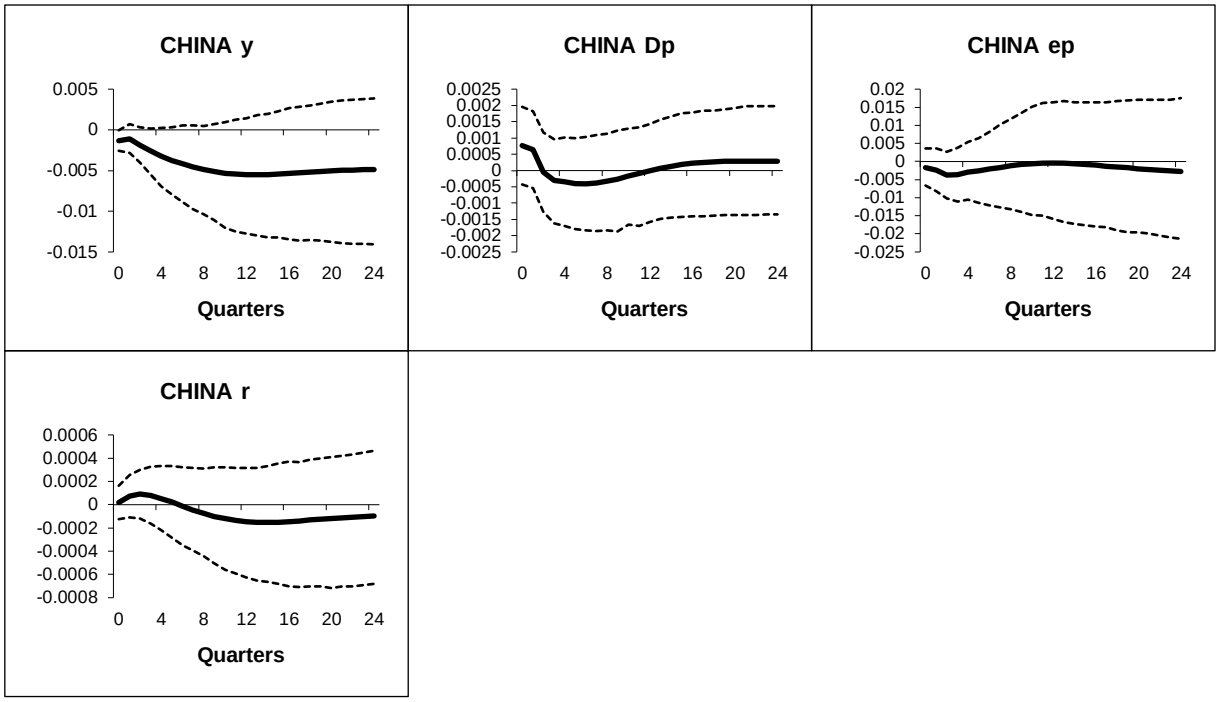

Figure B.1.49: China - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

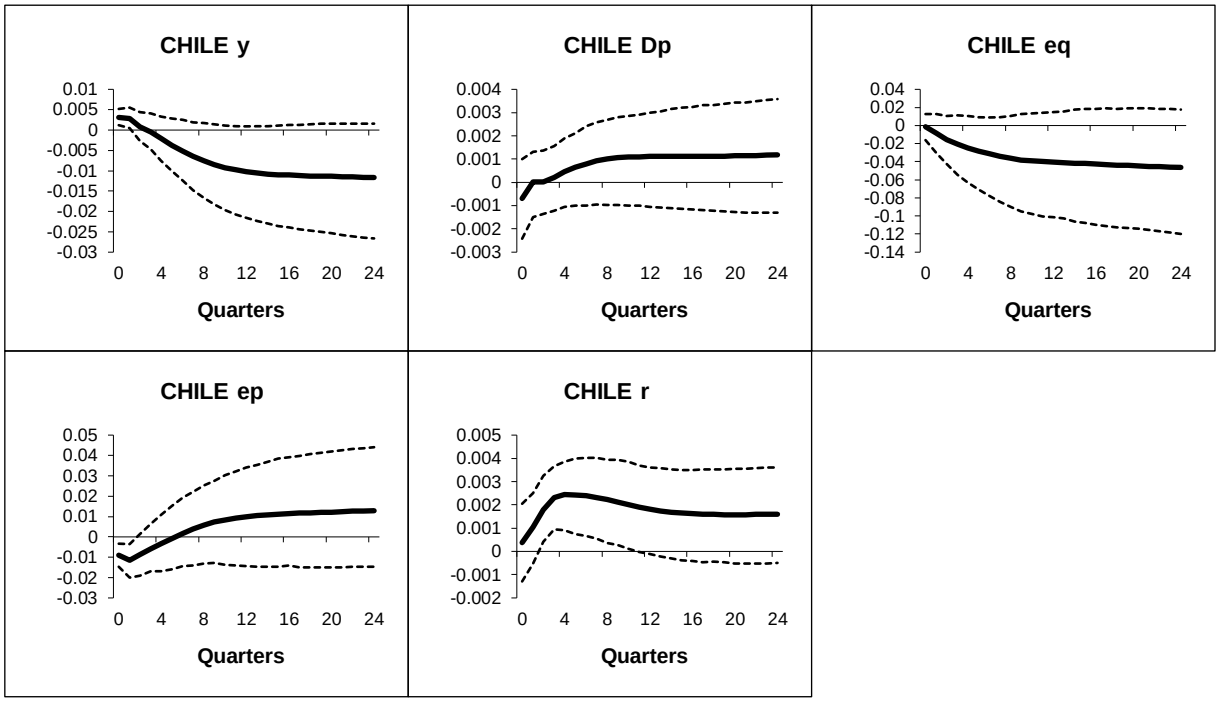

Figure B.1.50: Chile - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

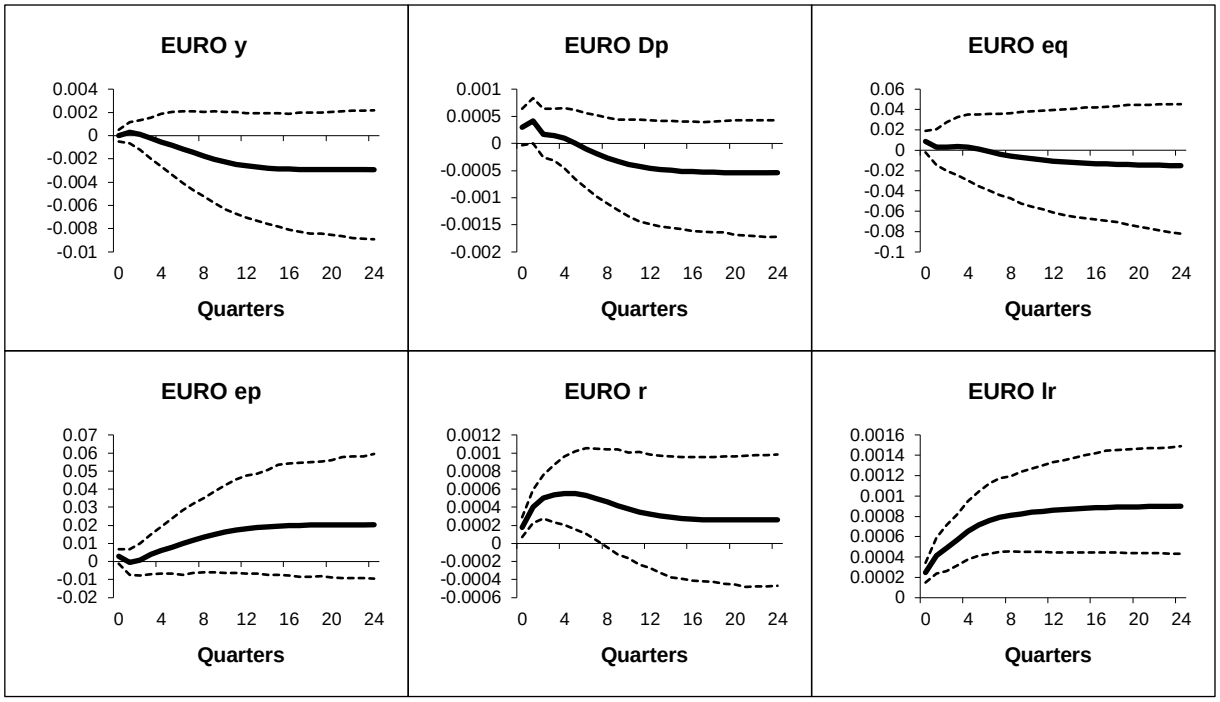

Figure B.1.51: Euro - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

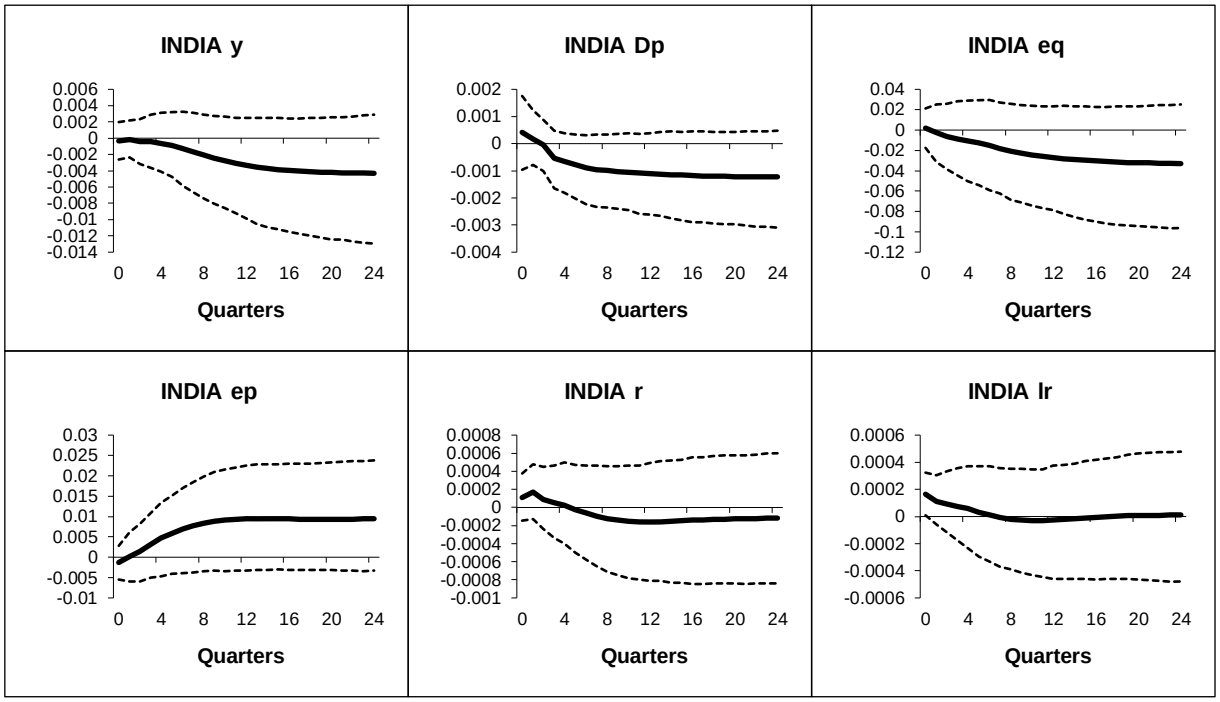

Figure B.1.52: India - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

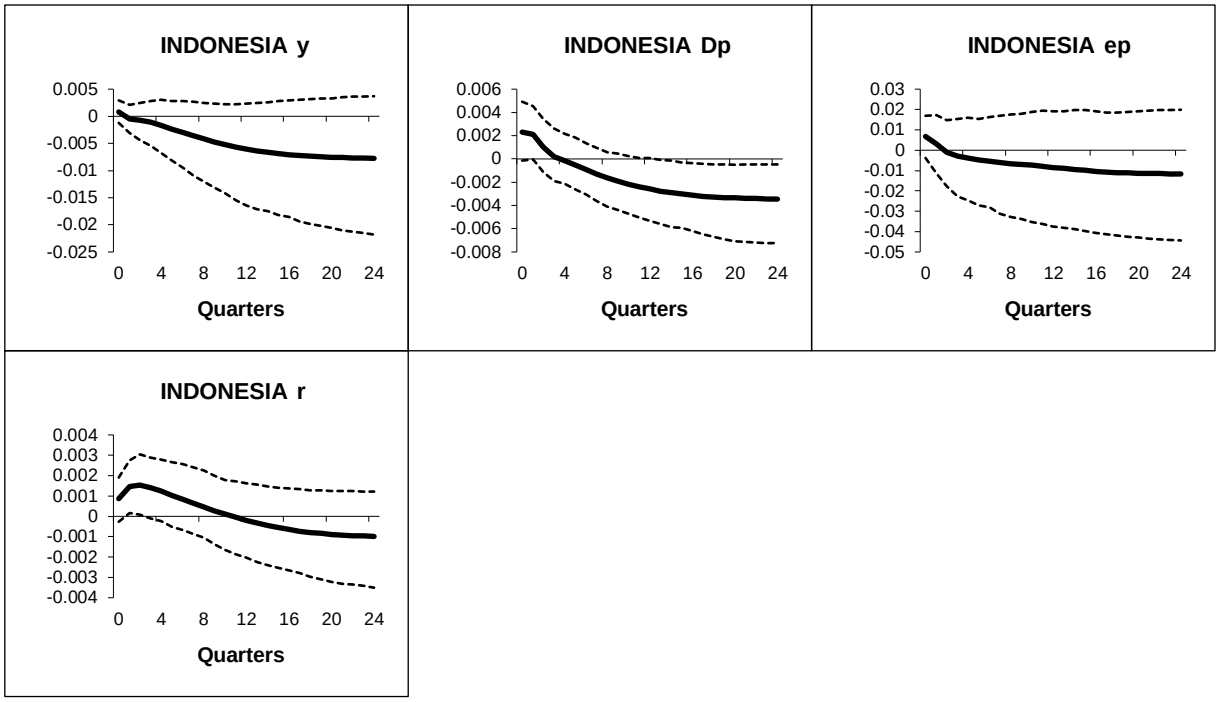

Figure B.1.53: Indonesia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

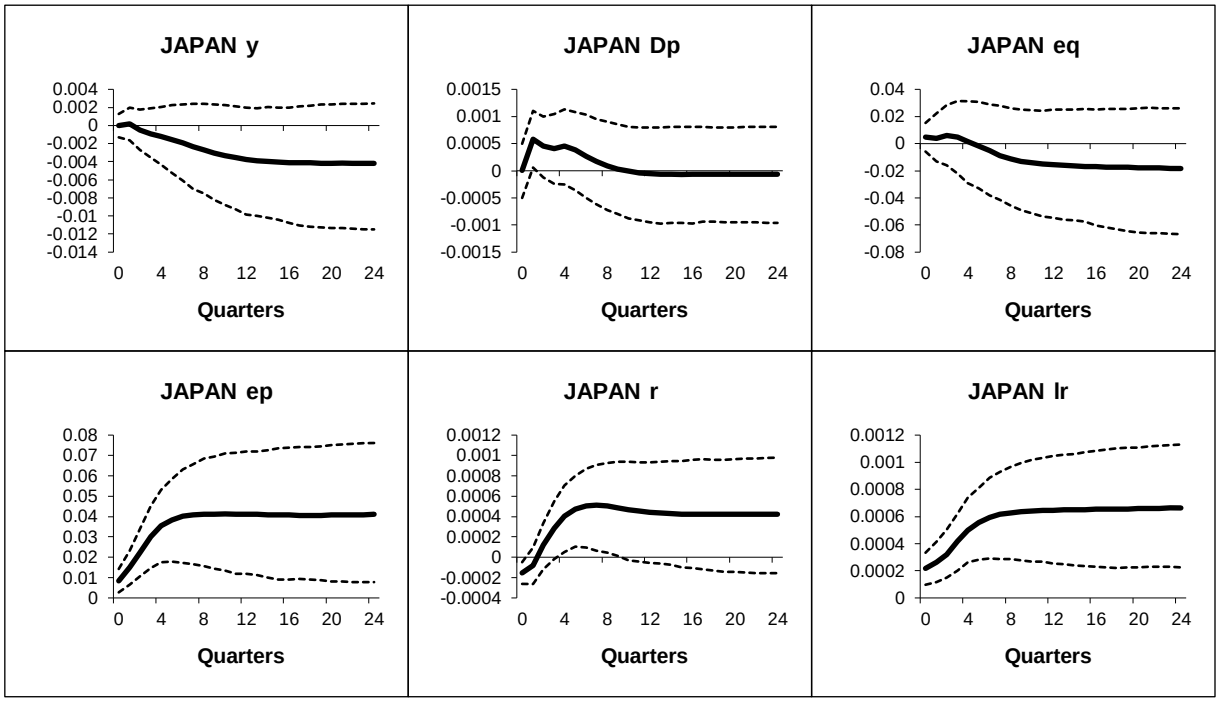

Figure B.1.54: Japan - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

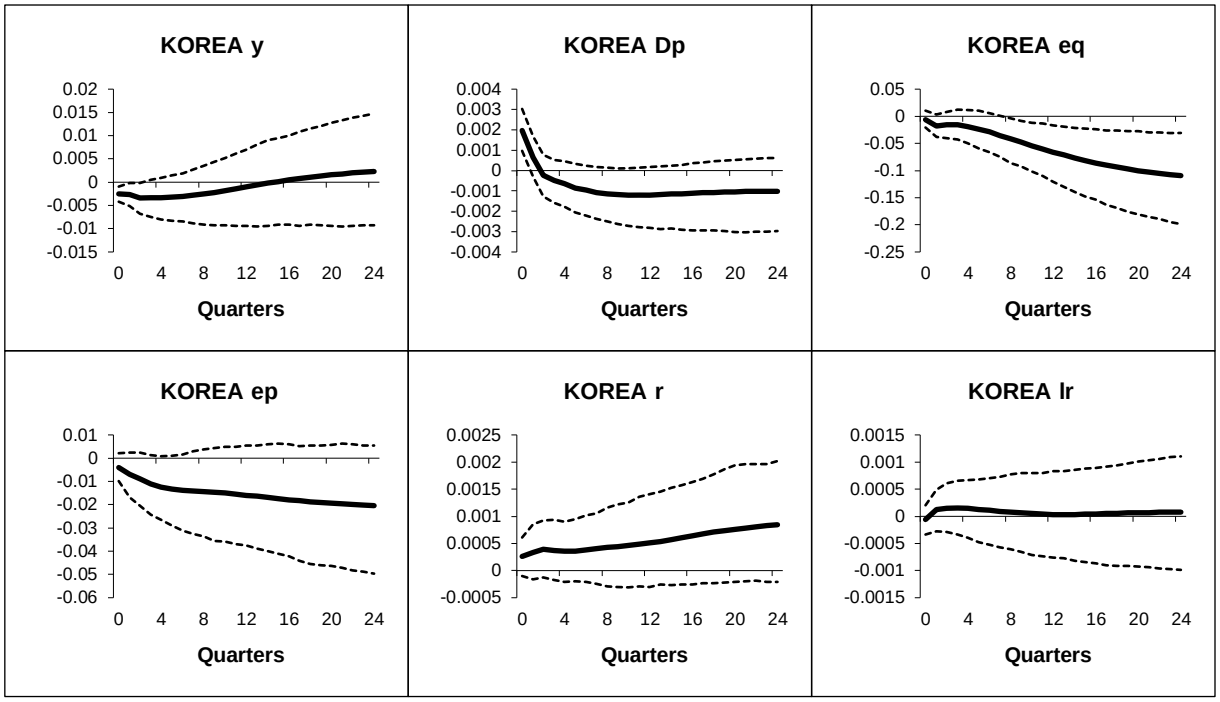

Figure B.1.55: Korea - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

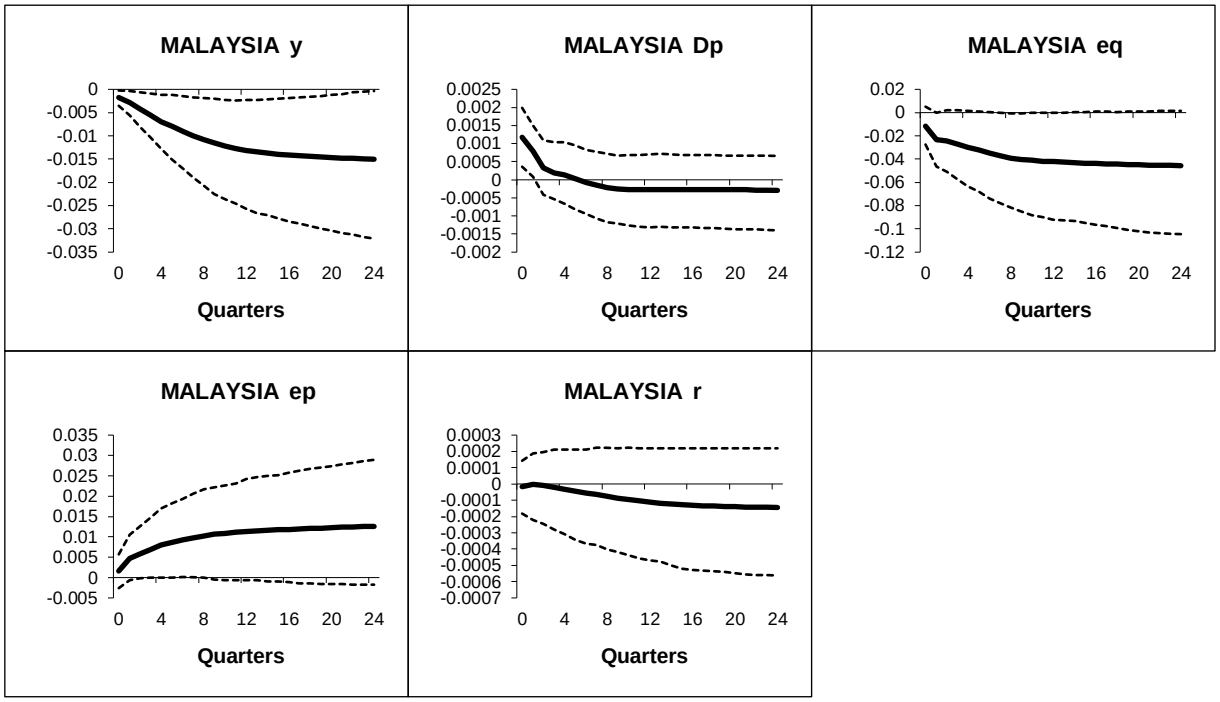

Figure B.1.56: Malaysia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

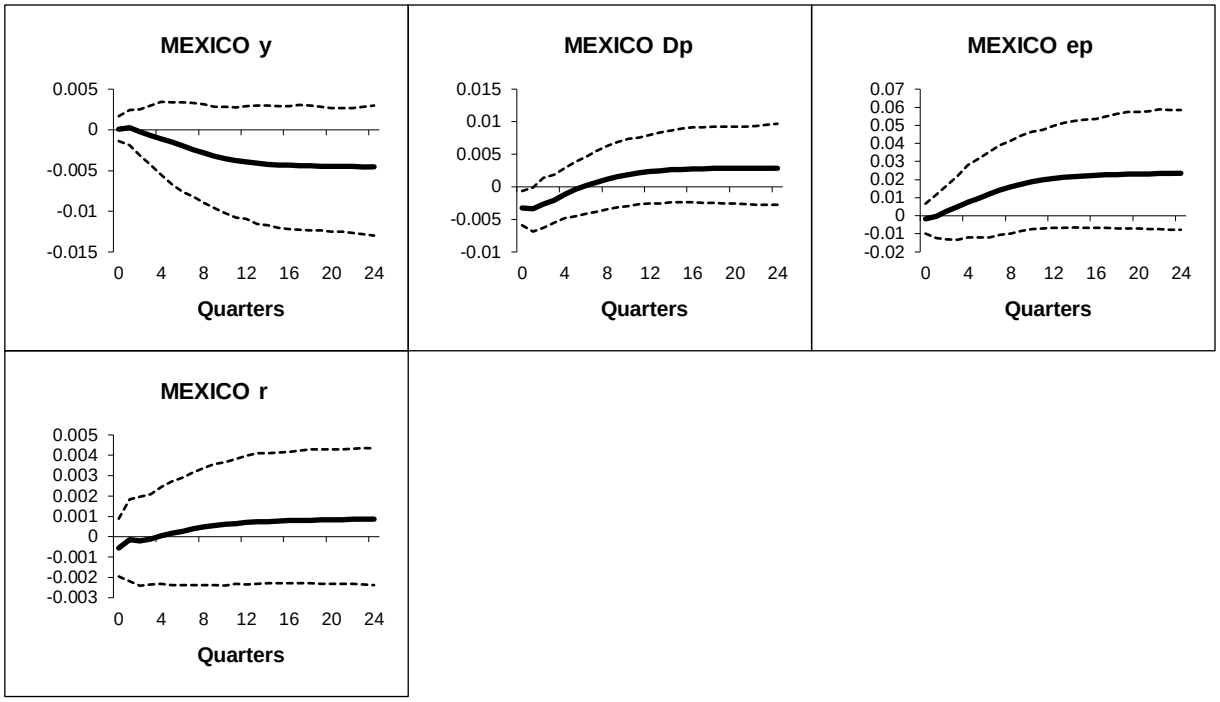

Figure B.1.57: Mexico - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

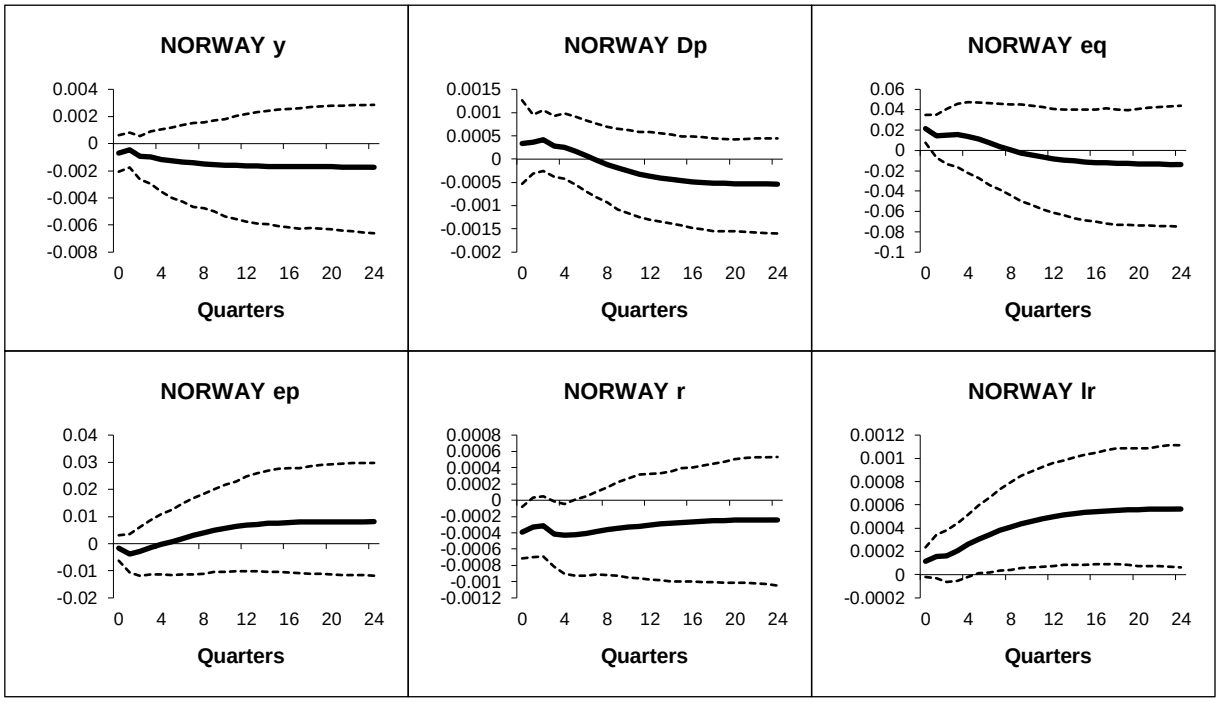

Figure B.1.58: Norway - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

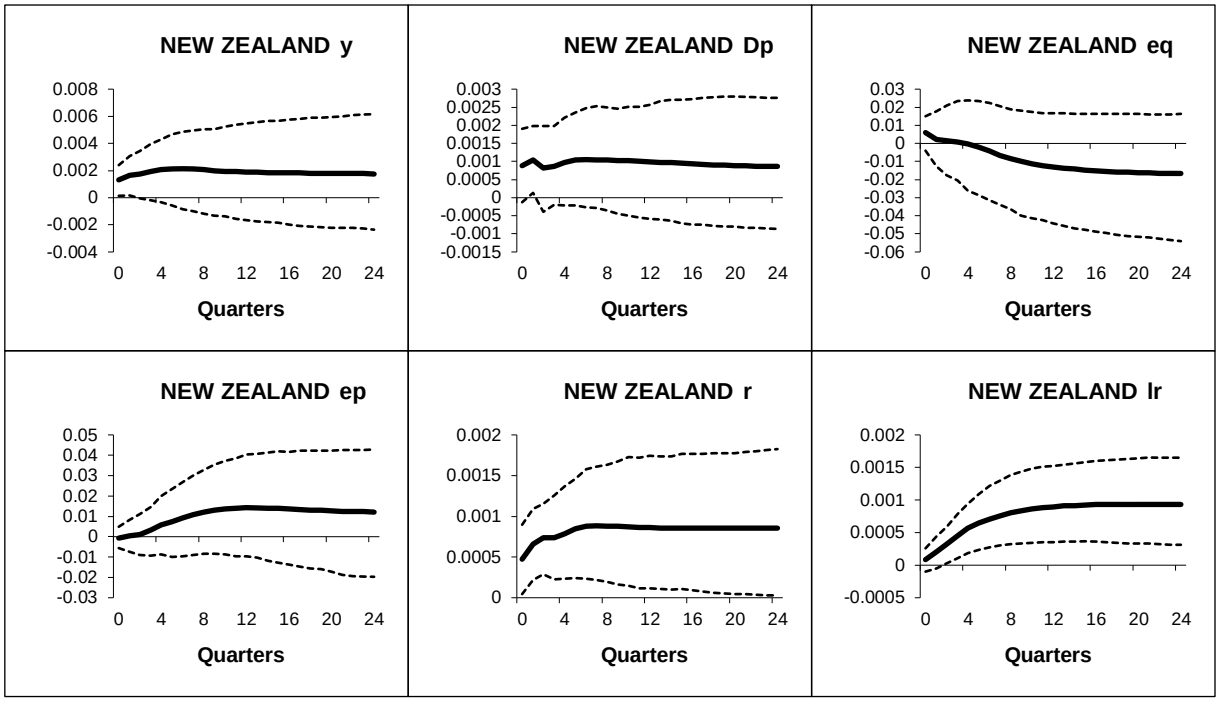

Figure B.1.59: New Zealand - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

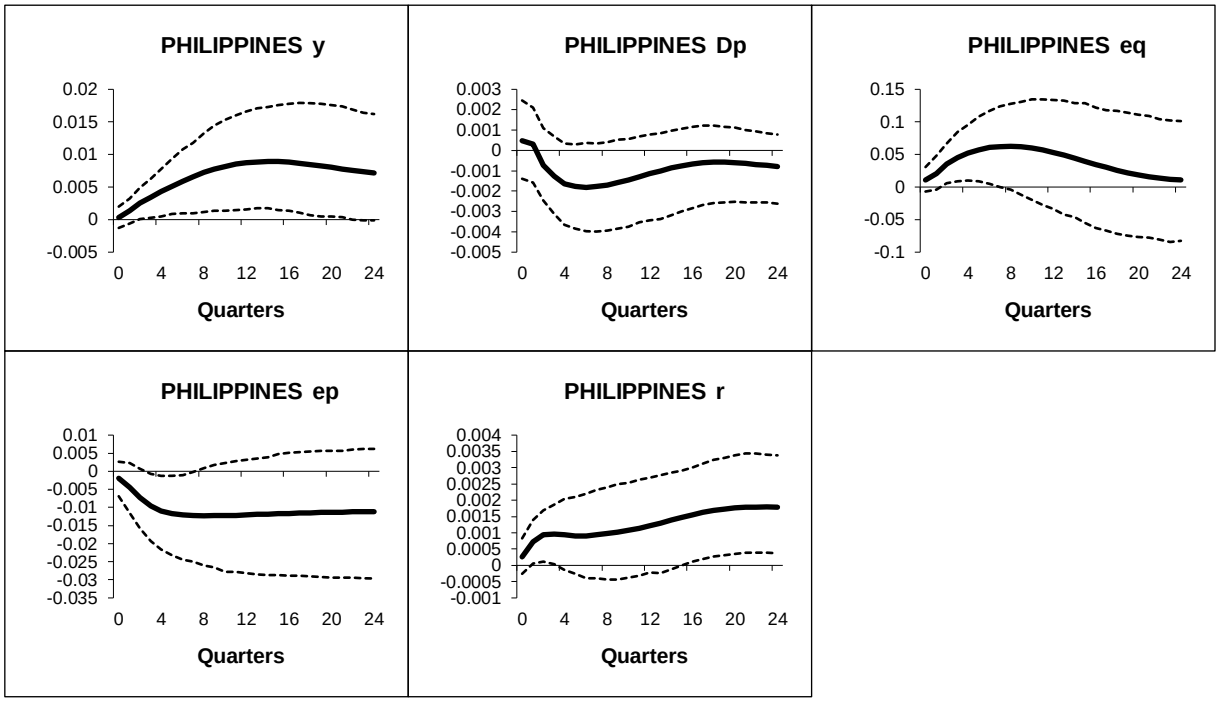

Figure B.1.60: Philippines - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

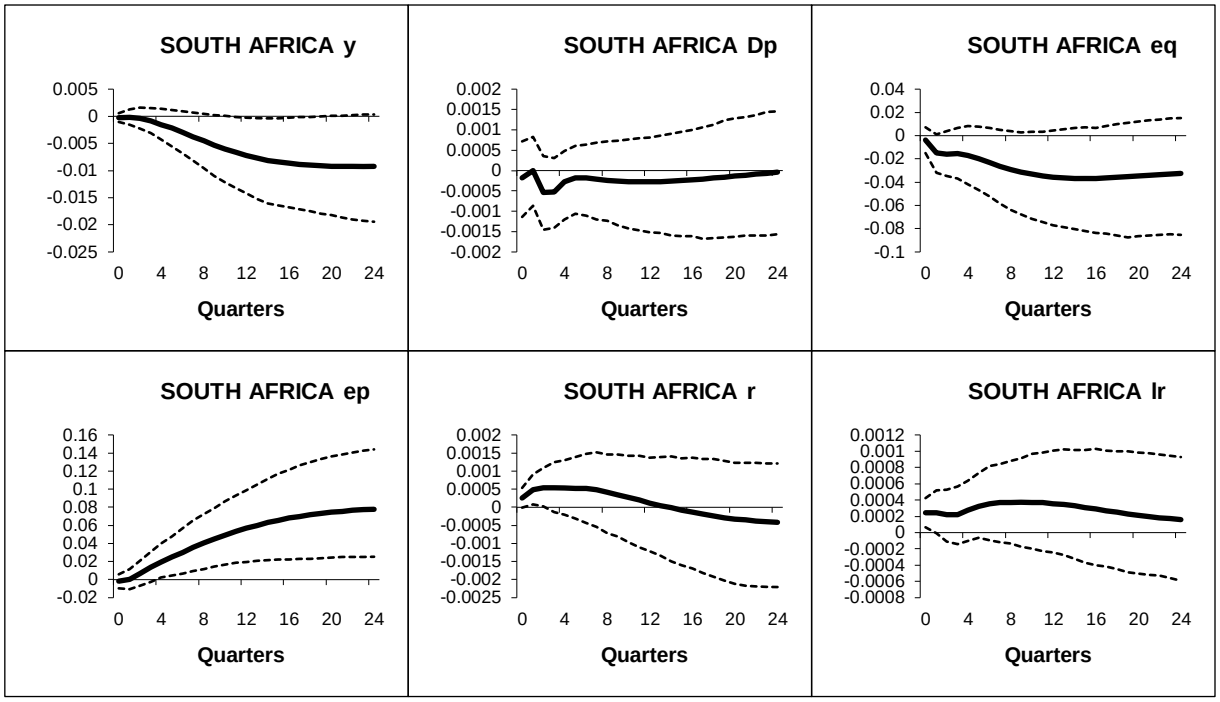

Figure B.1.61: South Africa - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

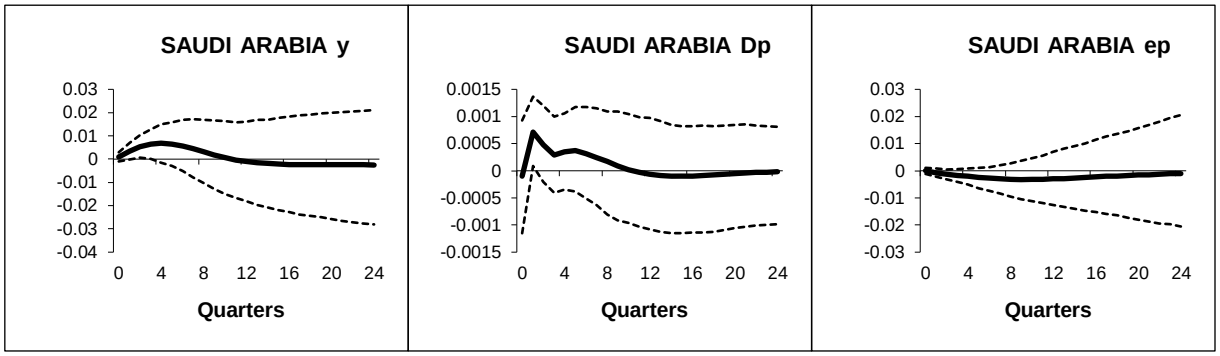

Figure B.1.62: Saudi Arabia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

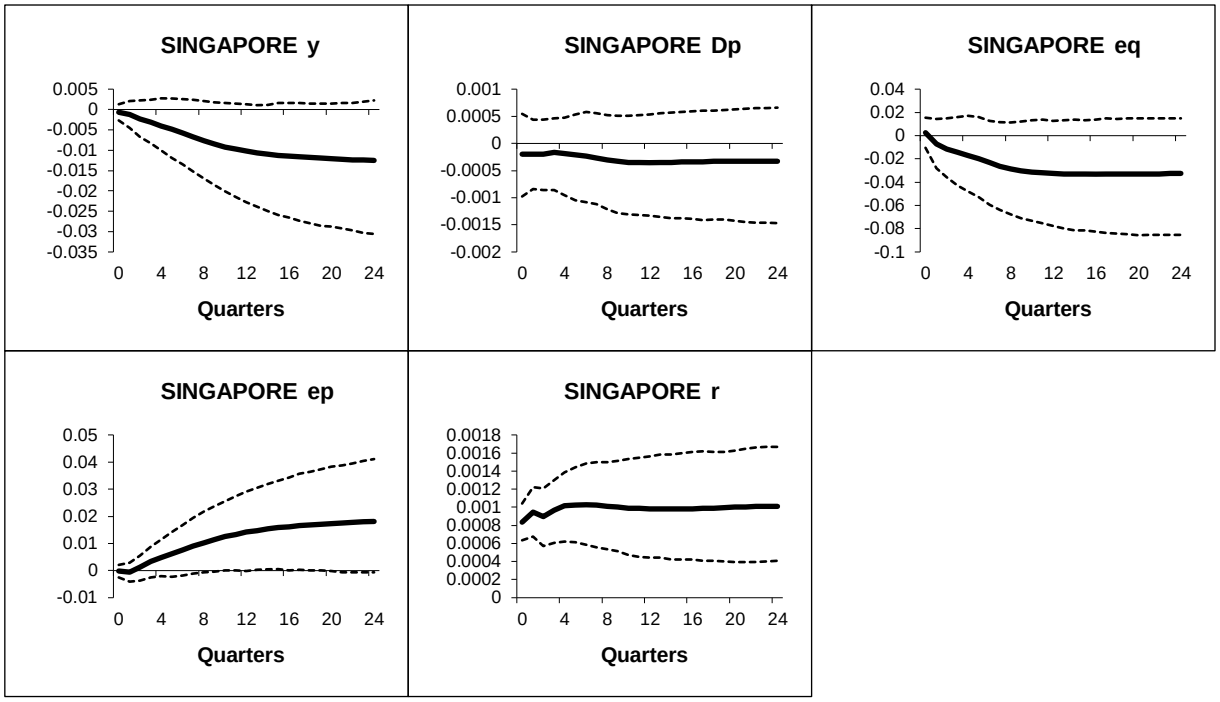

Figure B.1.63: Singapore - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

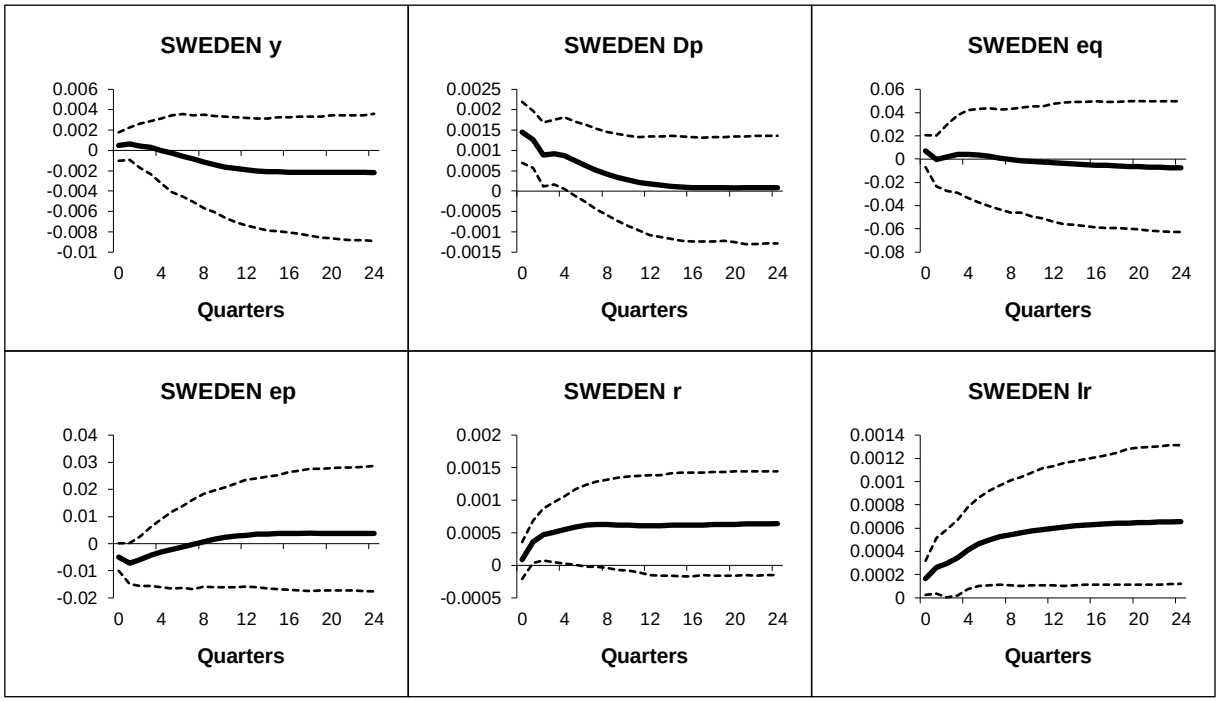

Figure B.1.64: Sweden - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

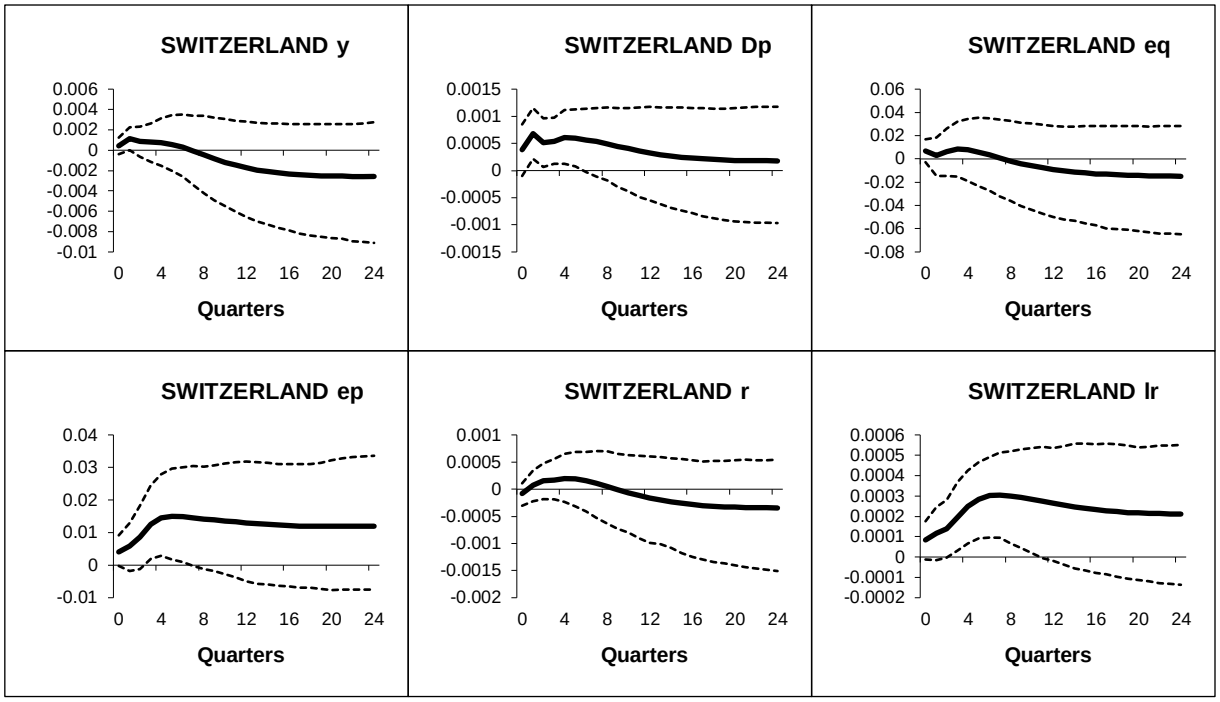

Figure B.1.65: Switzerland - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

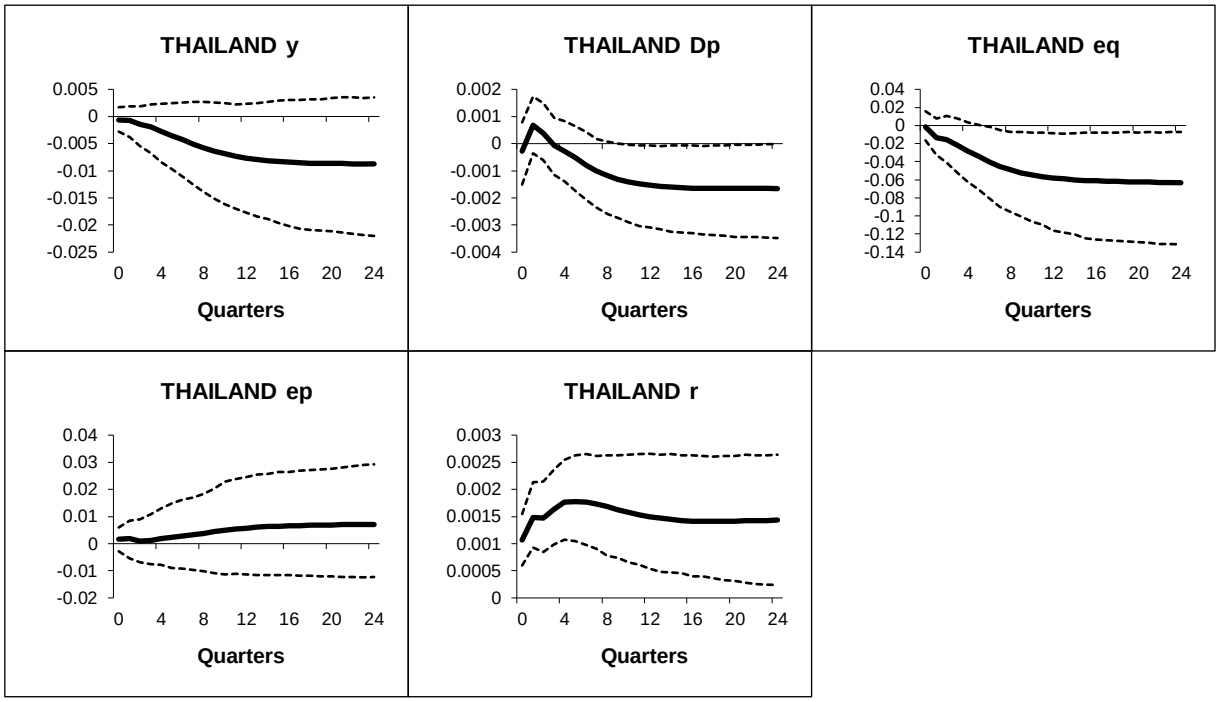

Figure B.1.66: Thailand - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

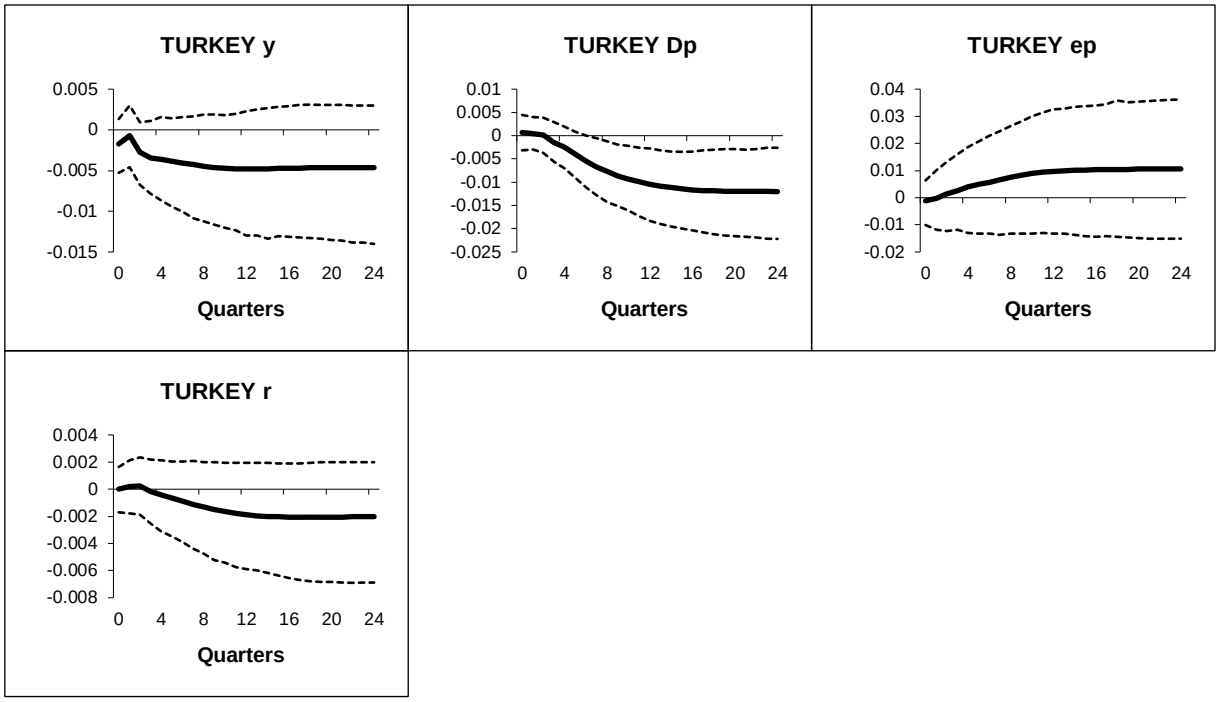

Figure B.1.67: Turkey - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

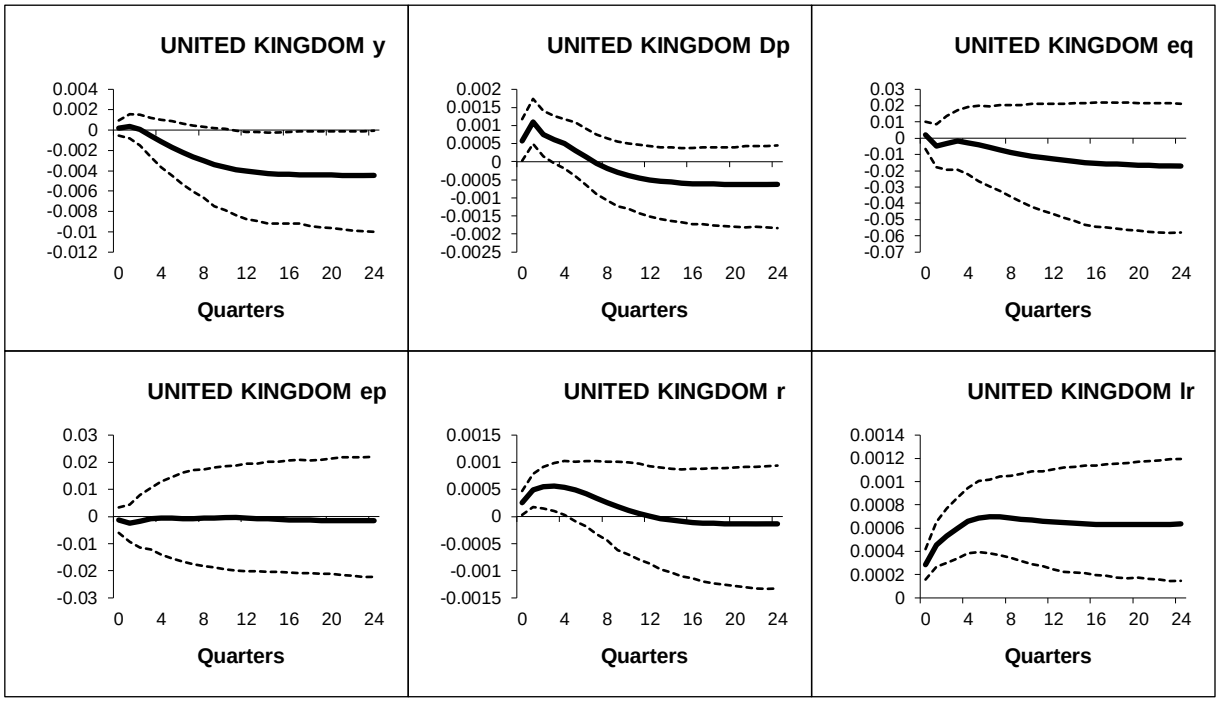

Figure B.1.68: UK - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

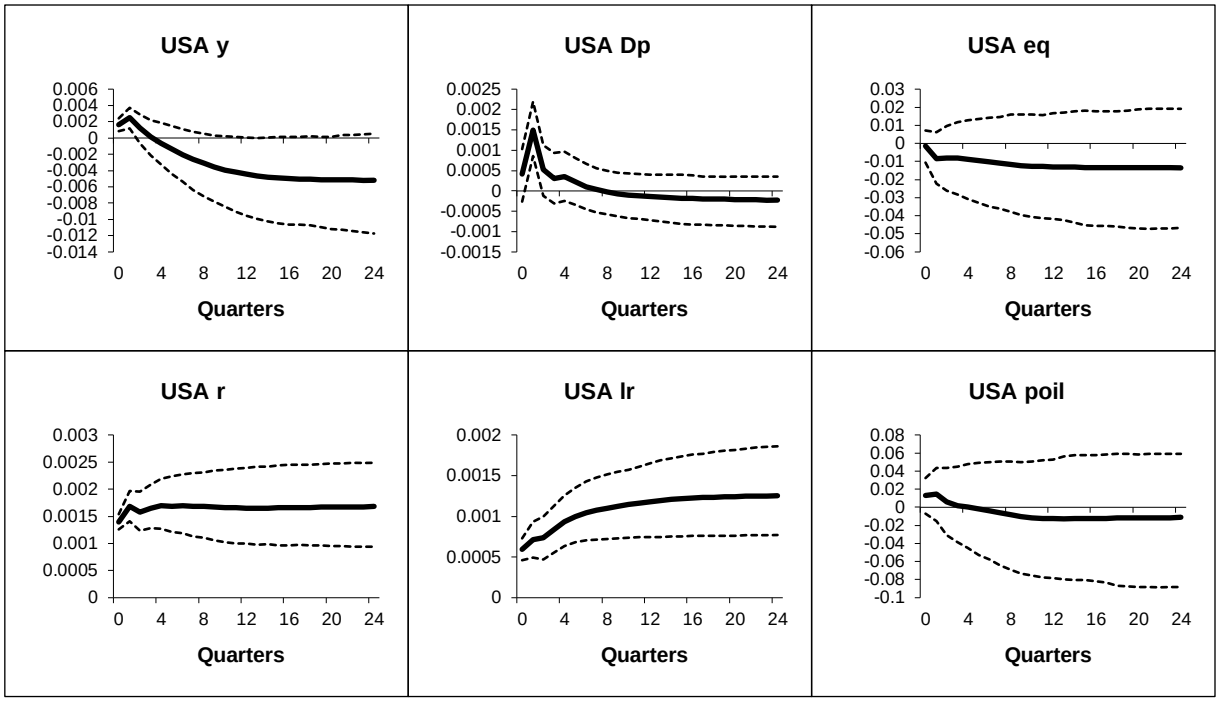

Figure B.1.69: USA - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). CV-CX weighted model IRFs.

APPENDIX B.2

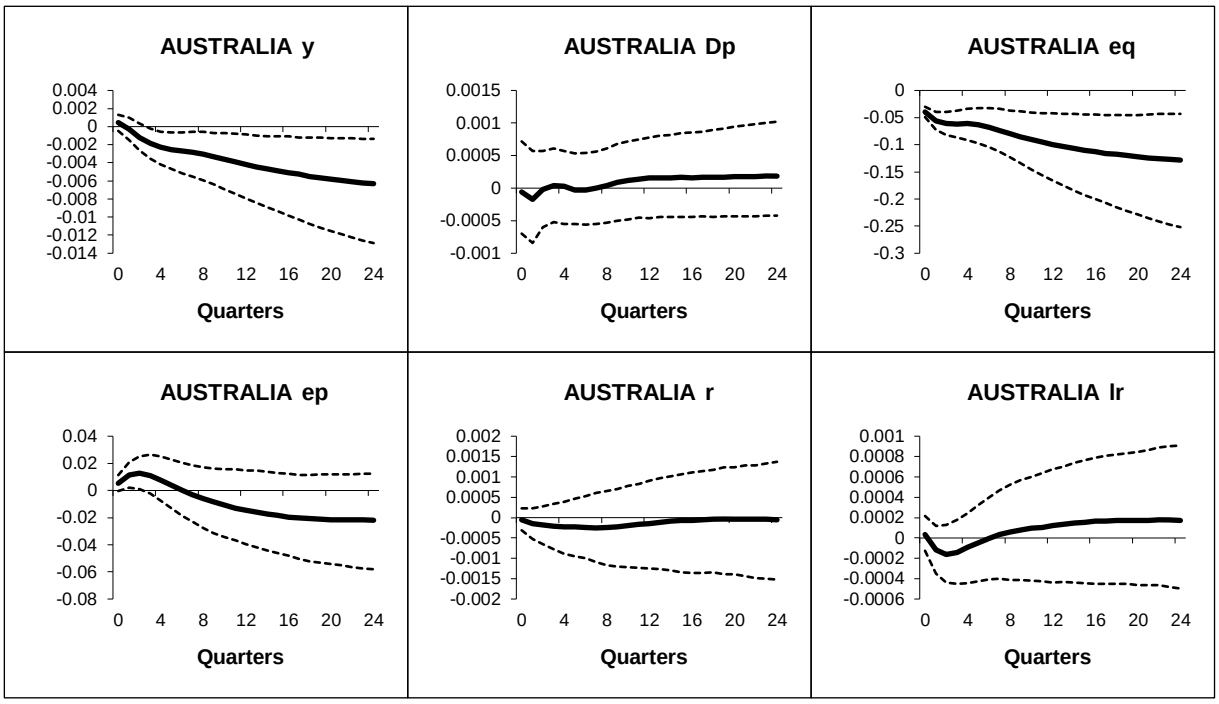

Figure B.2.1: Australia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

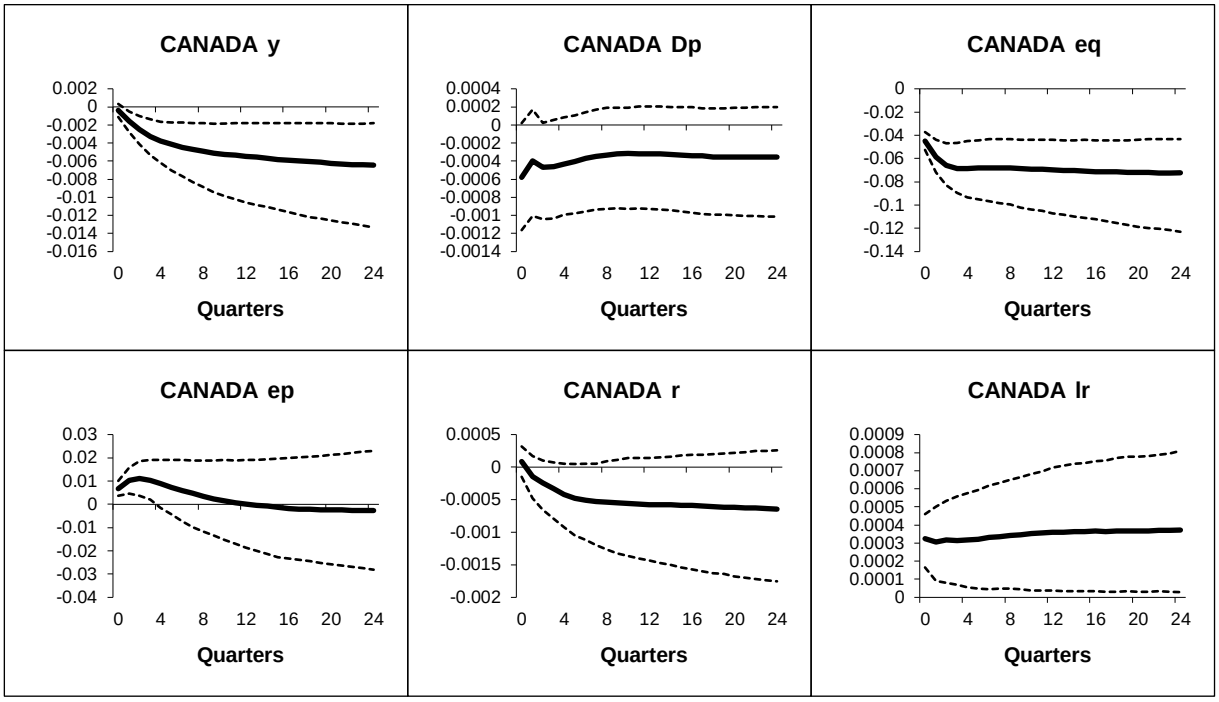

Figure B.2.2: Canada - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

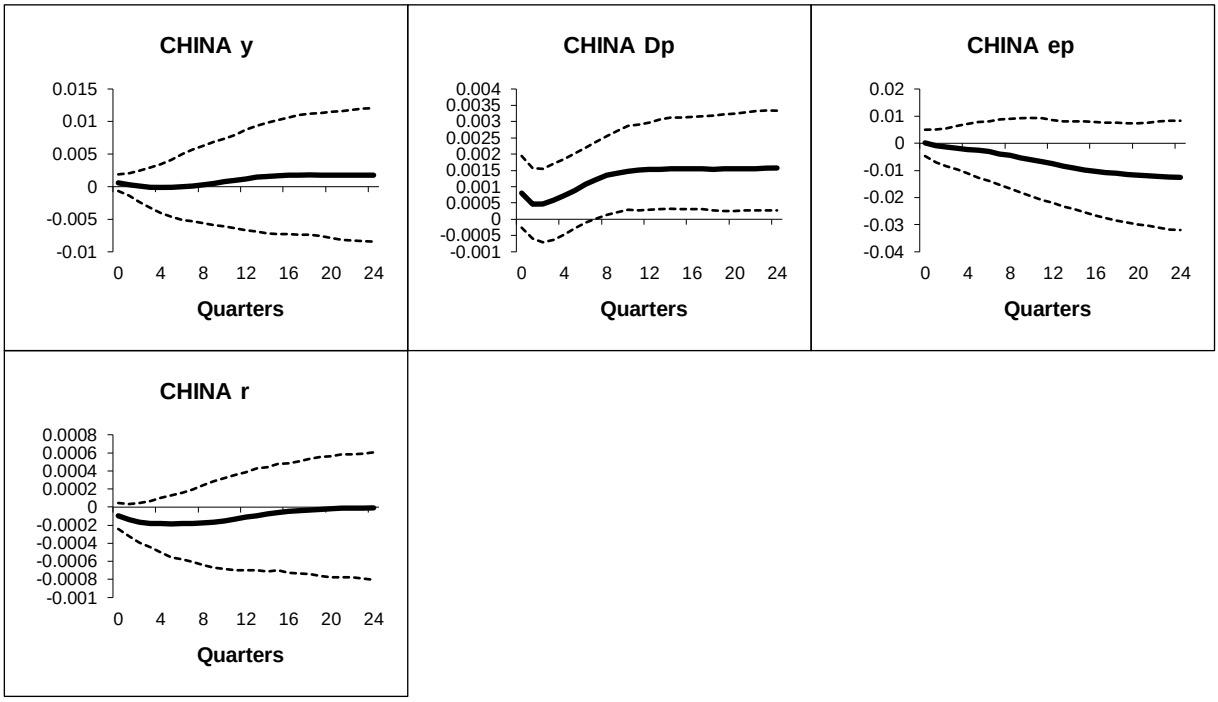

Figure B.2.3: China - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

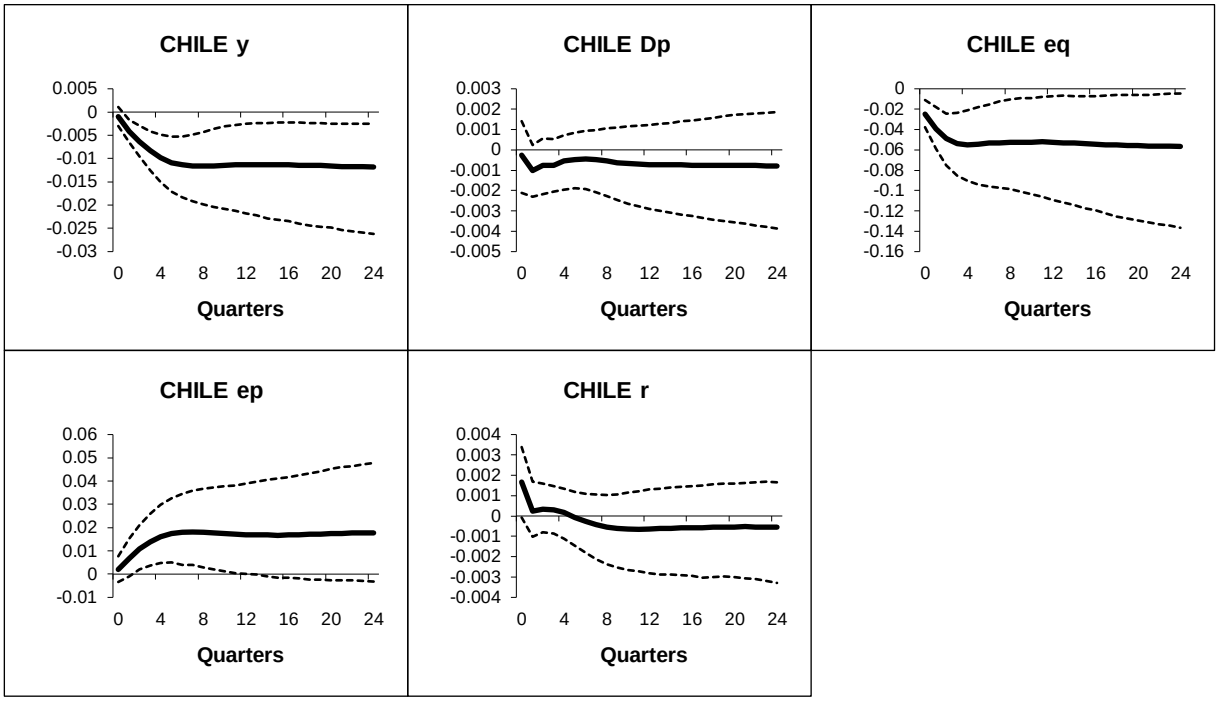

Figure B.2.4: Chile - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

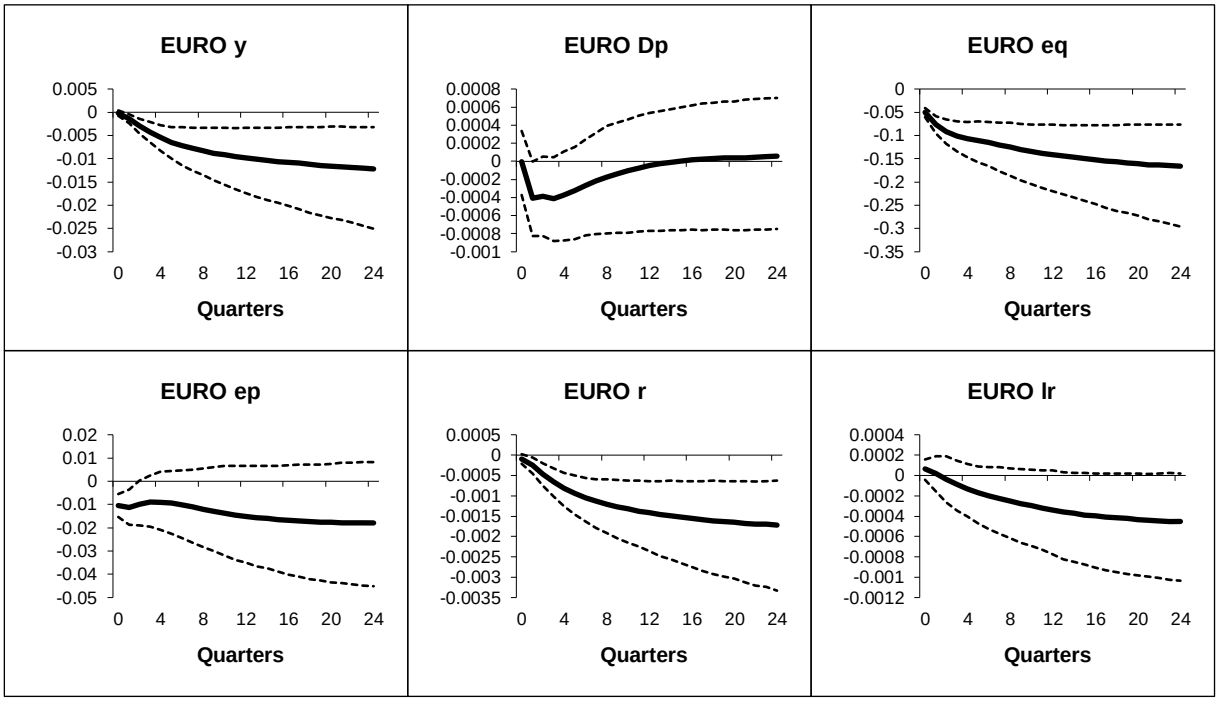

Figure B.2.5: Euro - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

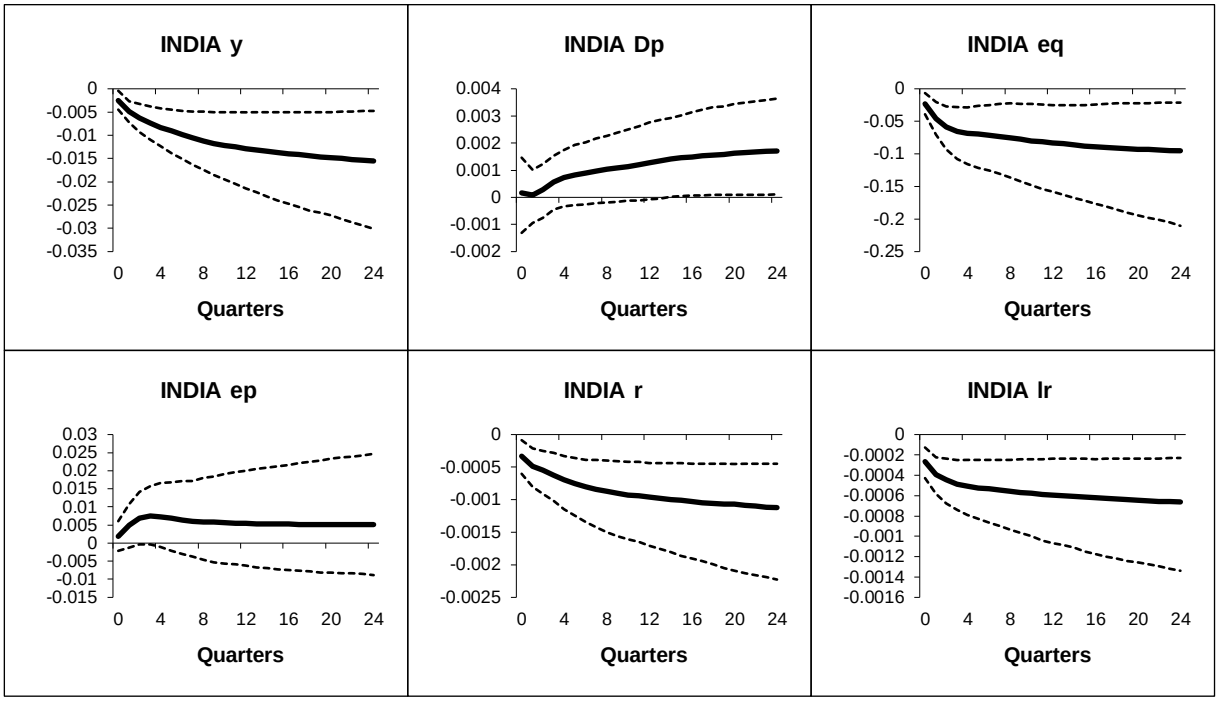

Figure B.2.6: India - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

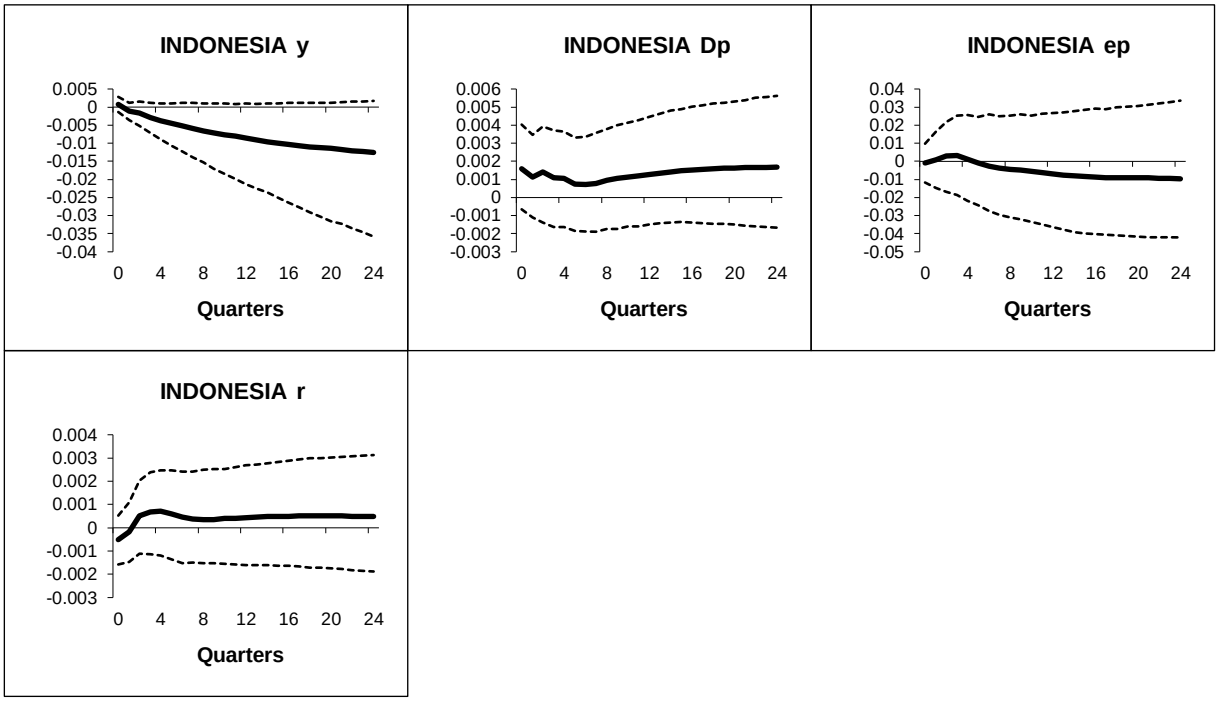

Figure B.2.7: Indonesia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

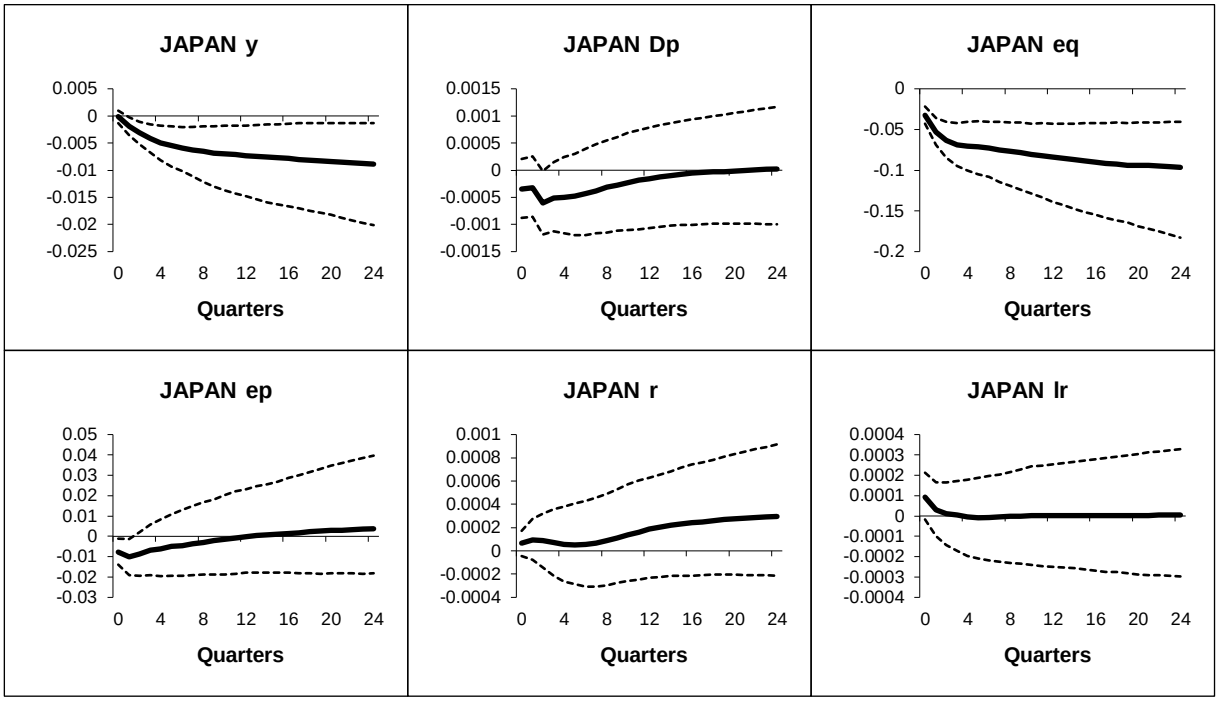

Figure B.2.8: Japan - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

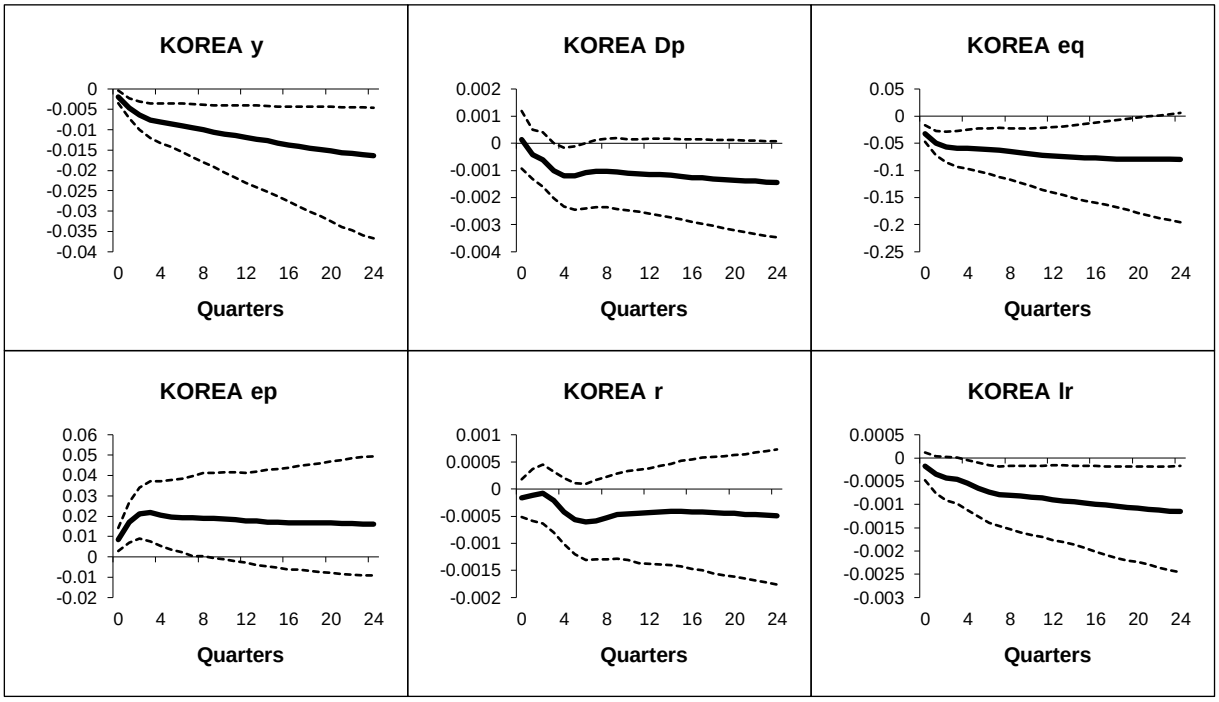

Figure B.2.9: Korea - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

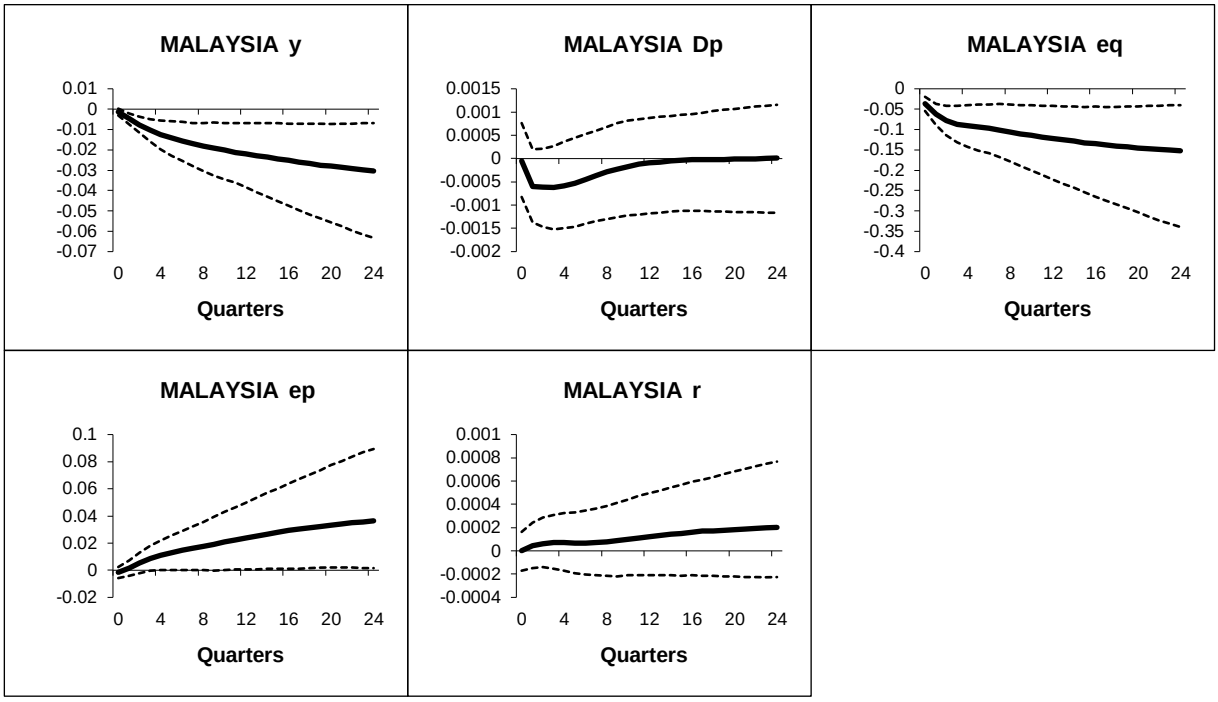

Figure B.2.10: Malaysia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

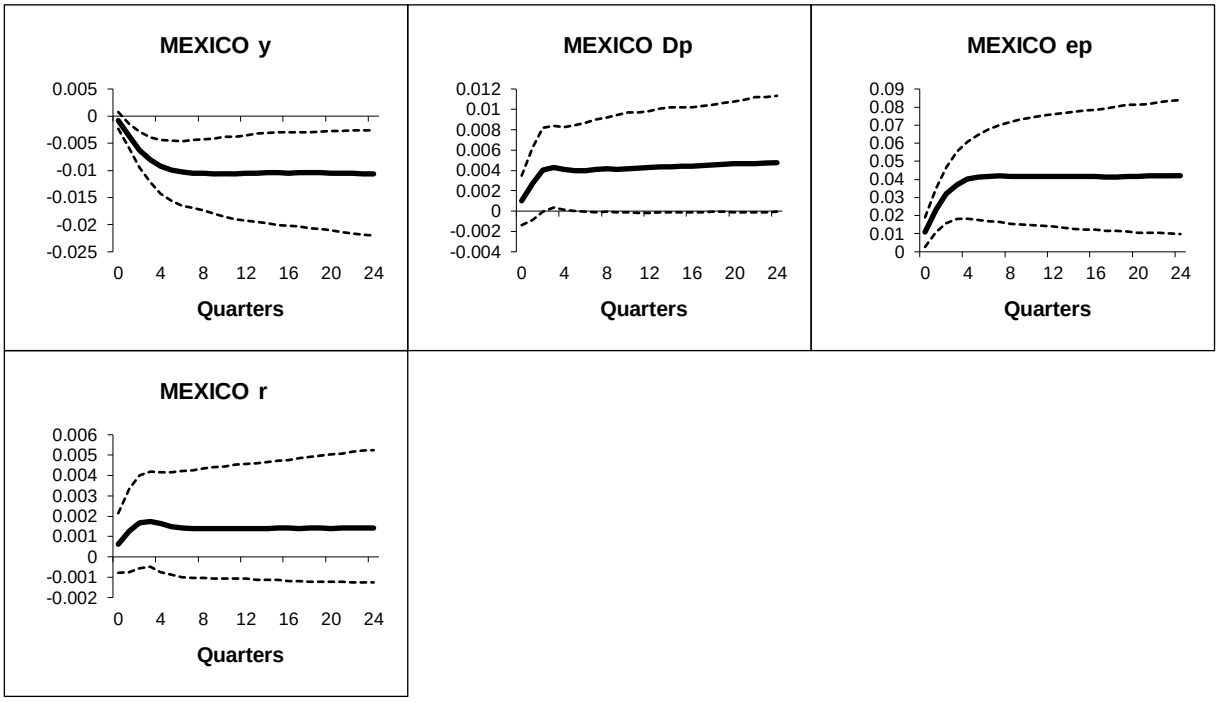

Figure B.2.11: Mexico - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

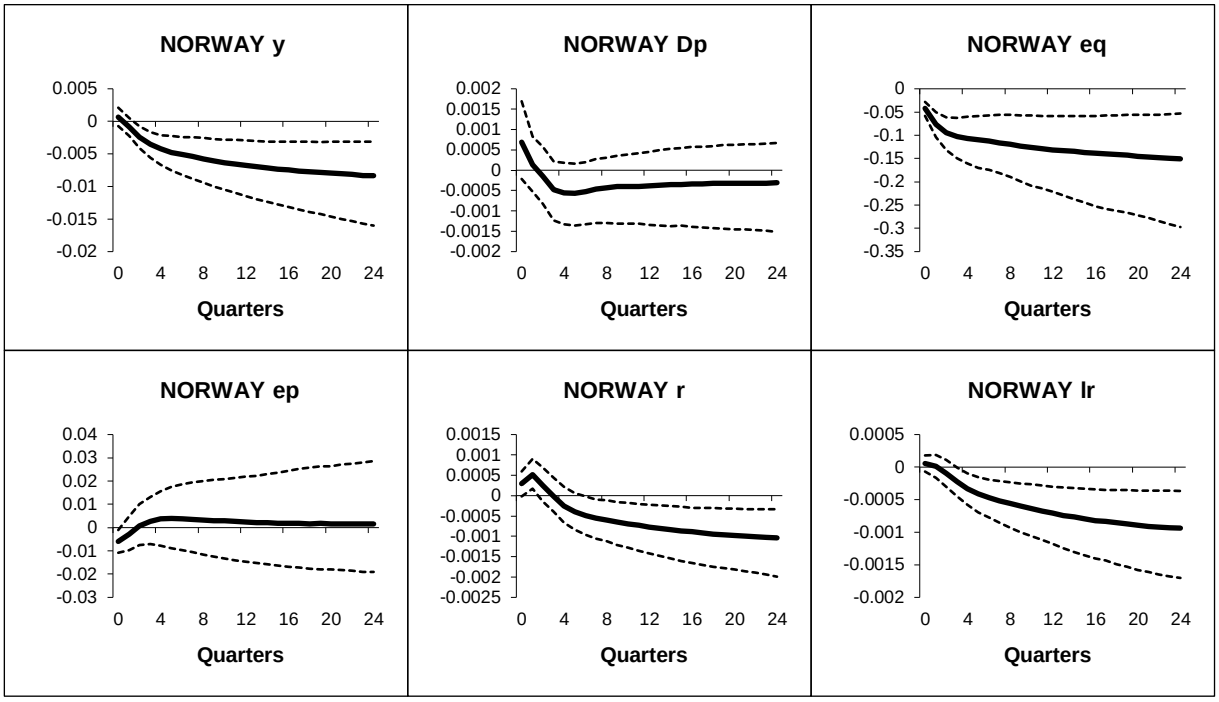

Figure B.2.12: Norway - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

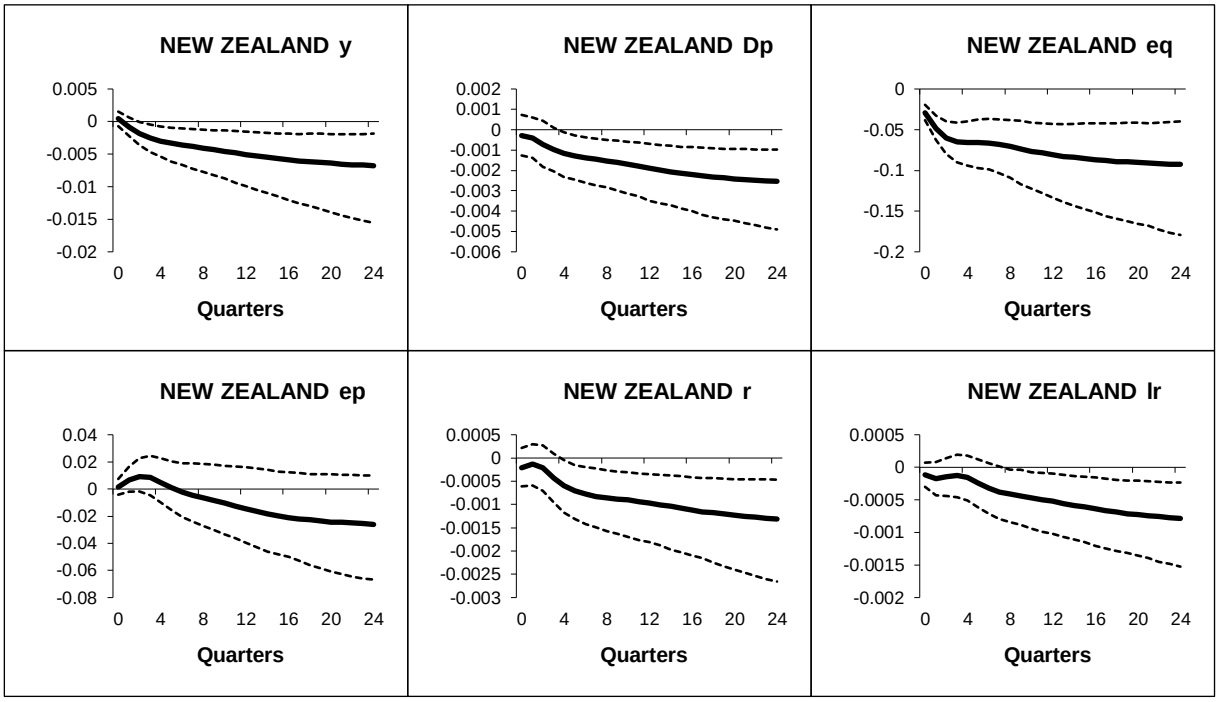

Figure B.2.13: New Zealand - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

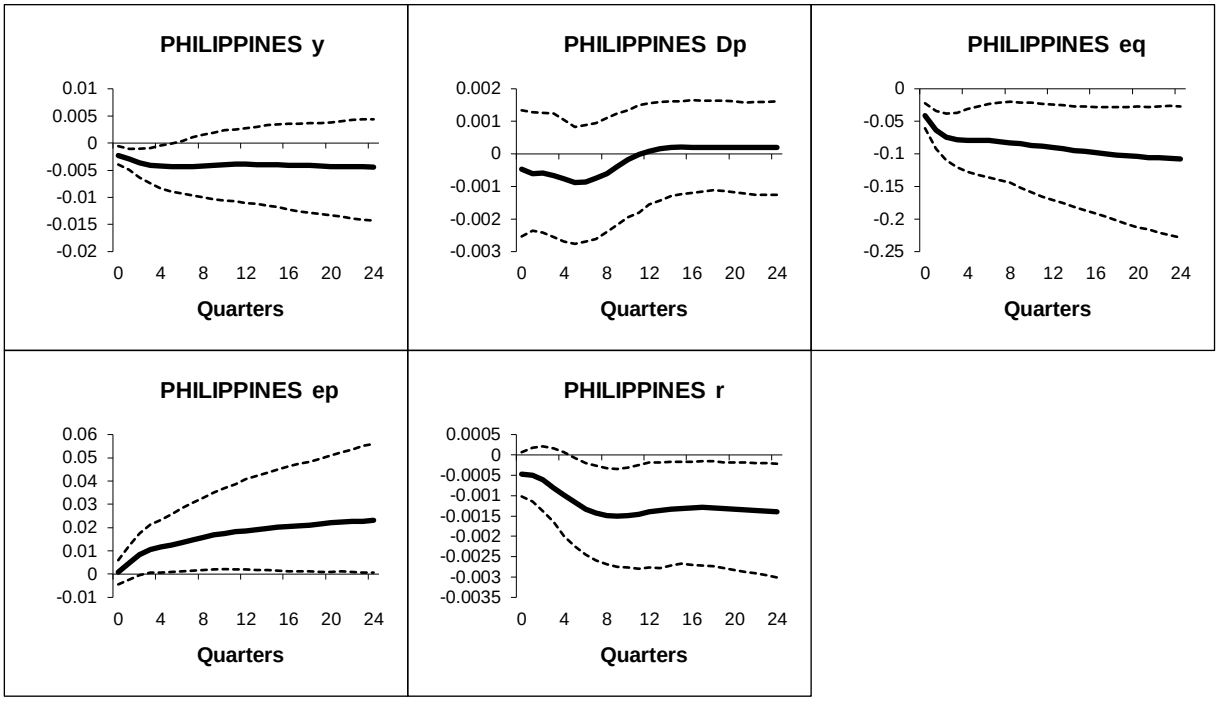

Figure B.2.14: Philippines - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

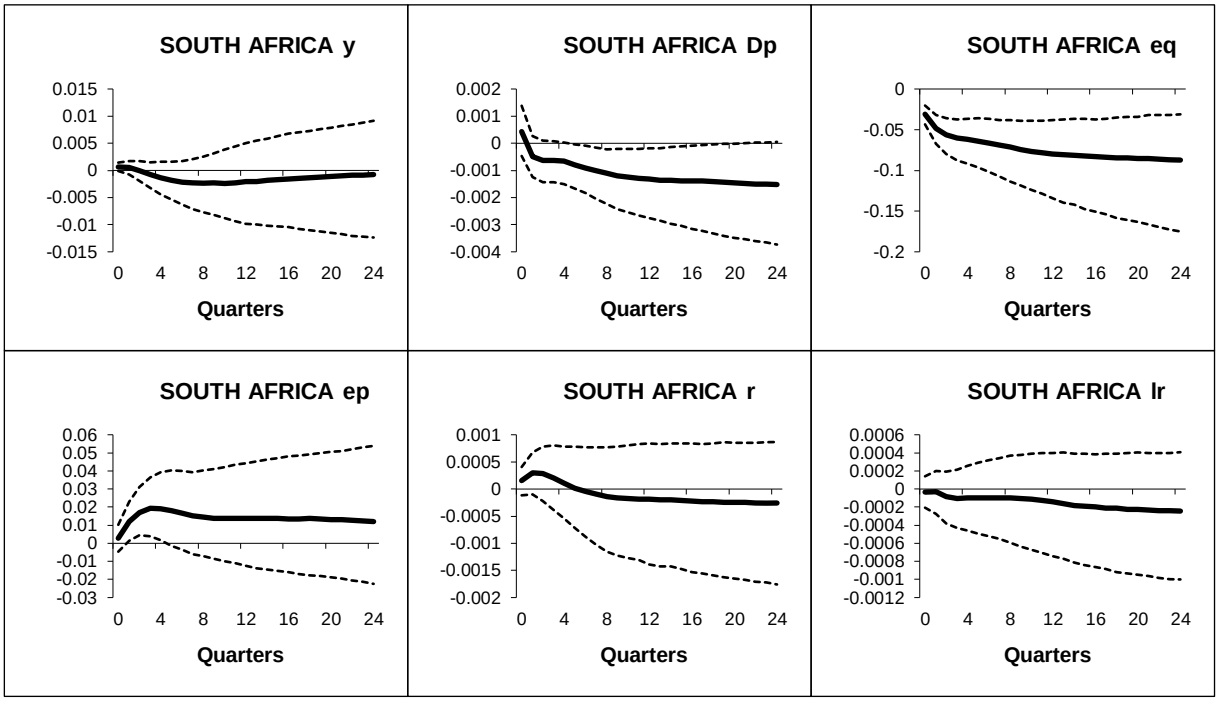

Figure B.2.15: South Africa - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

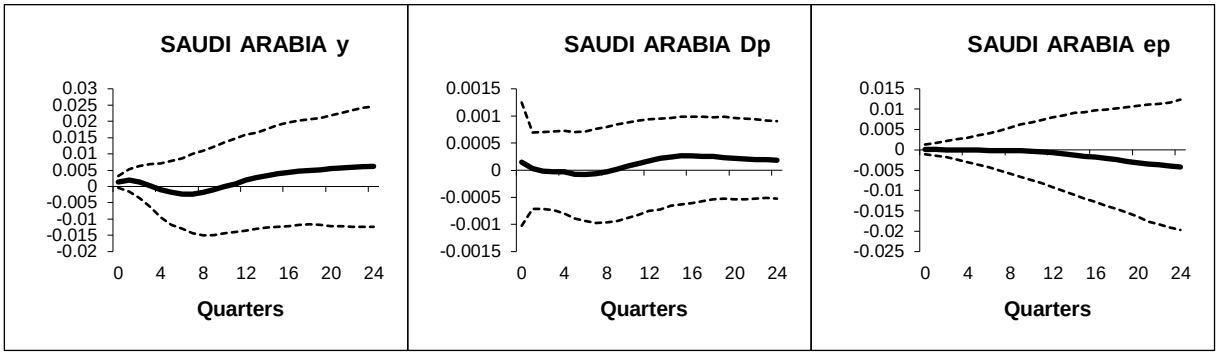

Figure B.2.16: Saudi Arabia - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

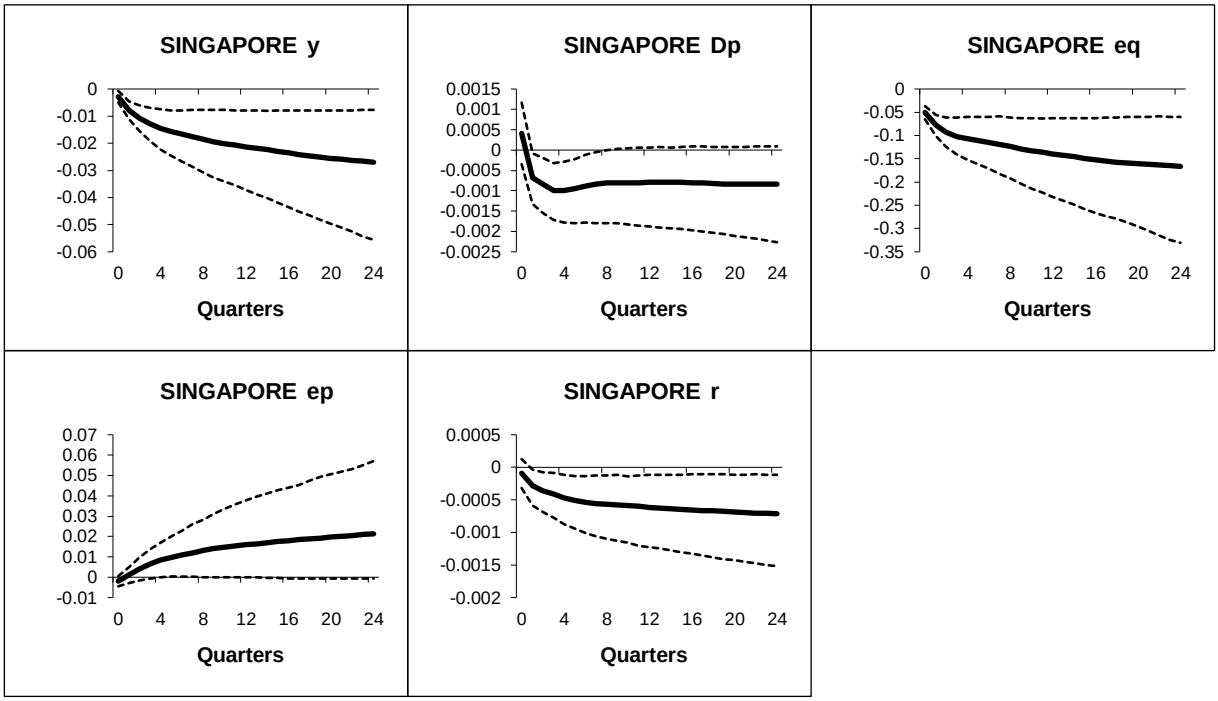

Figure B.2.17: Singapore - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

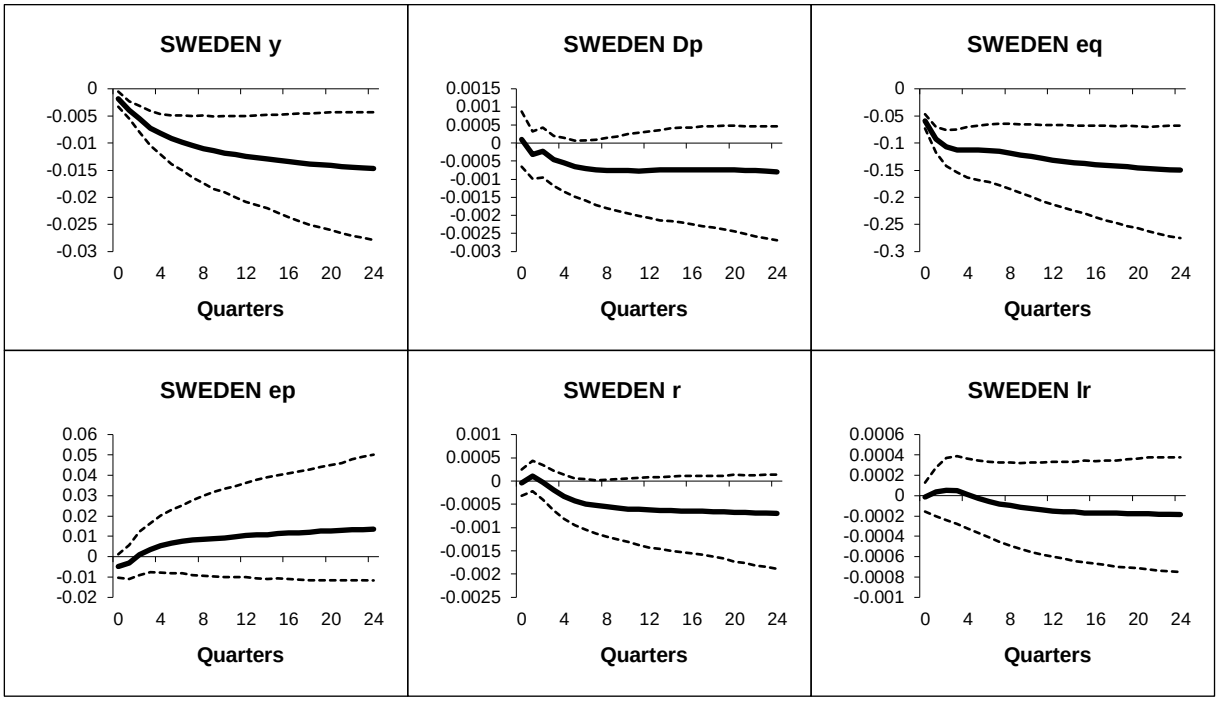

Figure B.2.18: Sweden - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

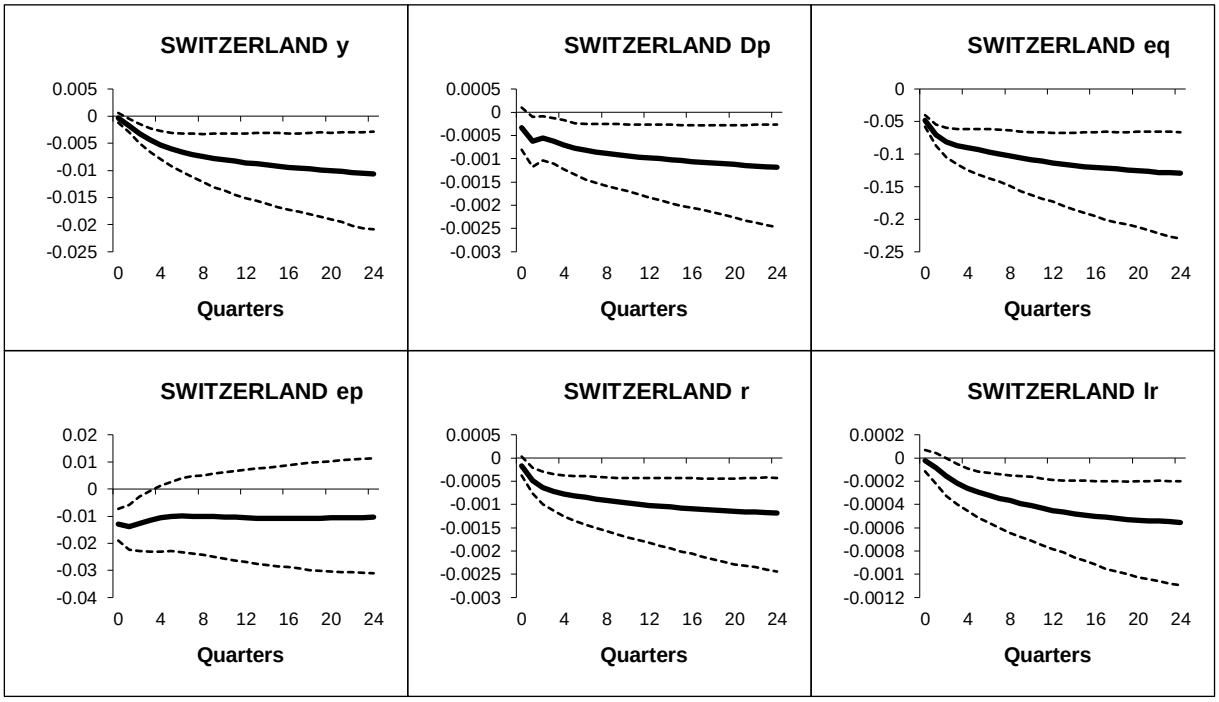

Figure B.2.19: Switzerland - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

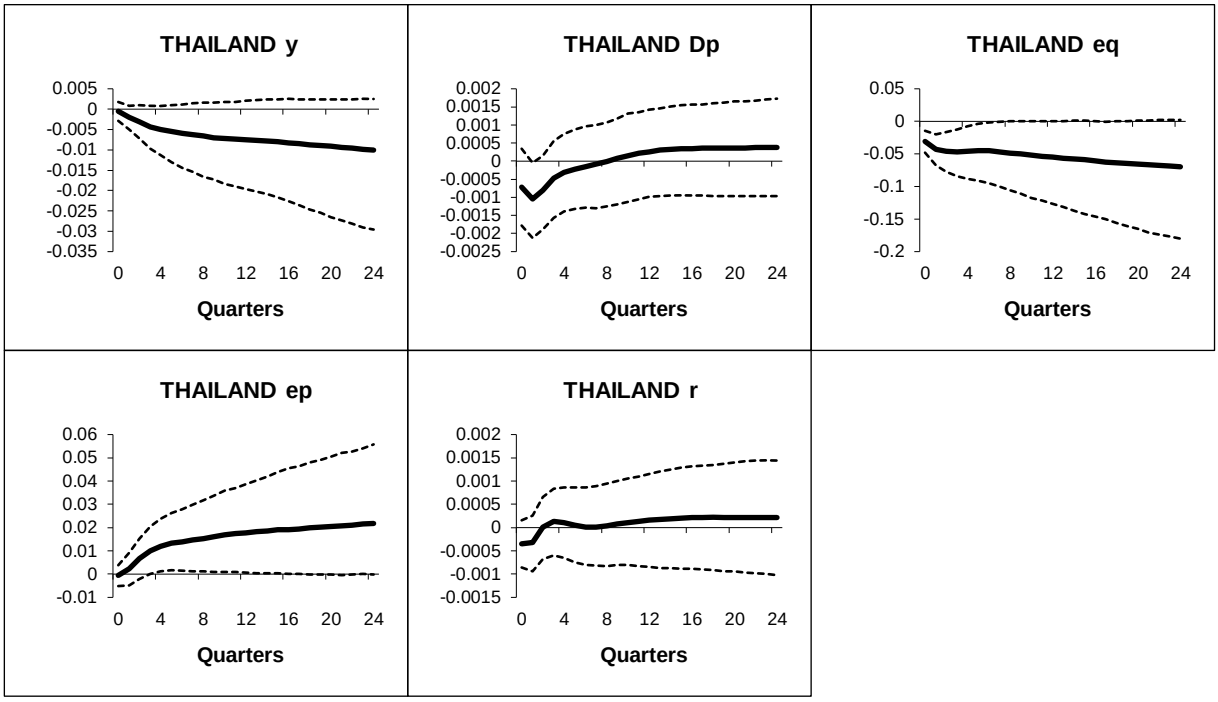

Figure B.2.20: Thailand - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

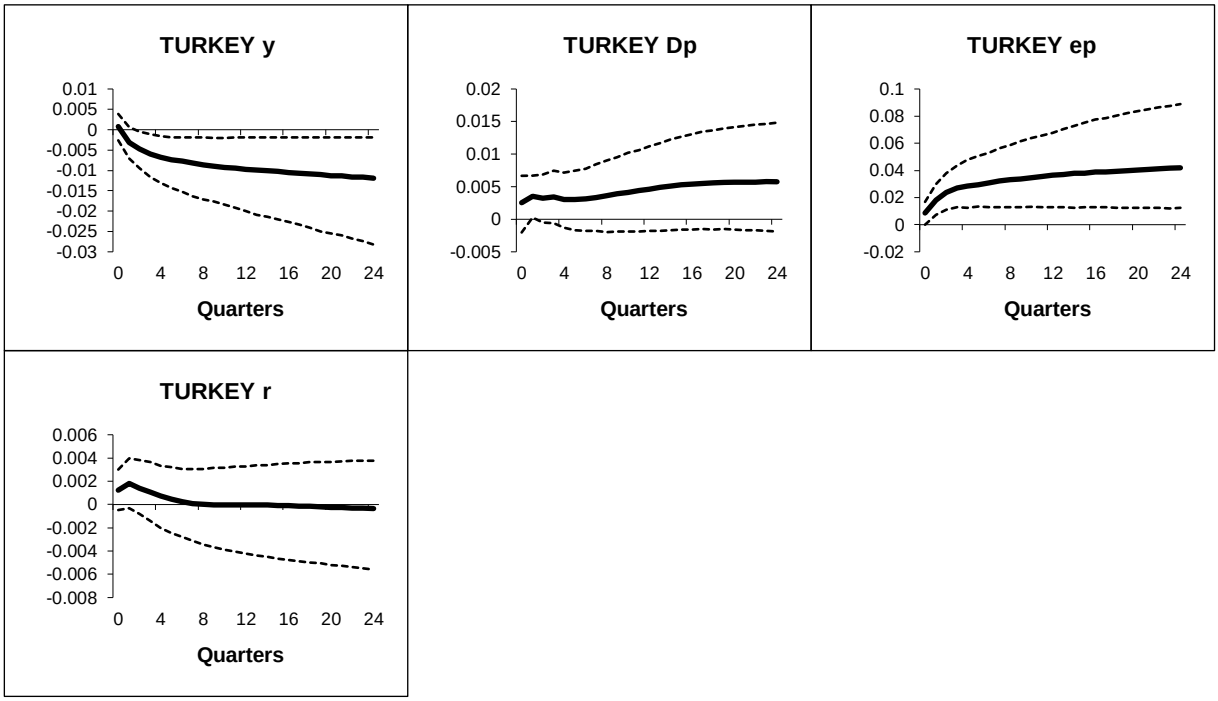

Figure B.2.21: Turkey - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

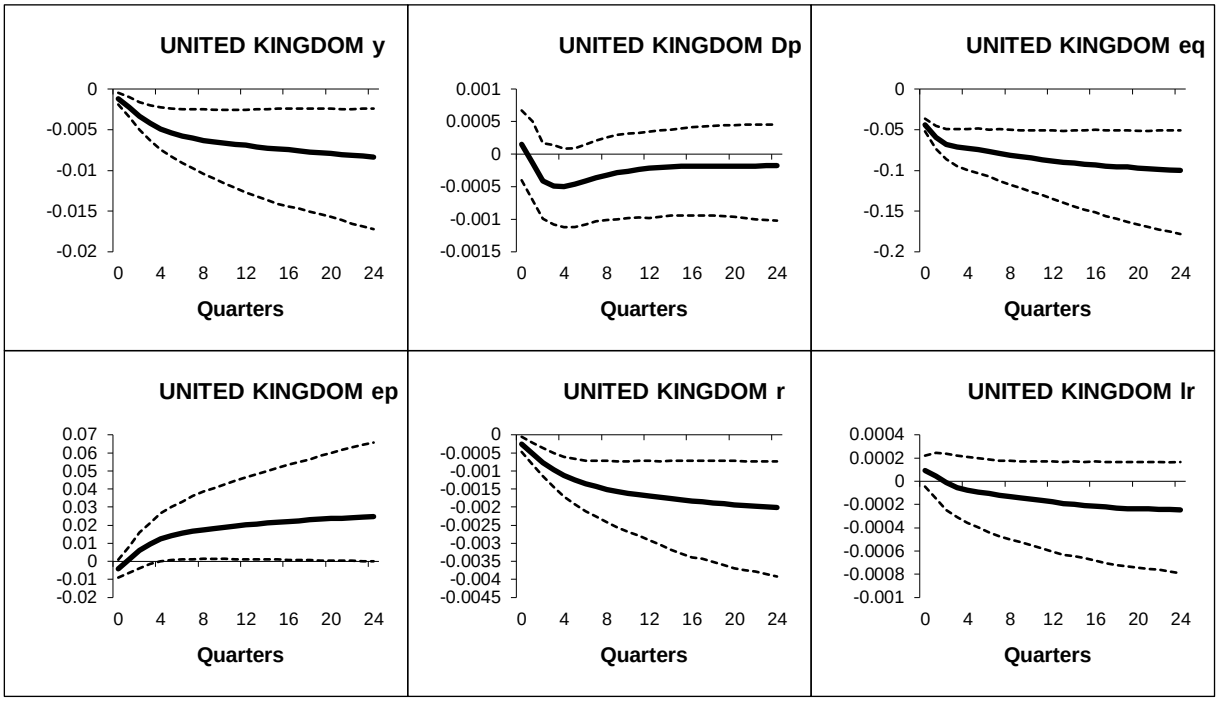

Figure B.2.22: UK - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

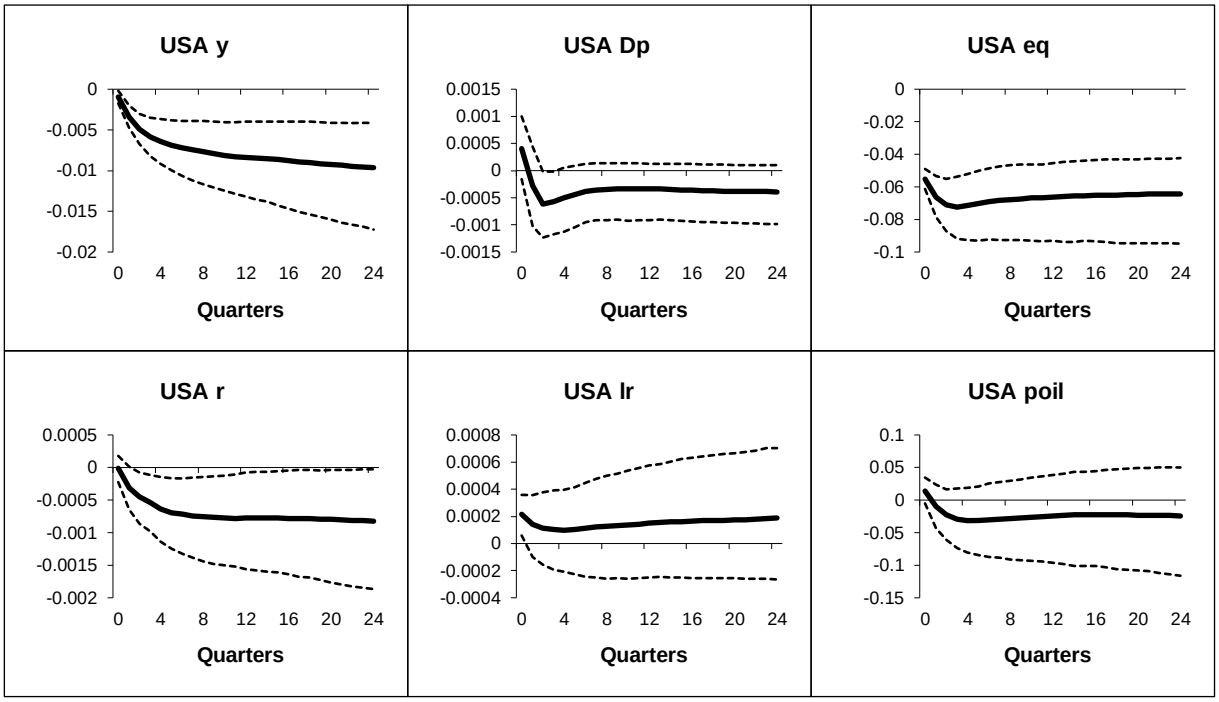

Figure B.2.23: USA - Generalized impulse responses of a negative (1 s.e.) shock to US equity prices (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

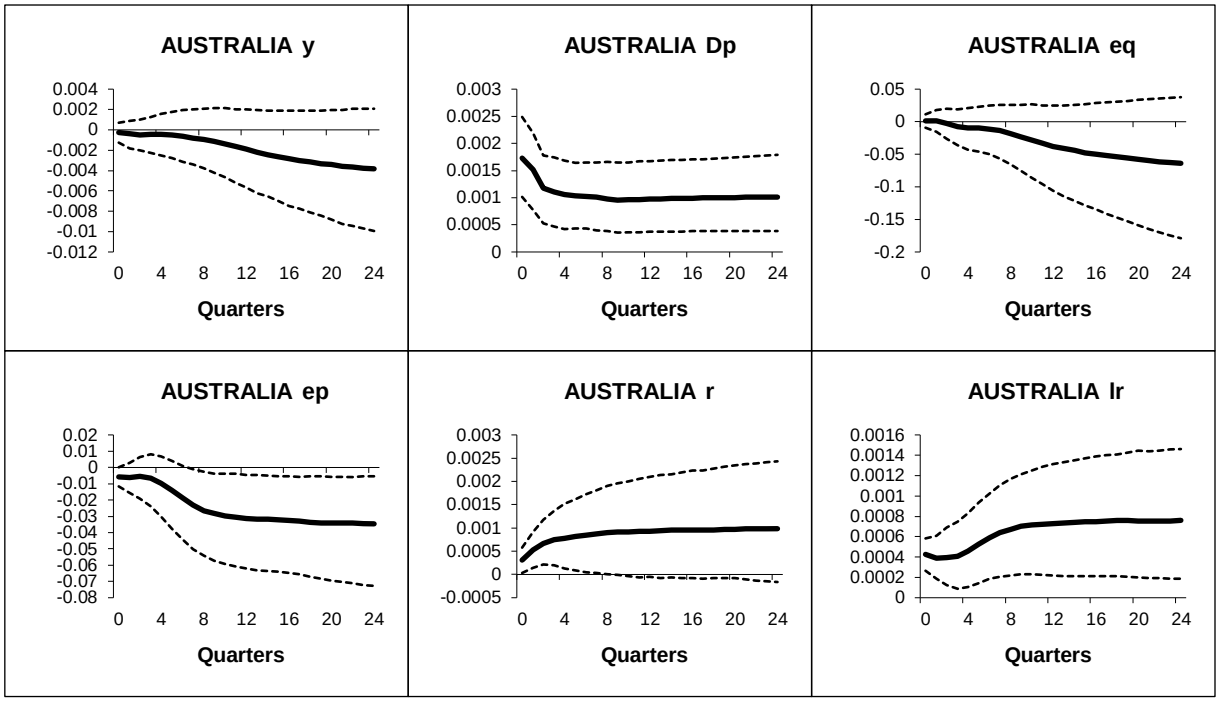

Figure B.2.24: Australia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

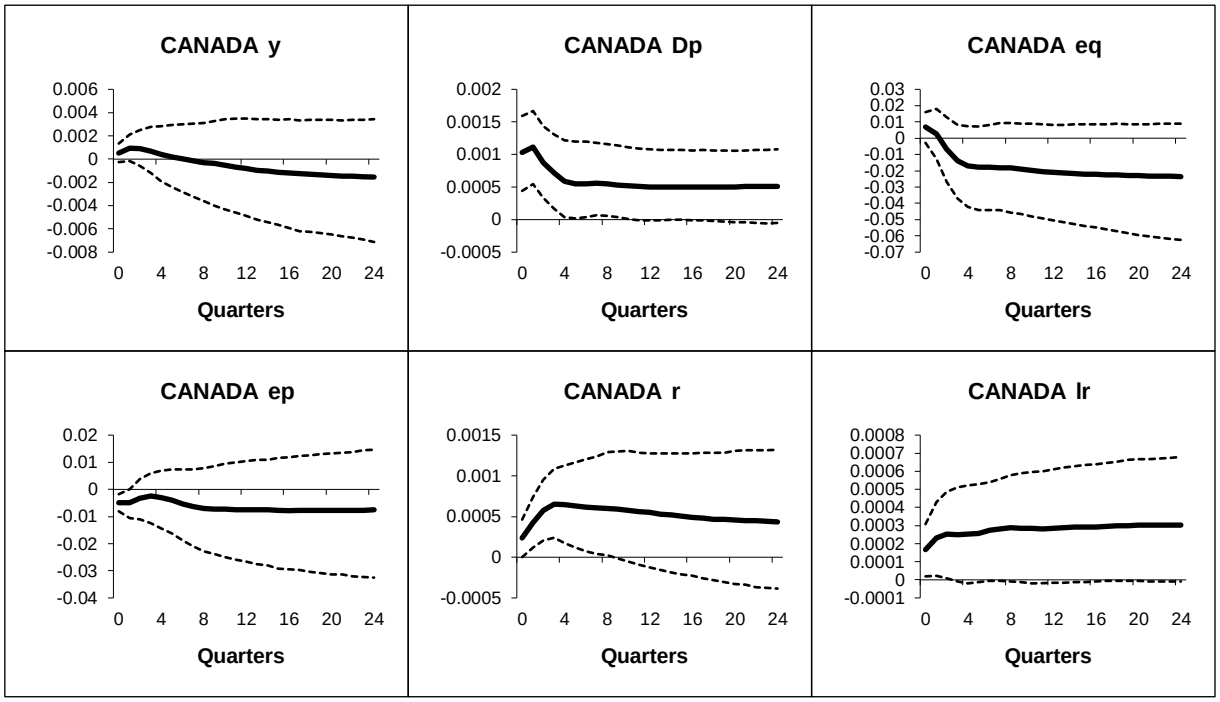

Figure B.2.25: Canada - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

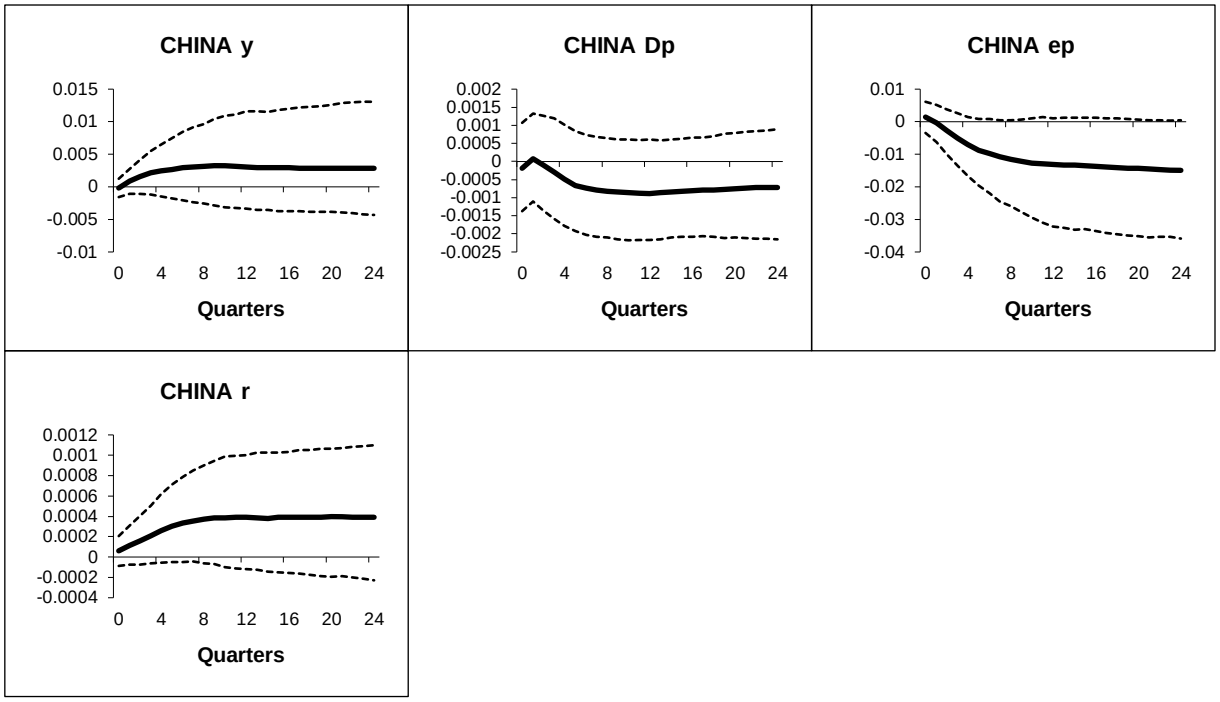

Figure B.2.26: China - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

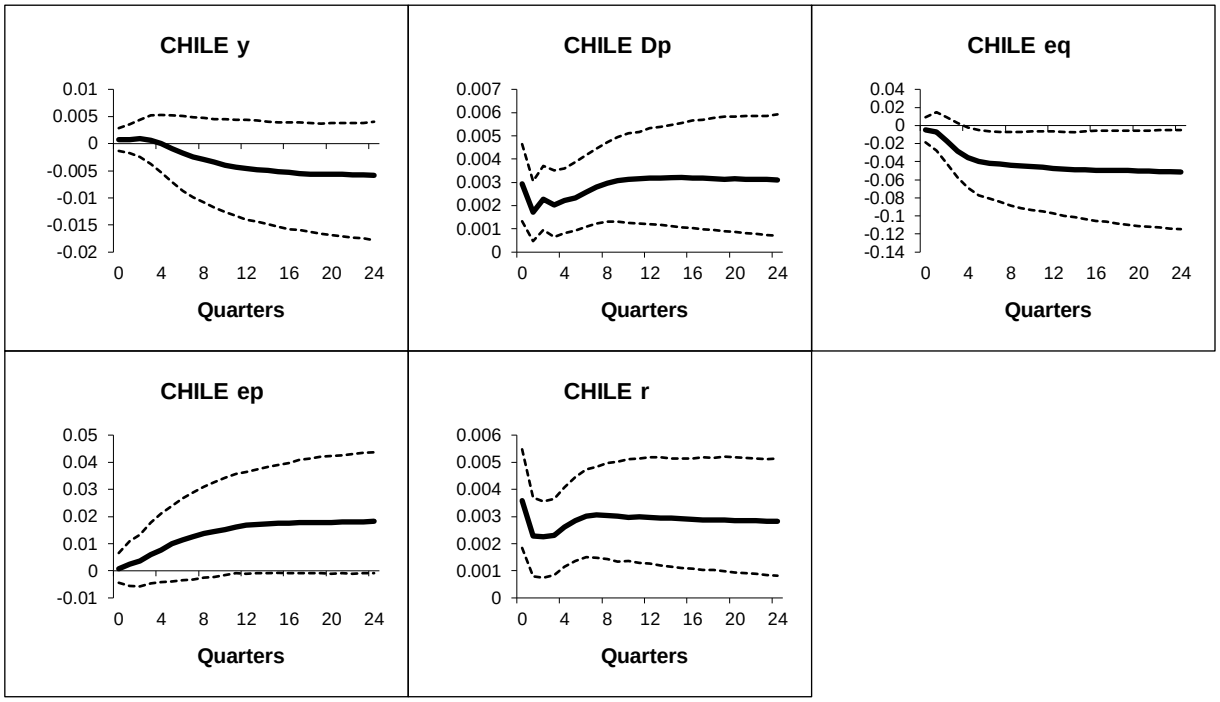

Figure B.2.27: Chile - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

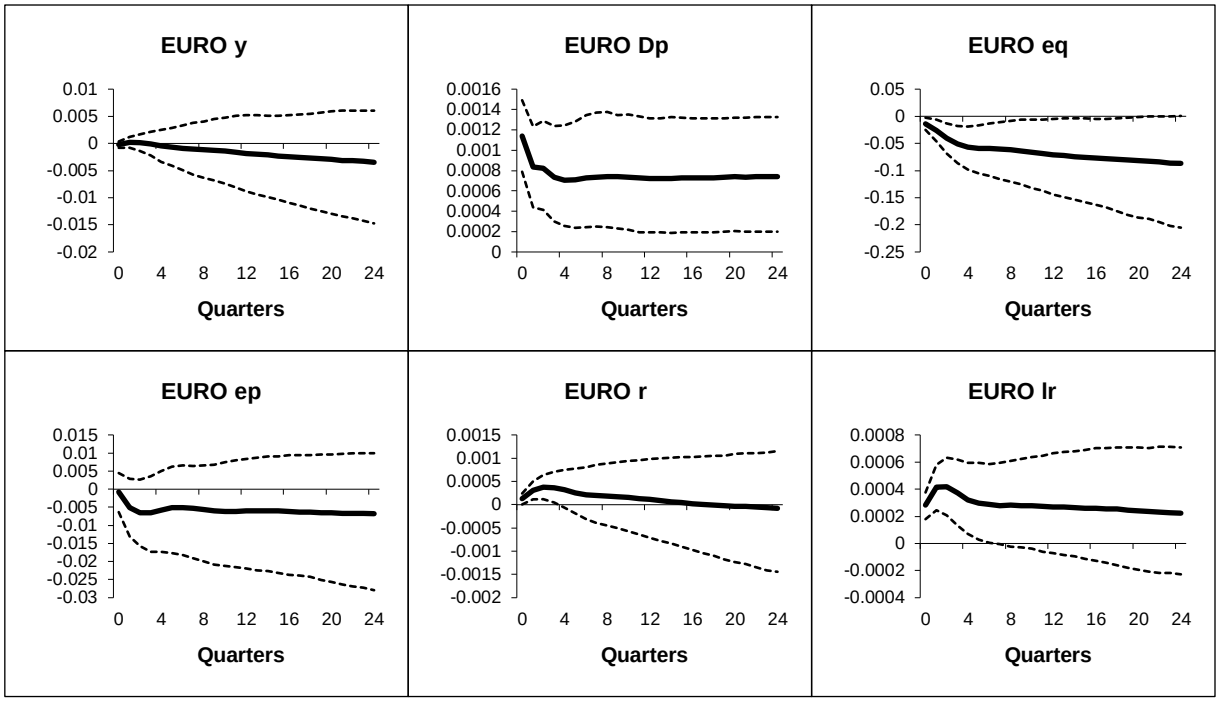

Figure B.2.28: Euro - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

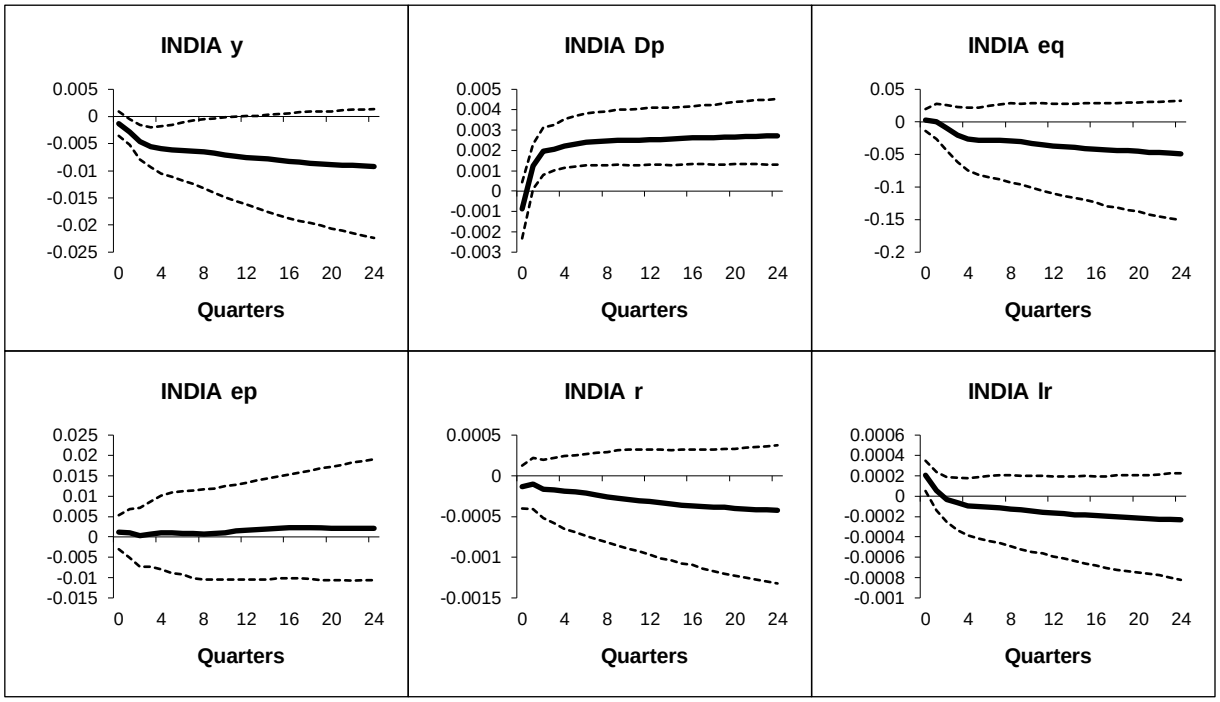

Figure B.2.29: India - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

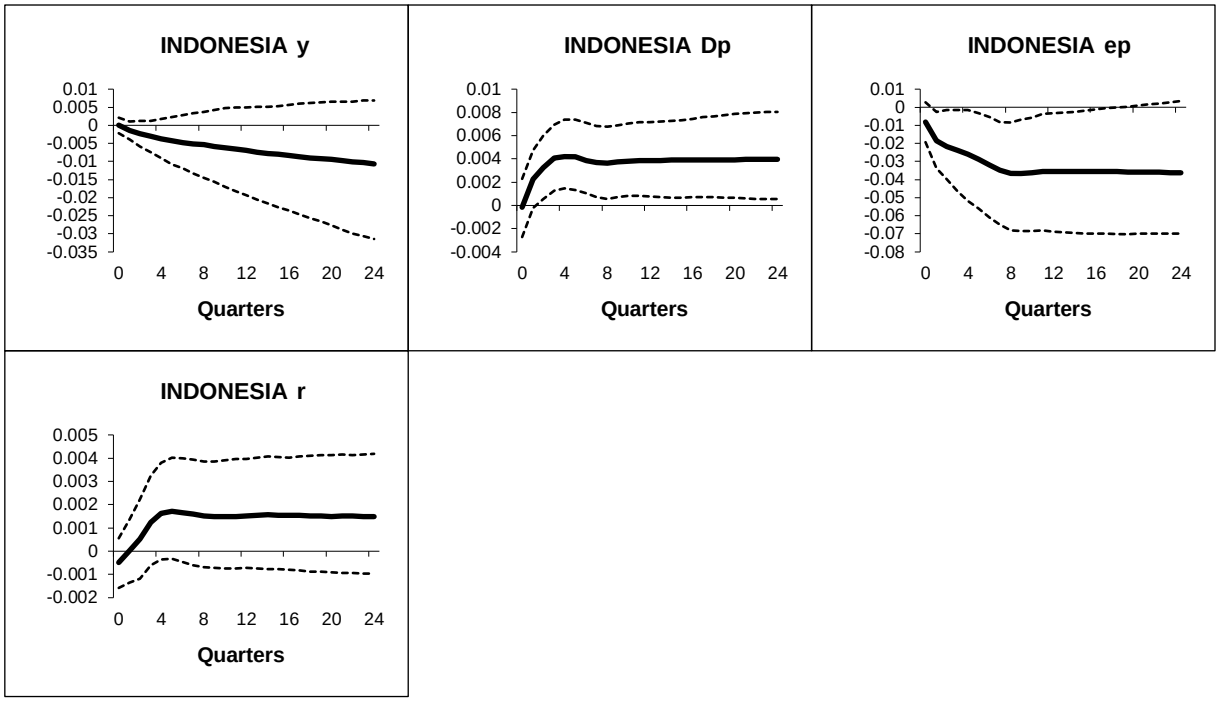

Figure B.2.30: Indonesia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

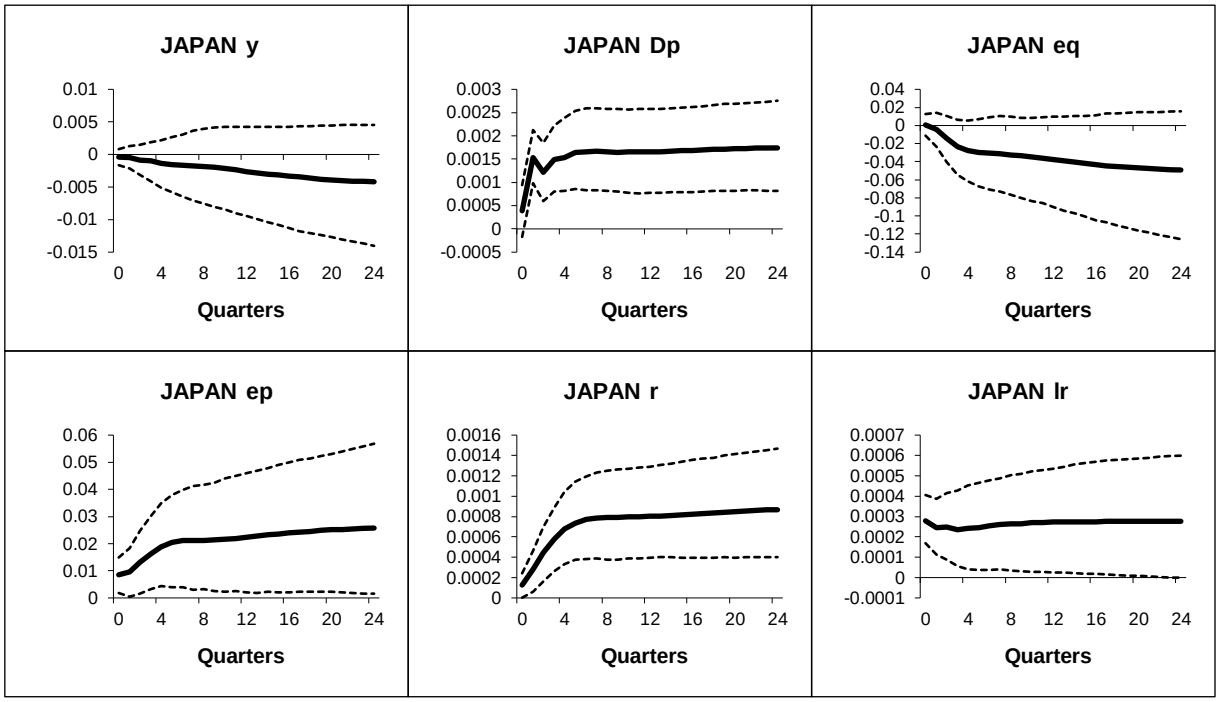

Figure B.2.31: Japan - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

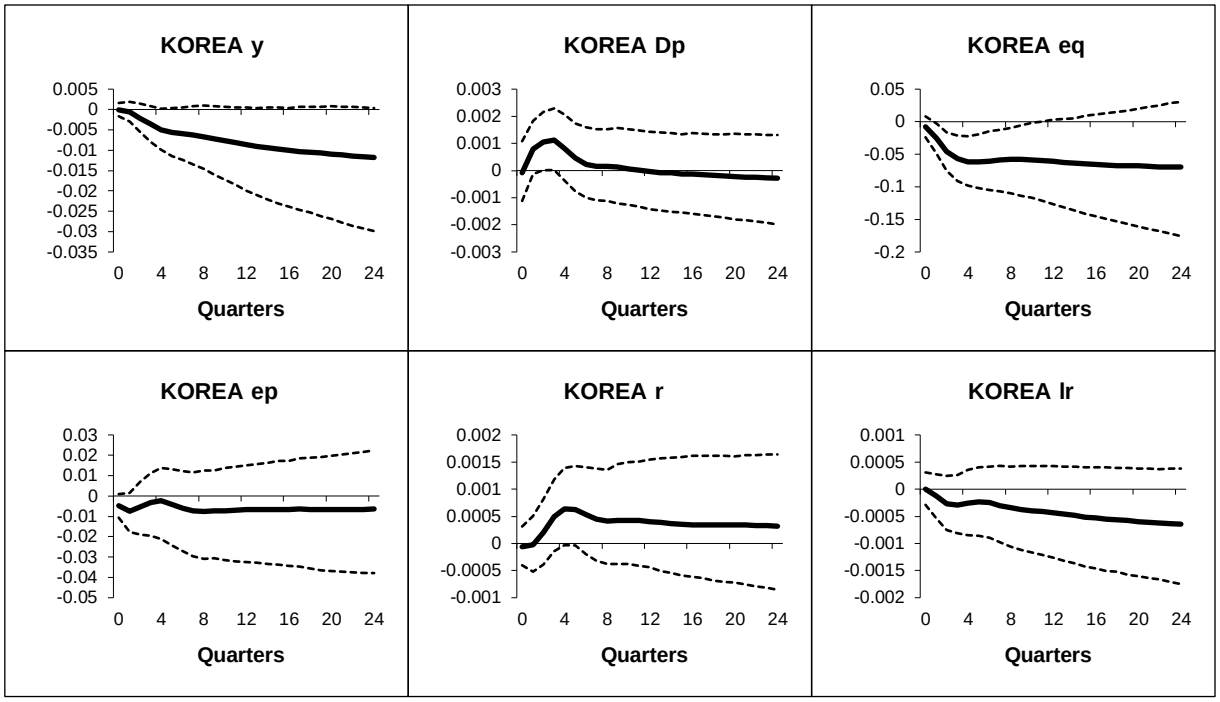

Figure B.2.32: Korea - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

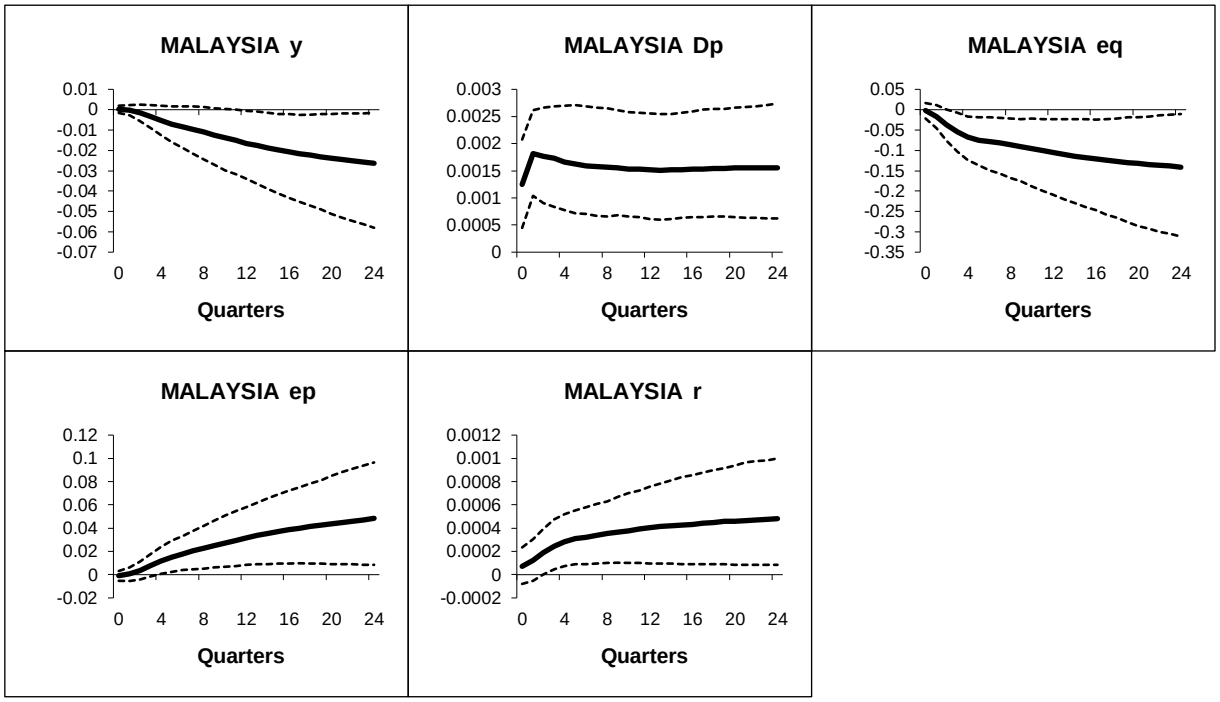

Figure B.2.33: Malaysia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

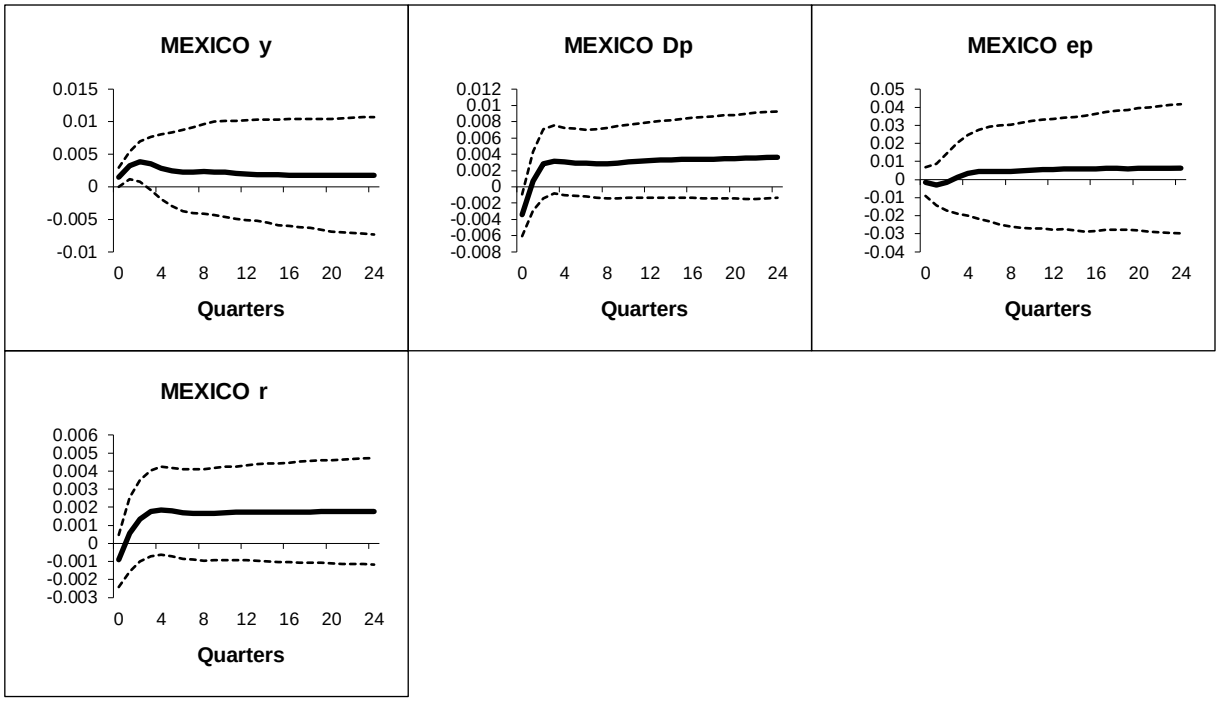

Figure B.2.34: Mexico - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

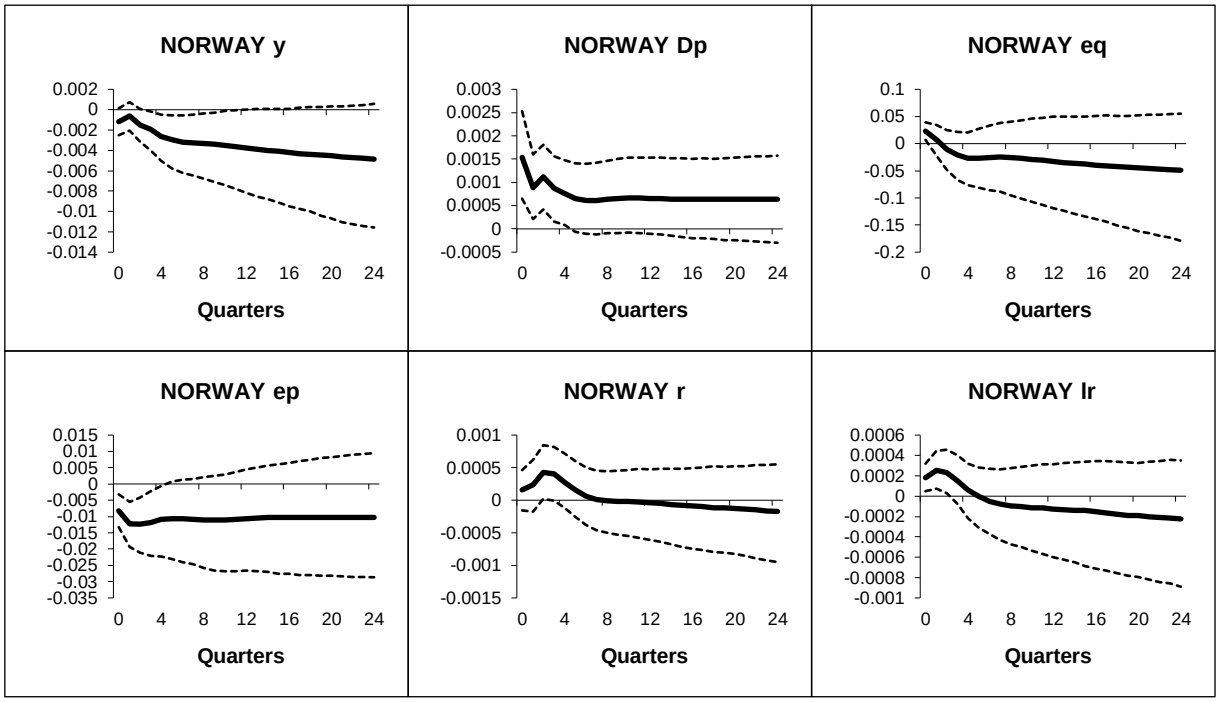

Figure B.2.35: Norway - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

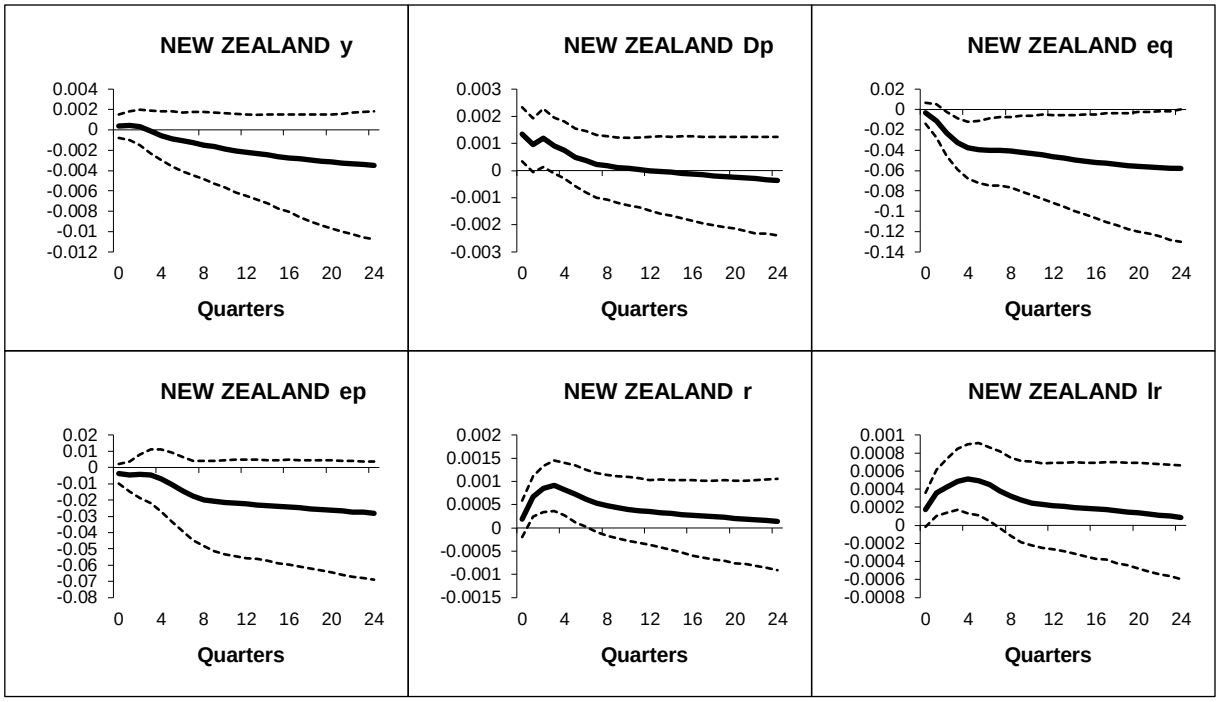

Figure B.2.36: New Zealand - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

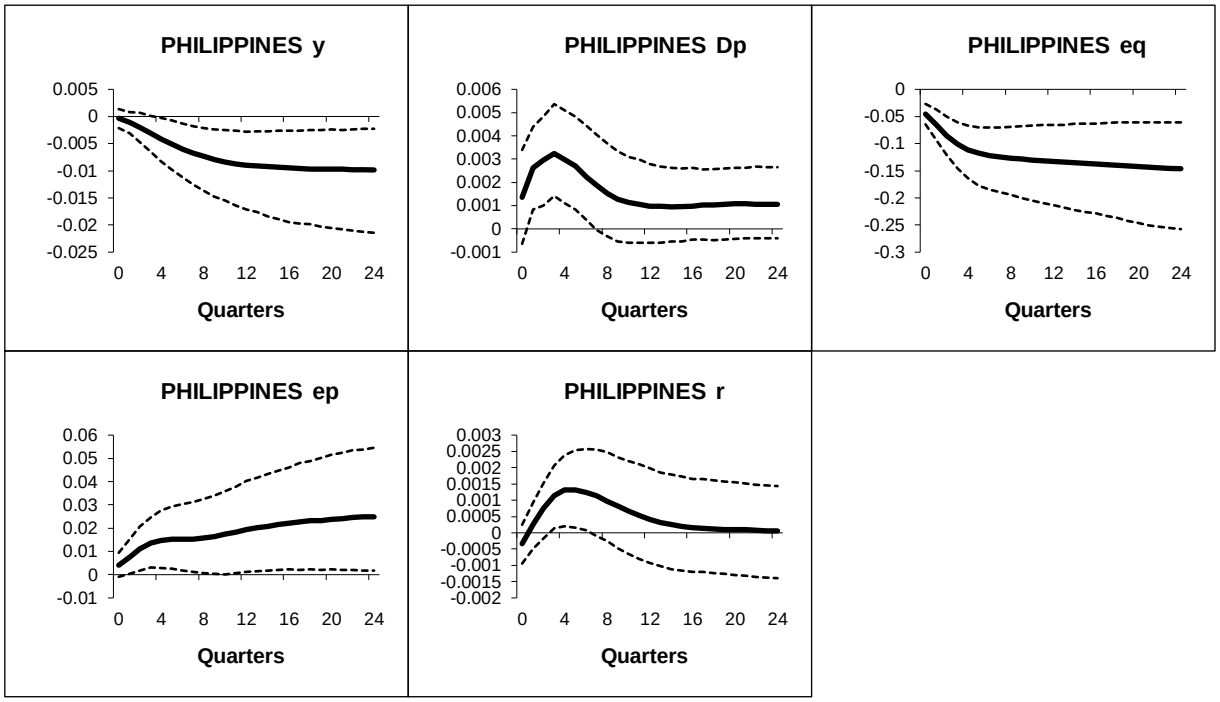

Figure B.2.37: Philippines - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

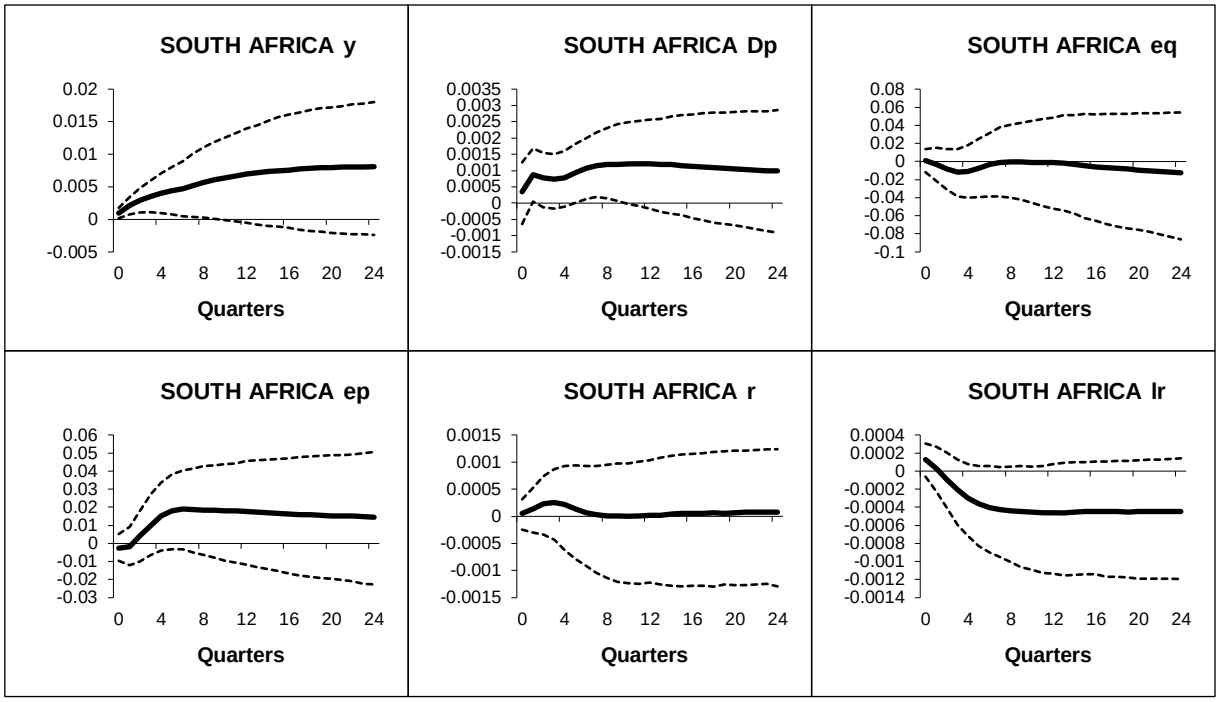

Figure B.2.38: South Africa - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

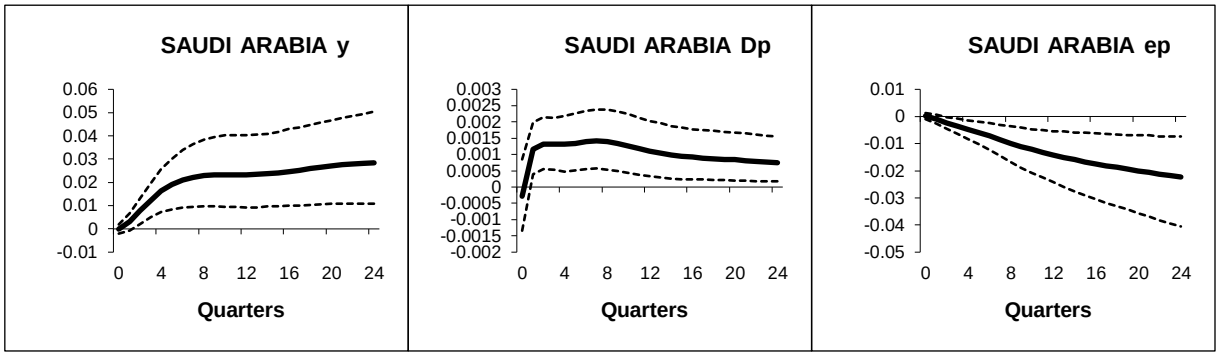

Figure B.2.39: Saudi Arabia - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

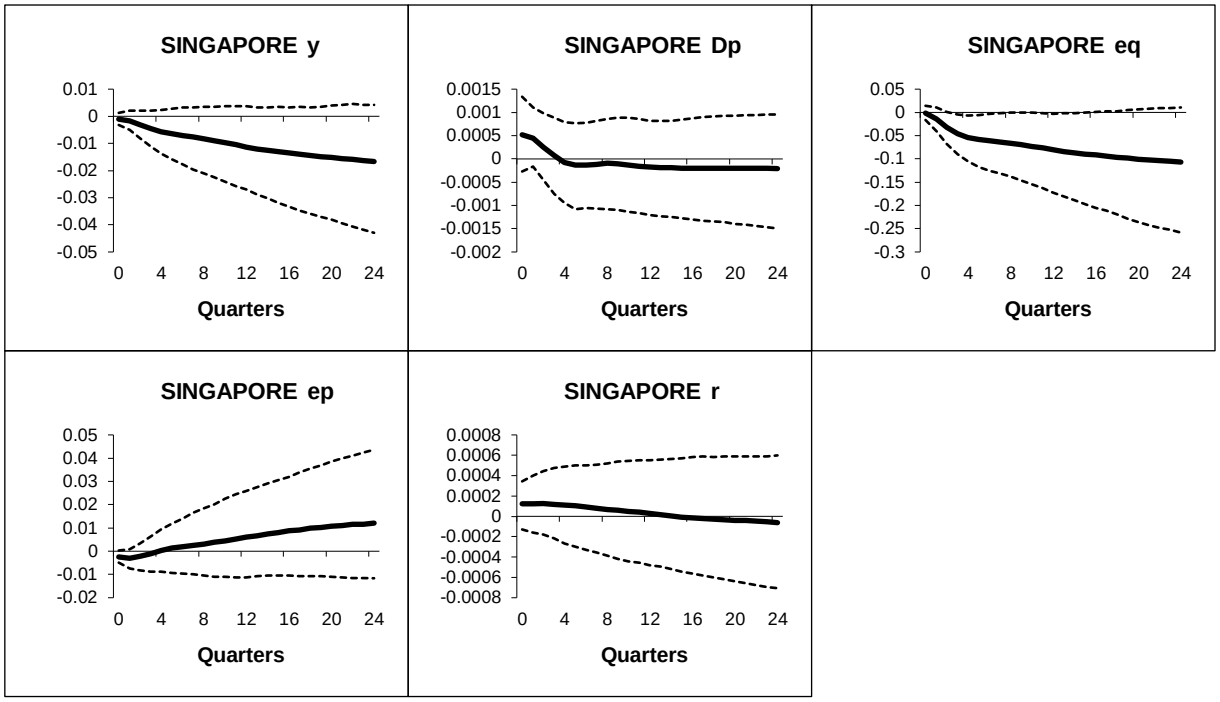

Figure B.2.40: Singapore - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

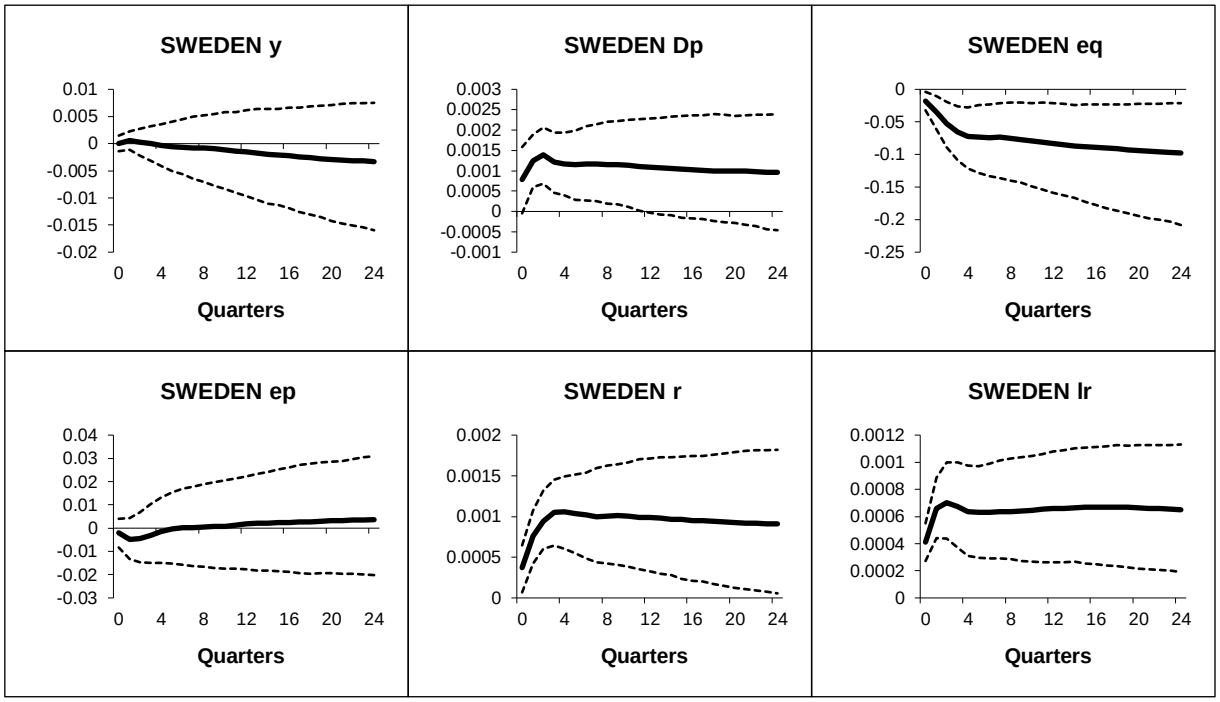

Figure B.2.41: Sweden - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

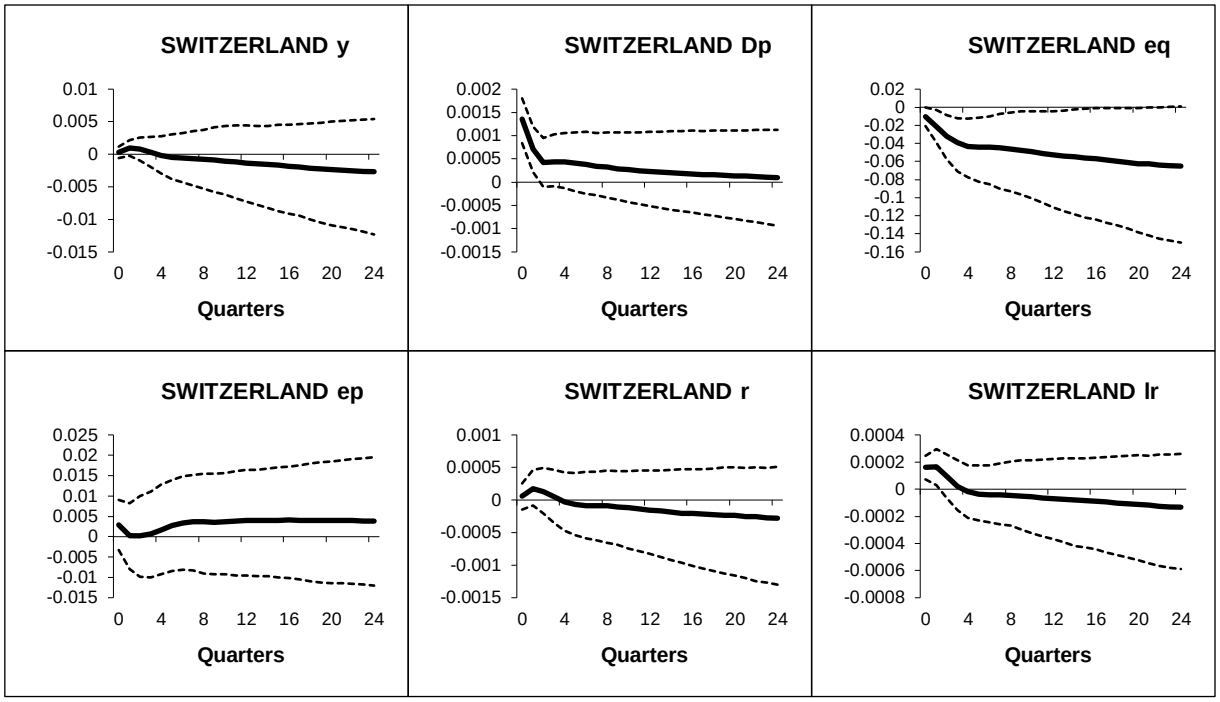

Figure B.2.42: Switzerland - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

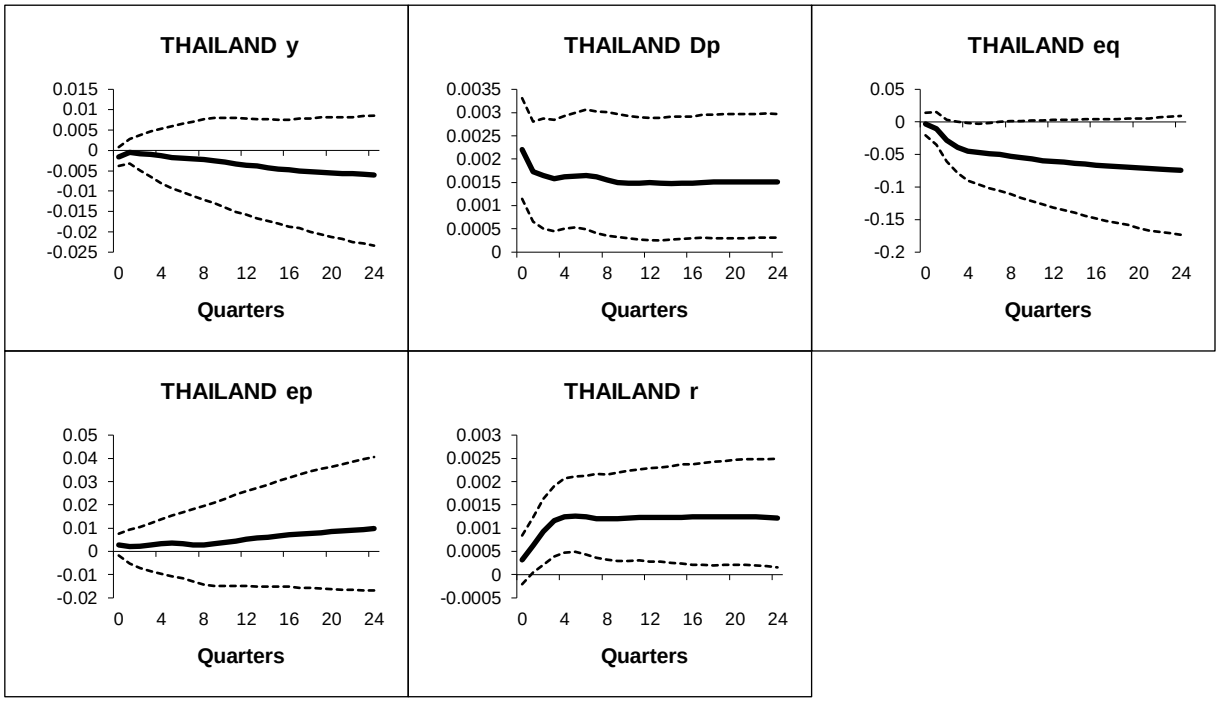

Figure B.2.43: Thailand - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

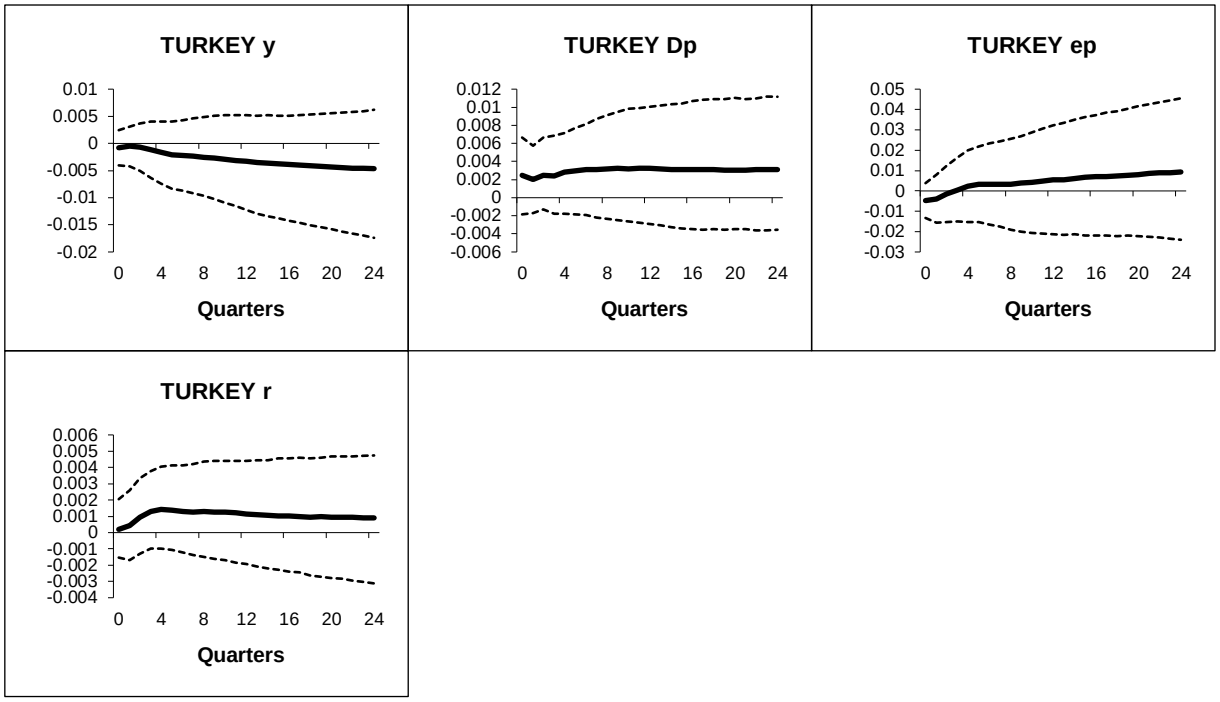

Figure B.2.44: Turkey - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

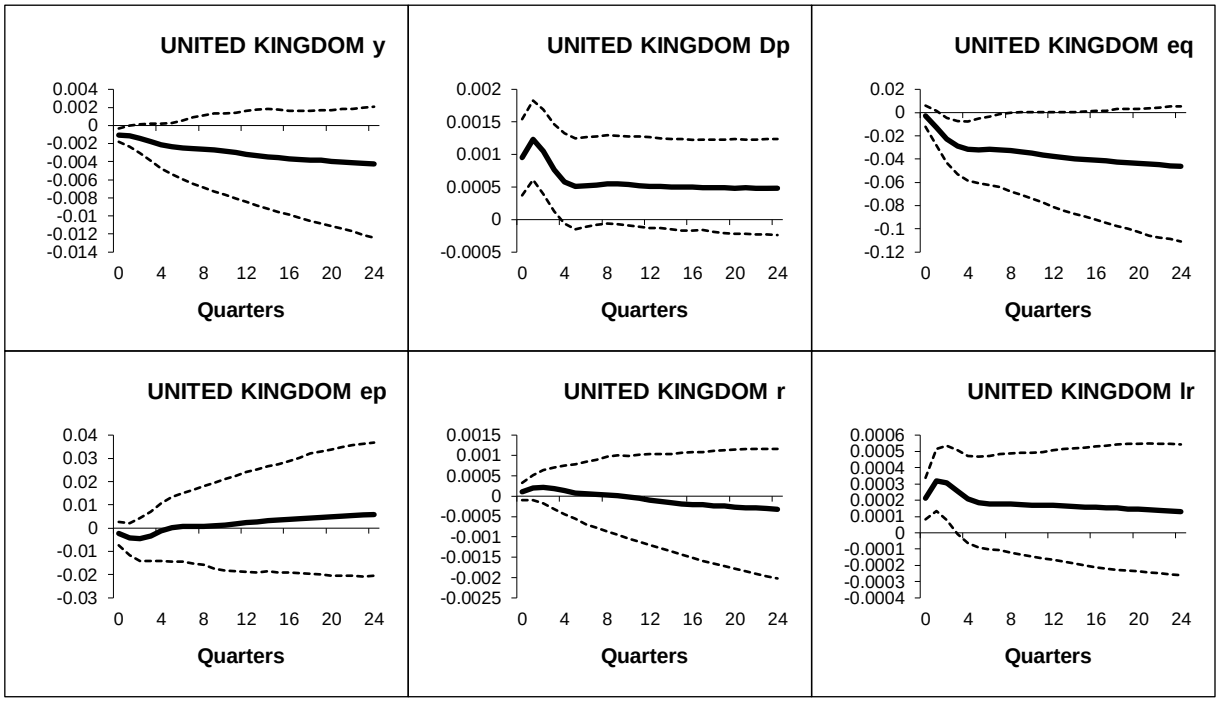

Figure B.2.45: UK - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

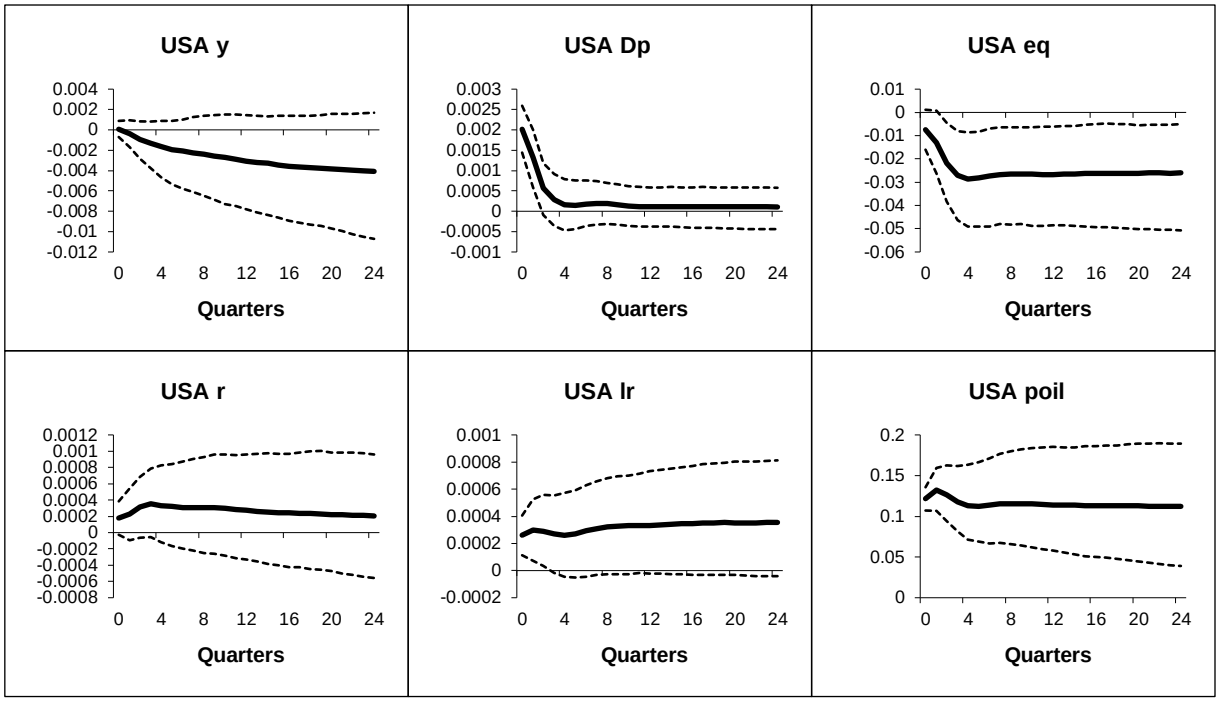

Figure B.2.46: USA - Generalized impulse responses of a positive (1 s.e.) shock to oil prices in the US model (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

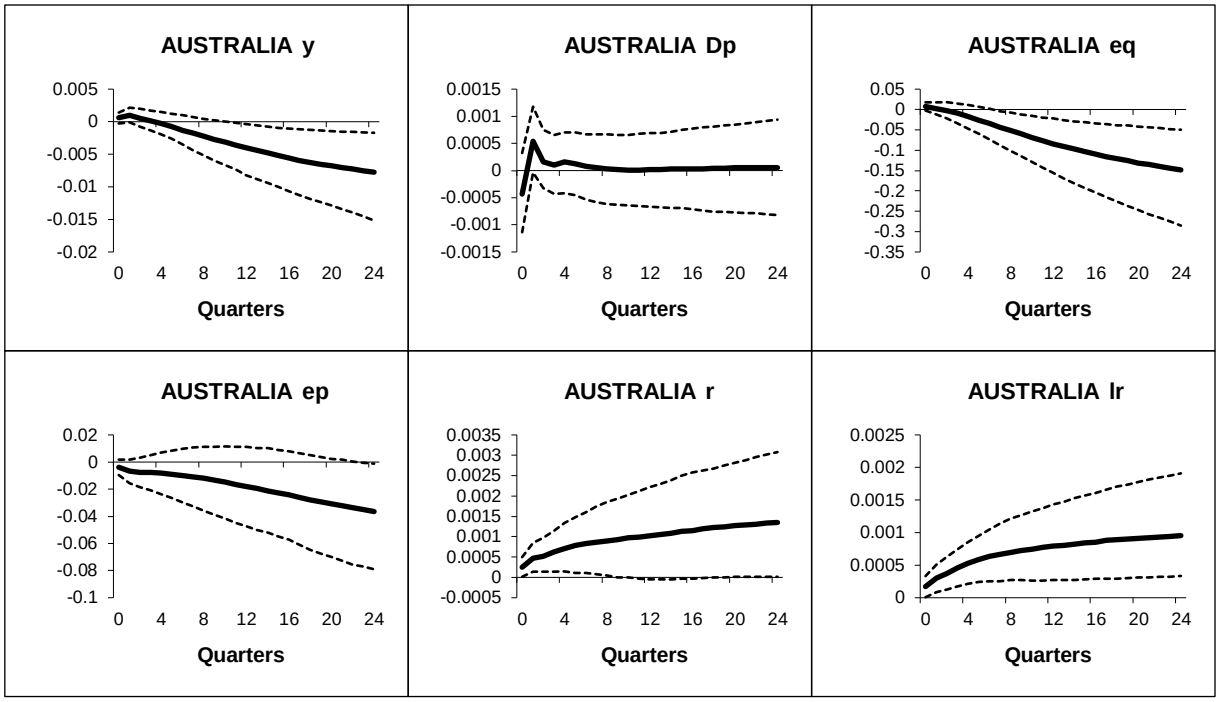

Figure B.2.47: Australia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

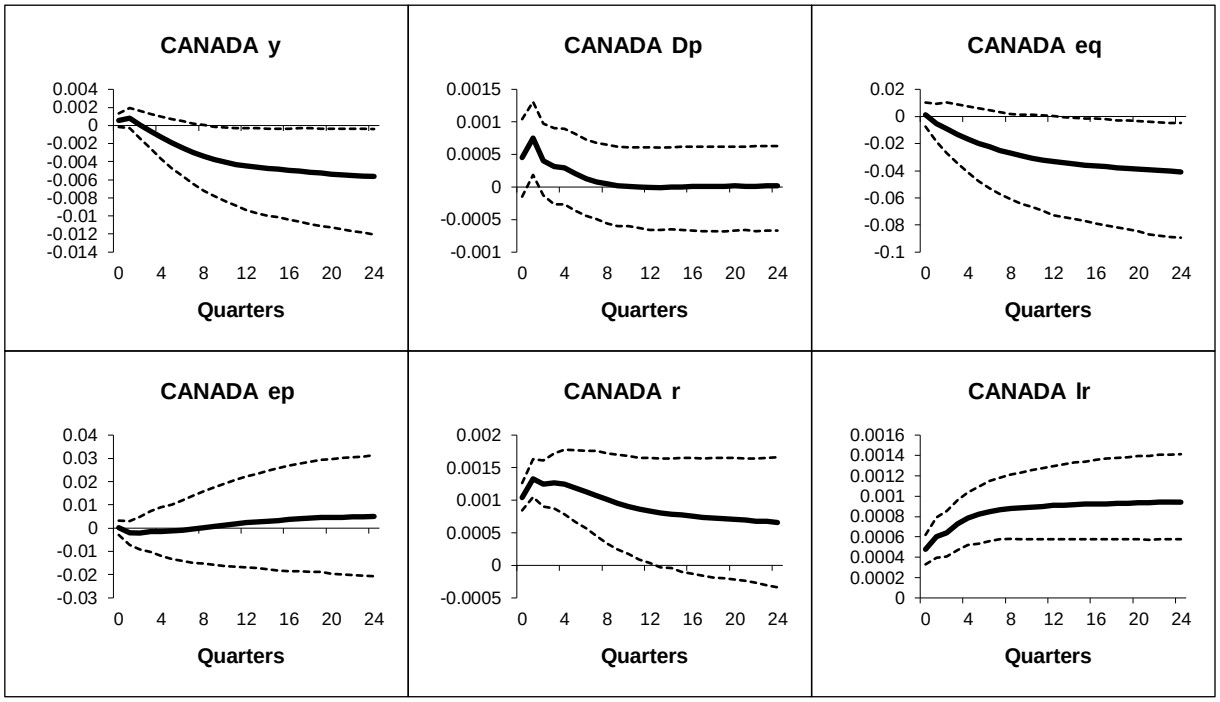

Figure B.2.48: Canada - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

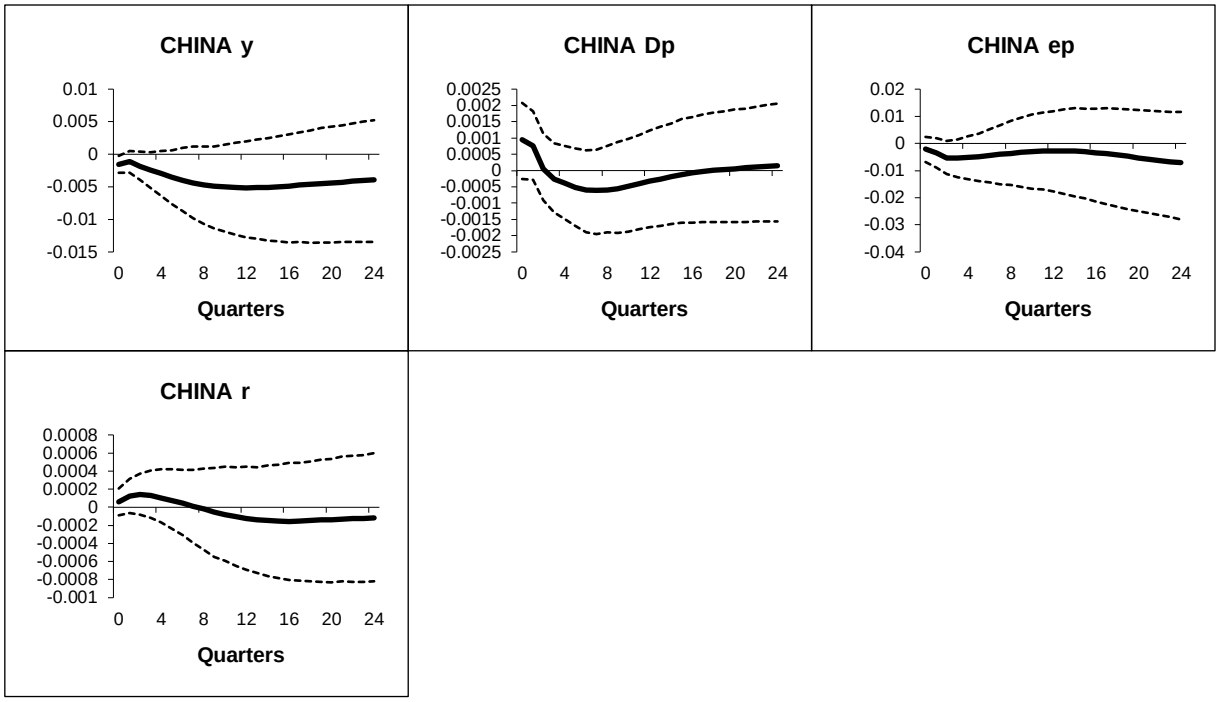

Figure B.2.49: China - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

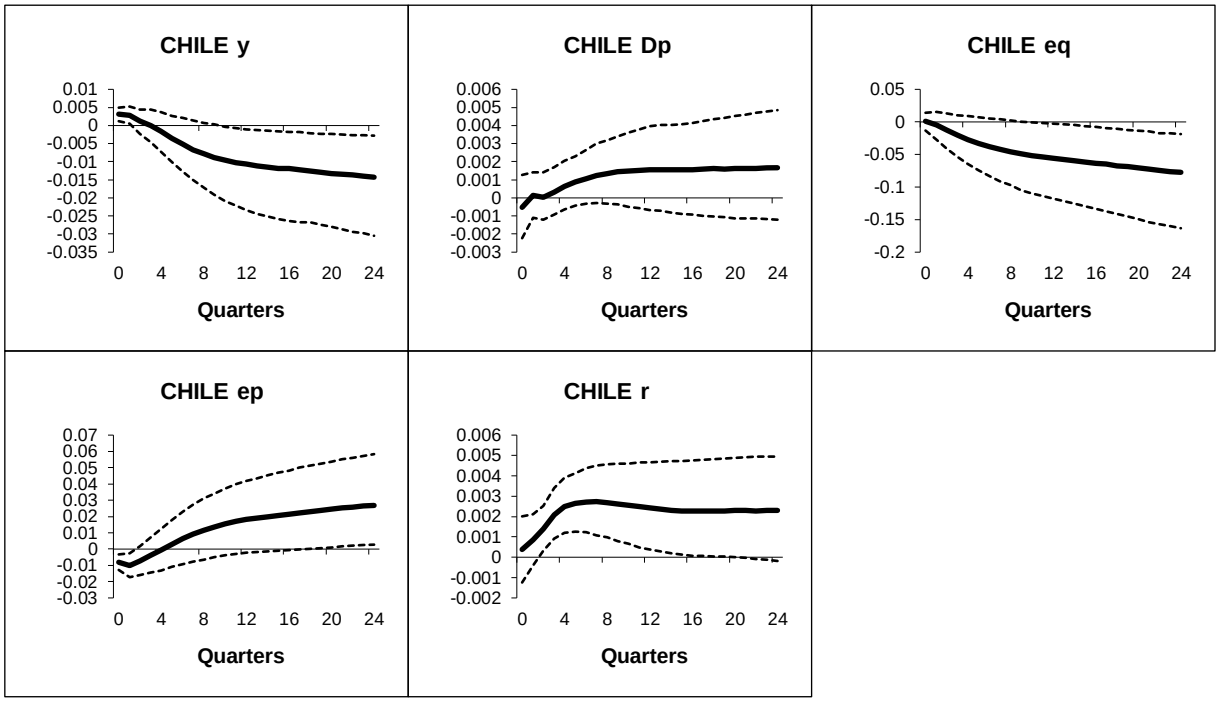

Figure B.2.50: Chile - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

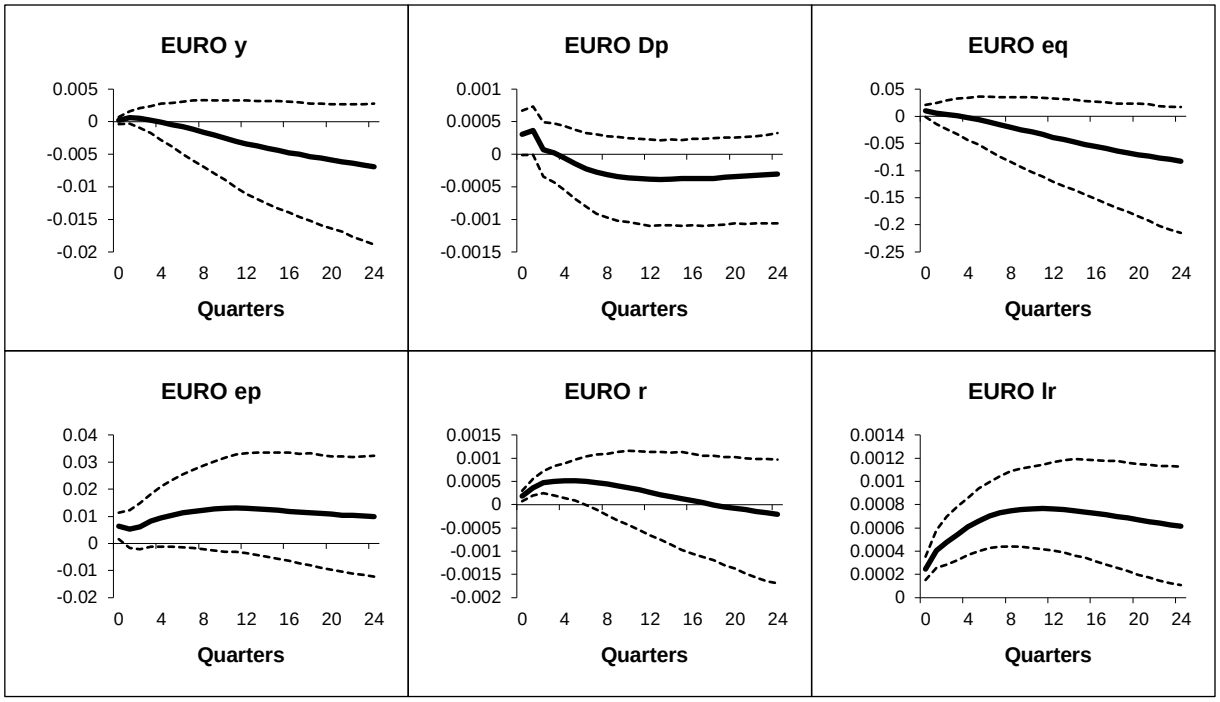

Figure B.2.51: Euro - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

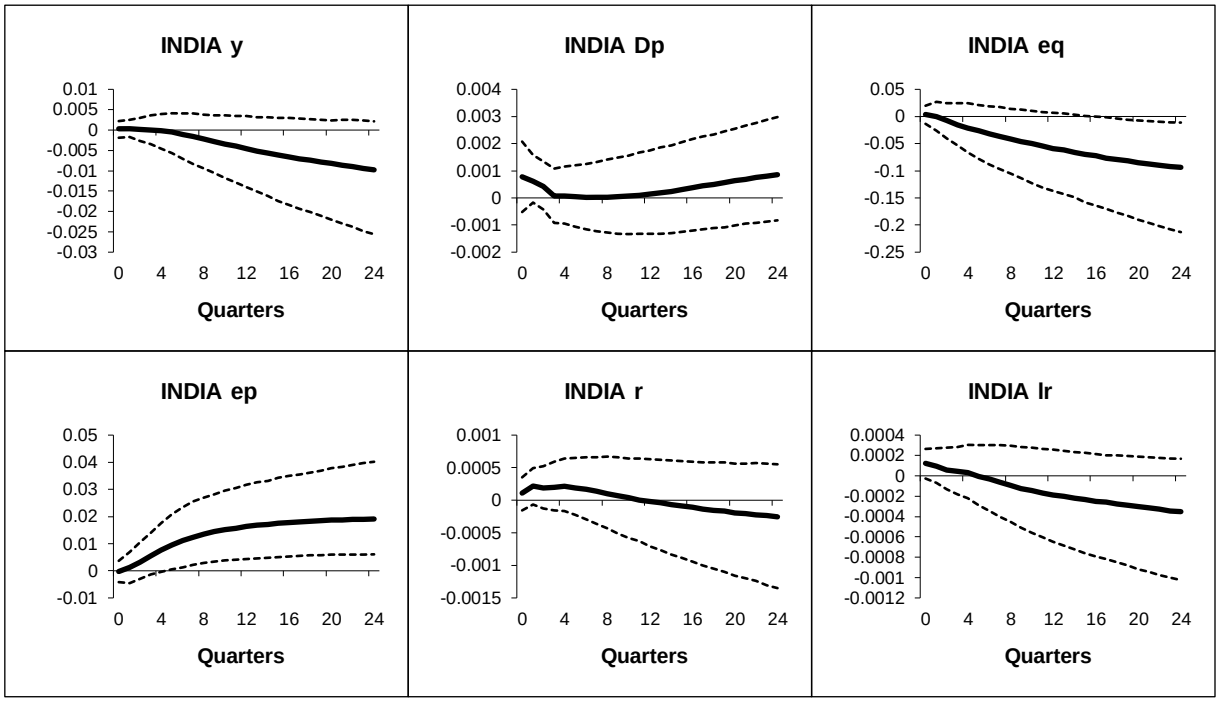

Figure B.2.52: India - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

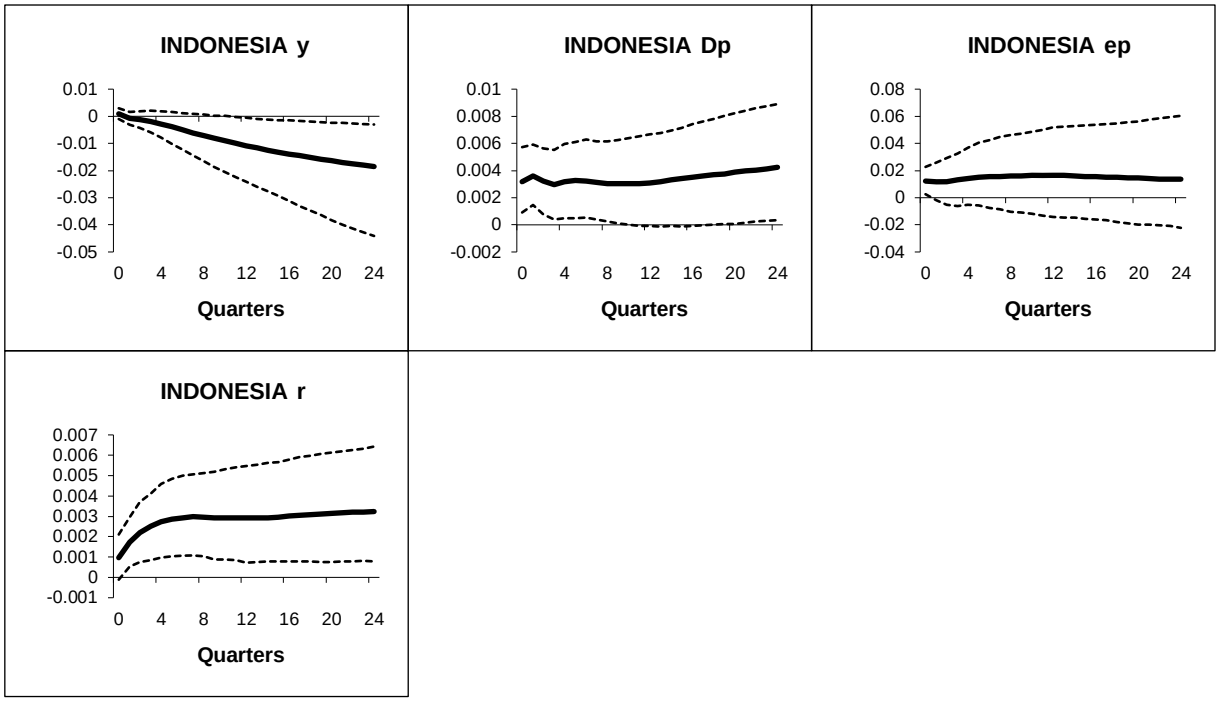

Figure B.2.53: Indonesia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

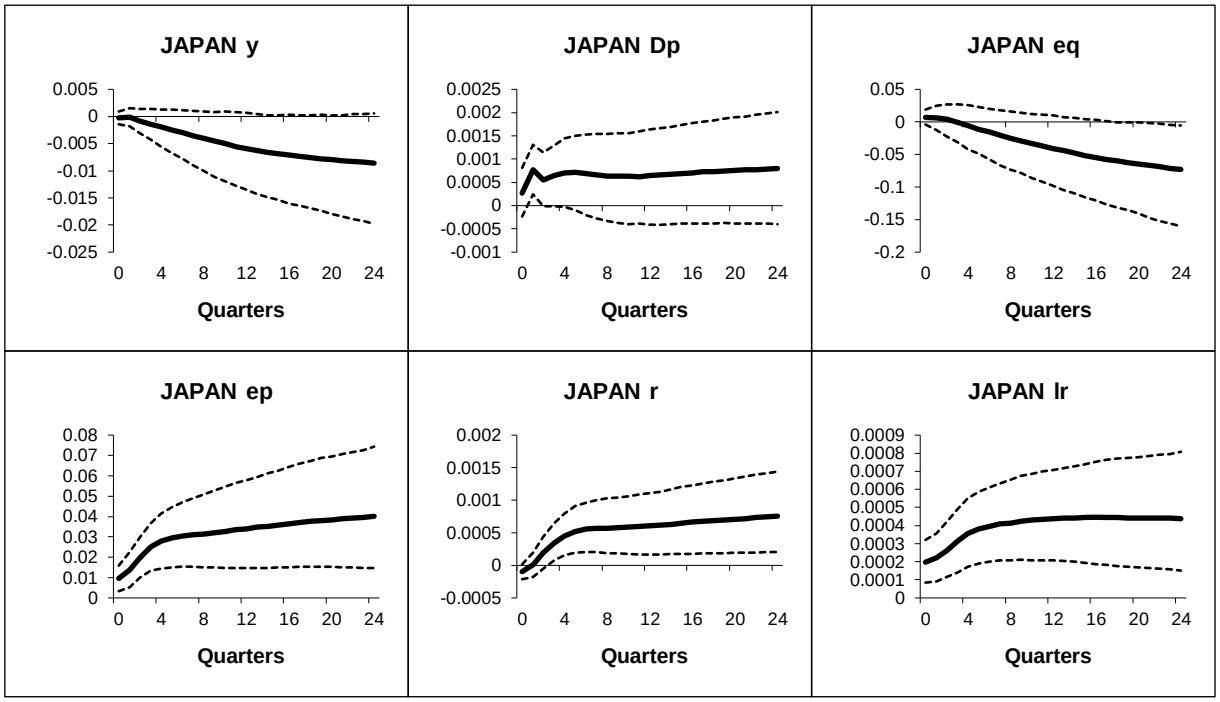

Figure B.2.54: Japan - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

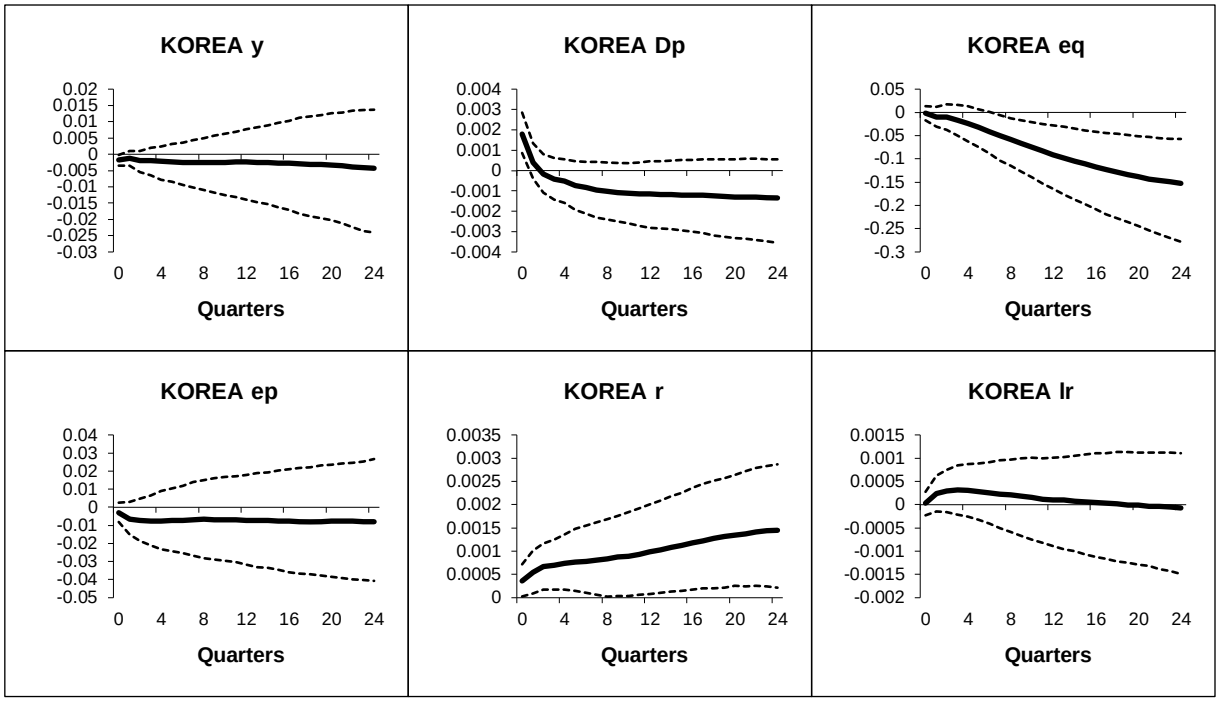

Figure B.2.55: Korea - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

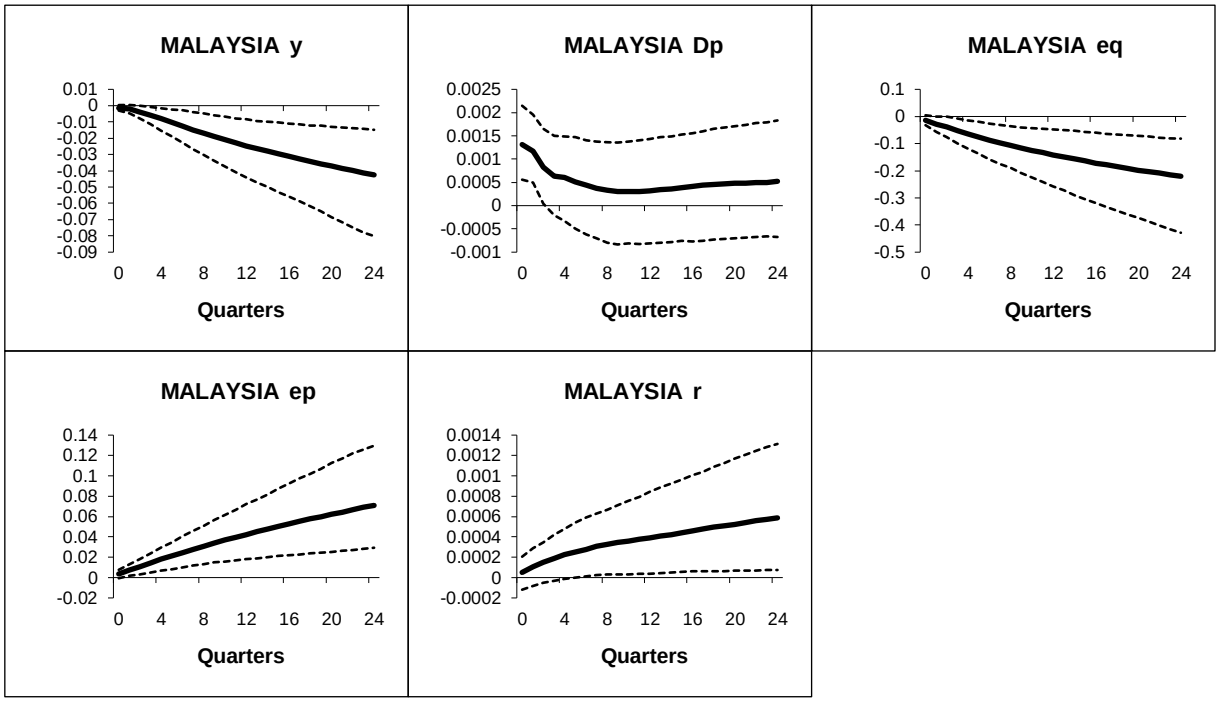

Figure B.2.56: Malaysia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

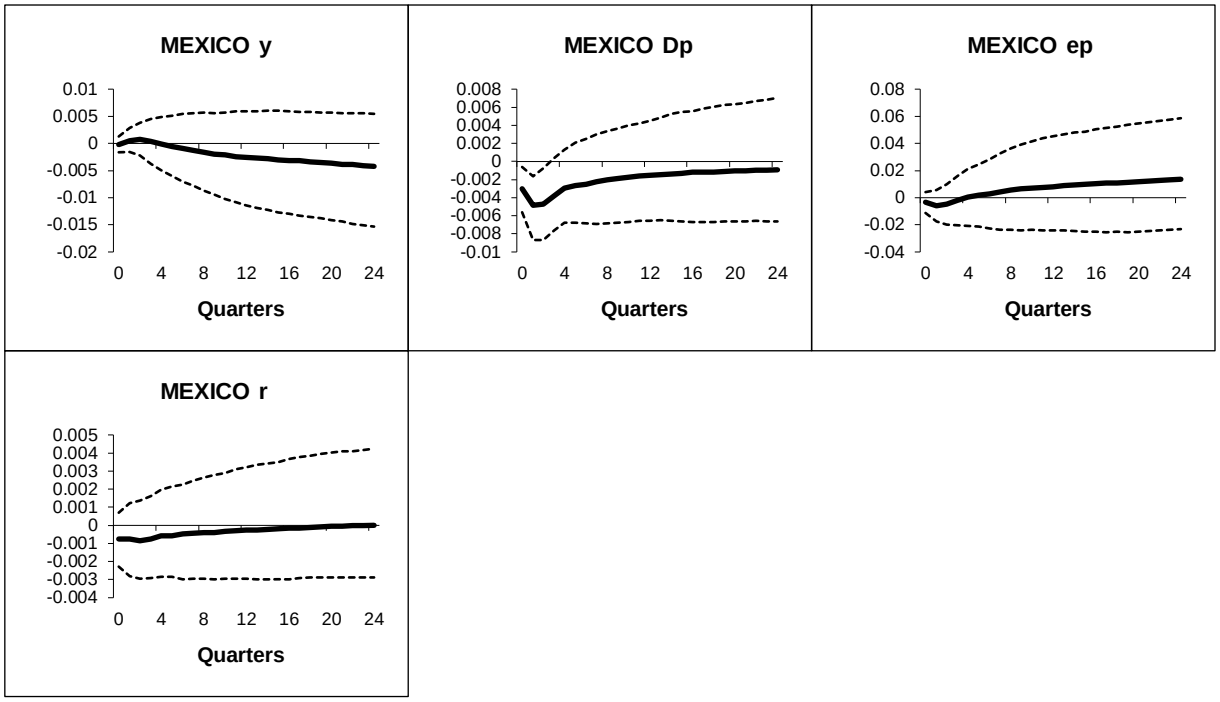

Figure B.2.57: Mexico - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

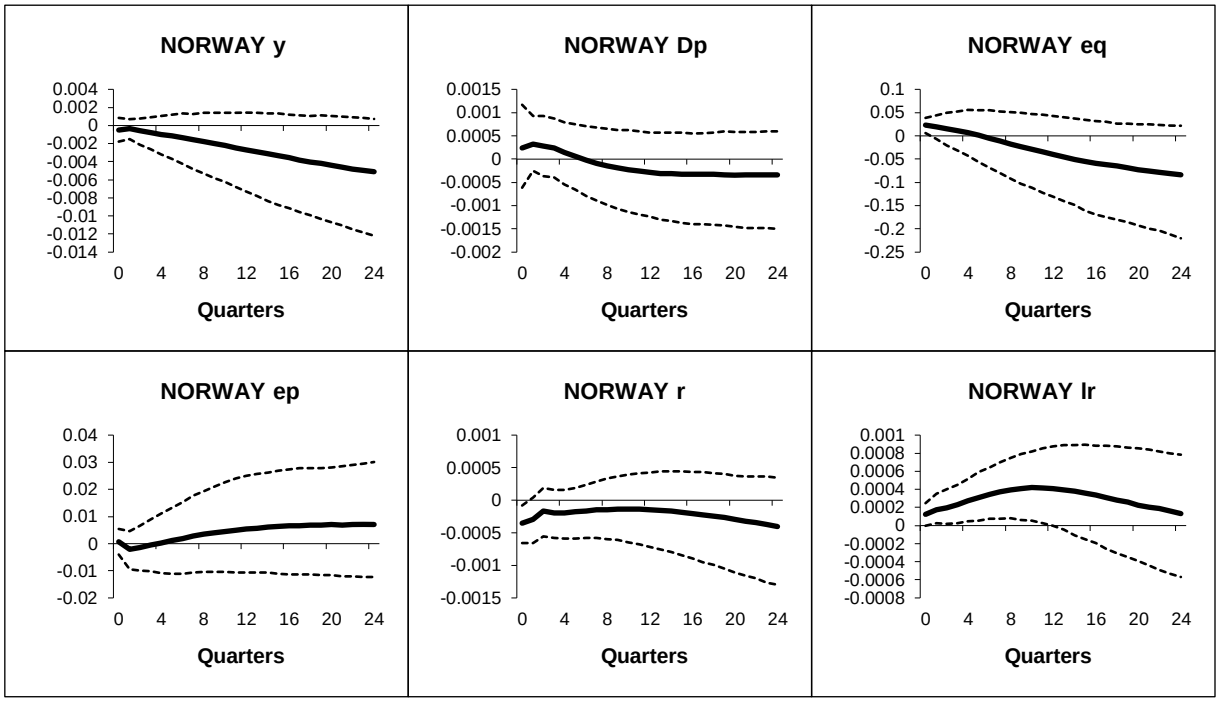

Figure B.2.58: Norway - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

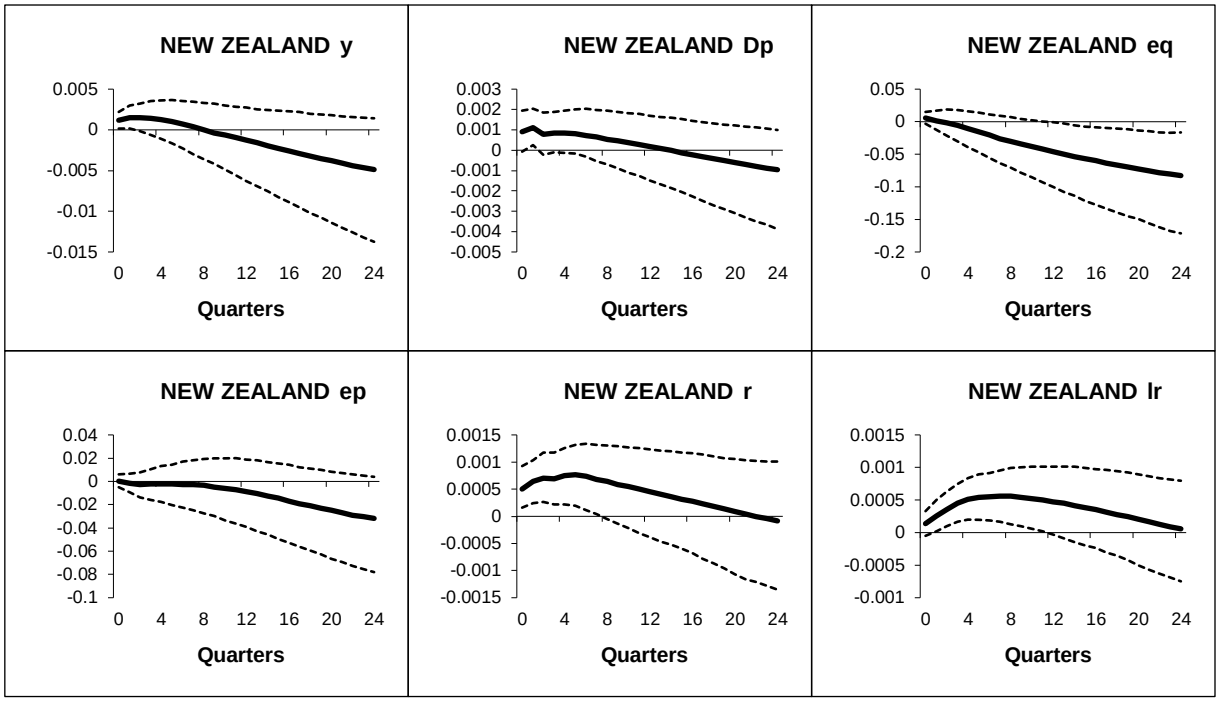

Figure B.2.59: New Zealand - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

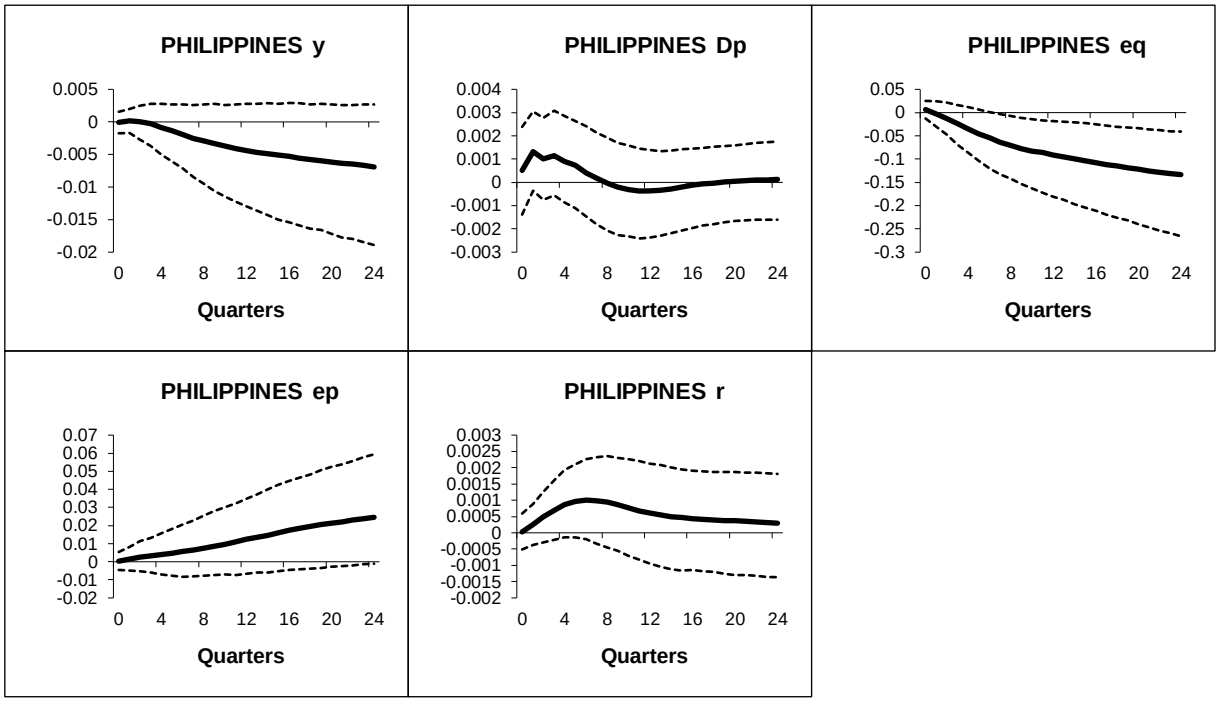

Figure B.2.60: Philippines - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

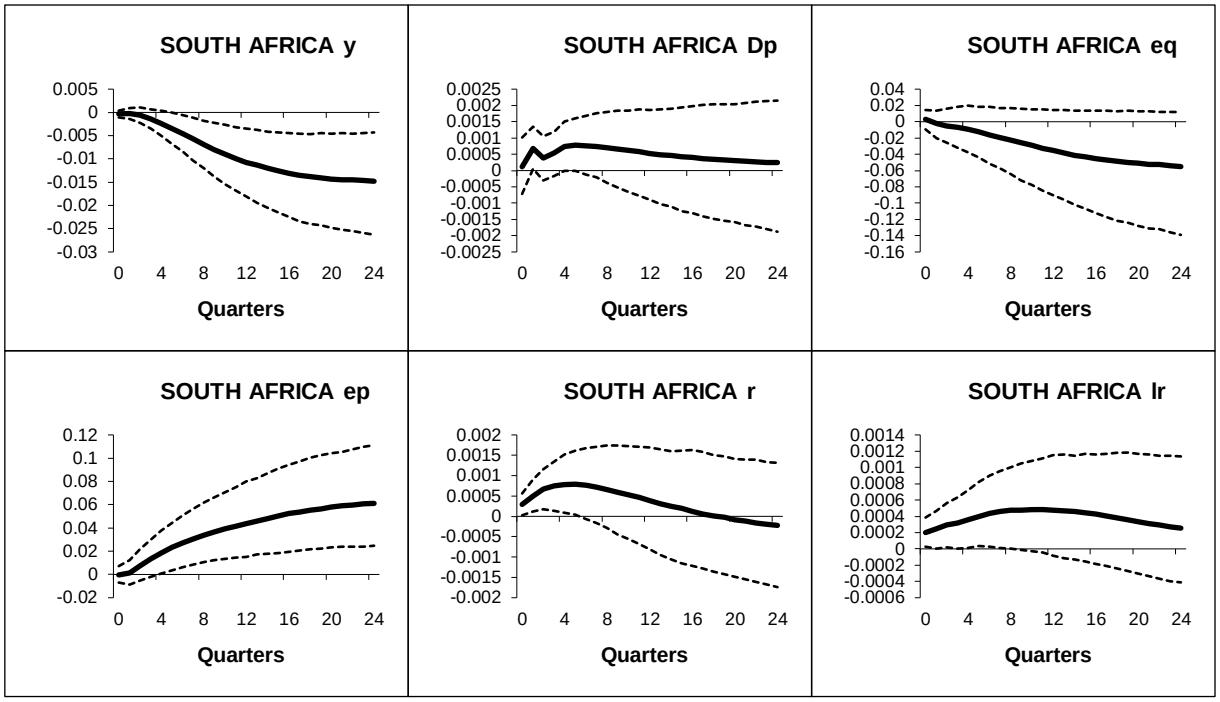

Figure B.2.61: South Africa - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

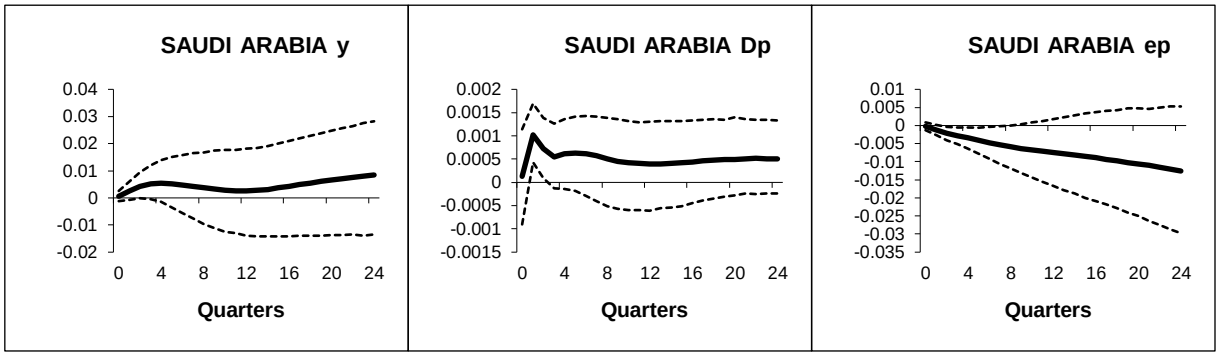

Figure B.2.62: Saudi Arabia - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

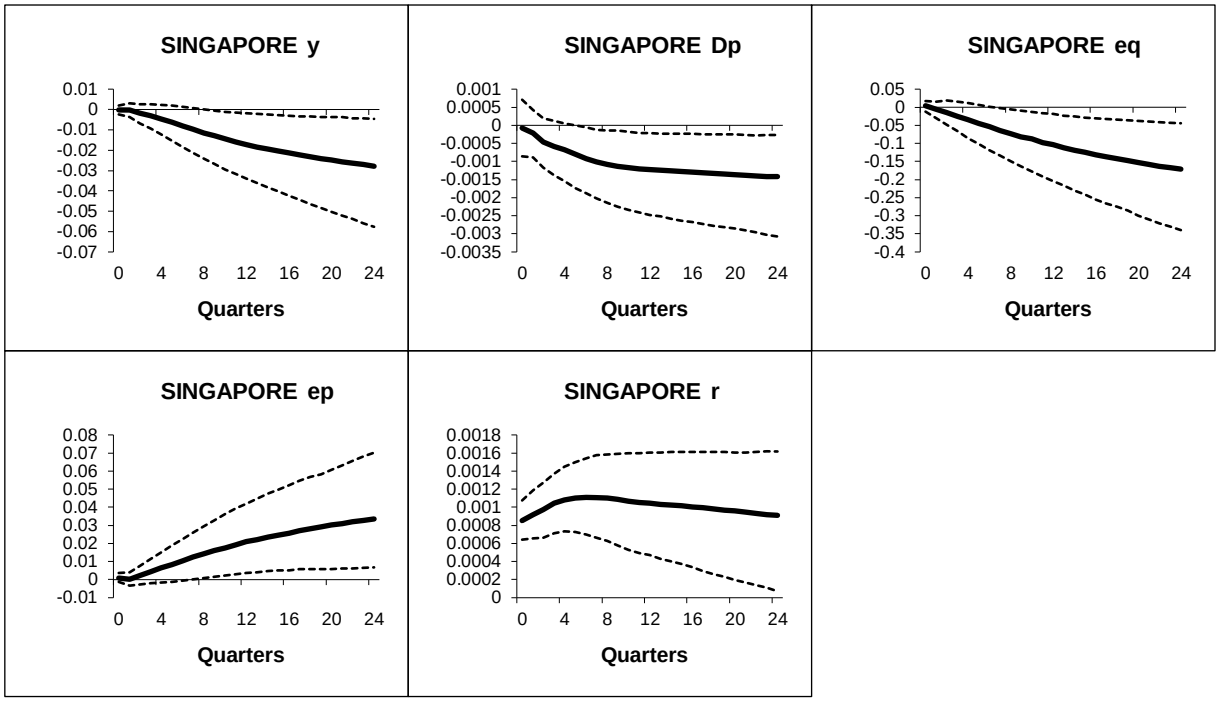

Figure B.2.63: Singapore - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

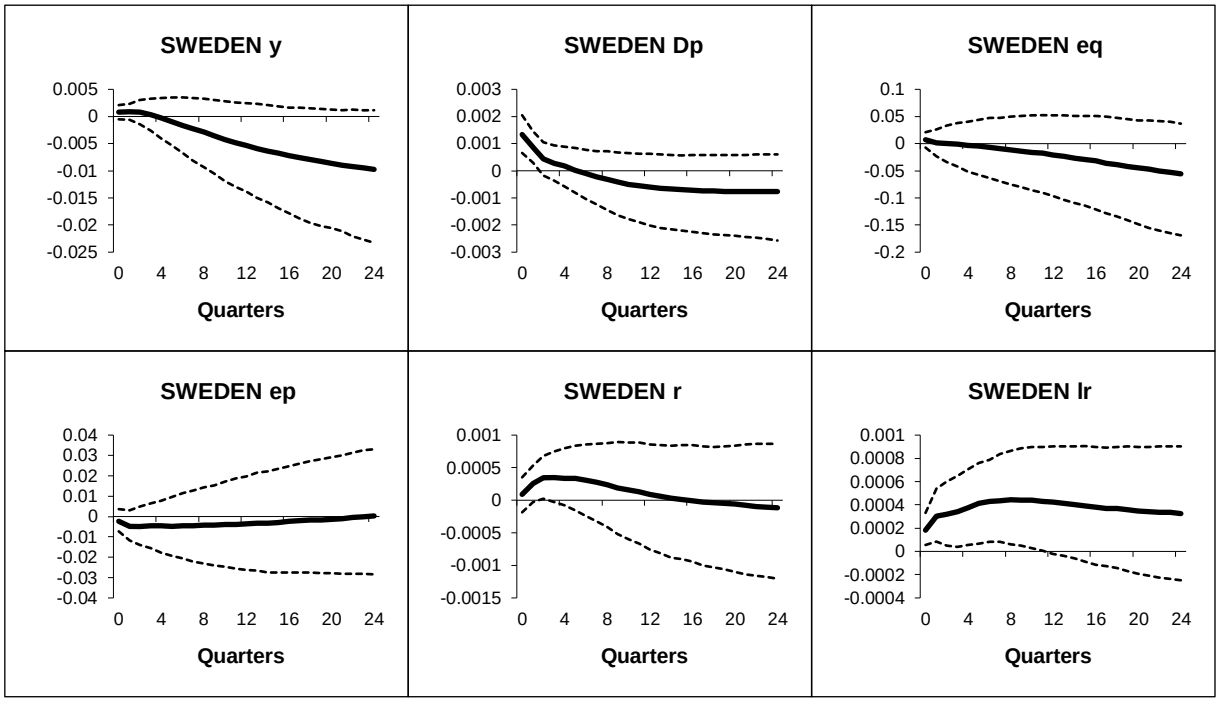

Figure B.2.64: Sweden - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

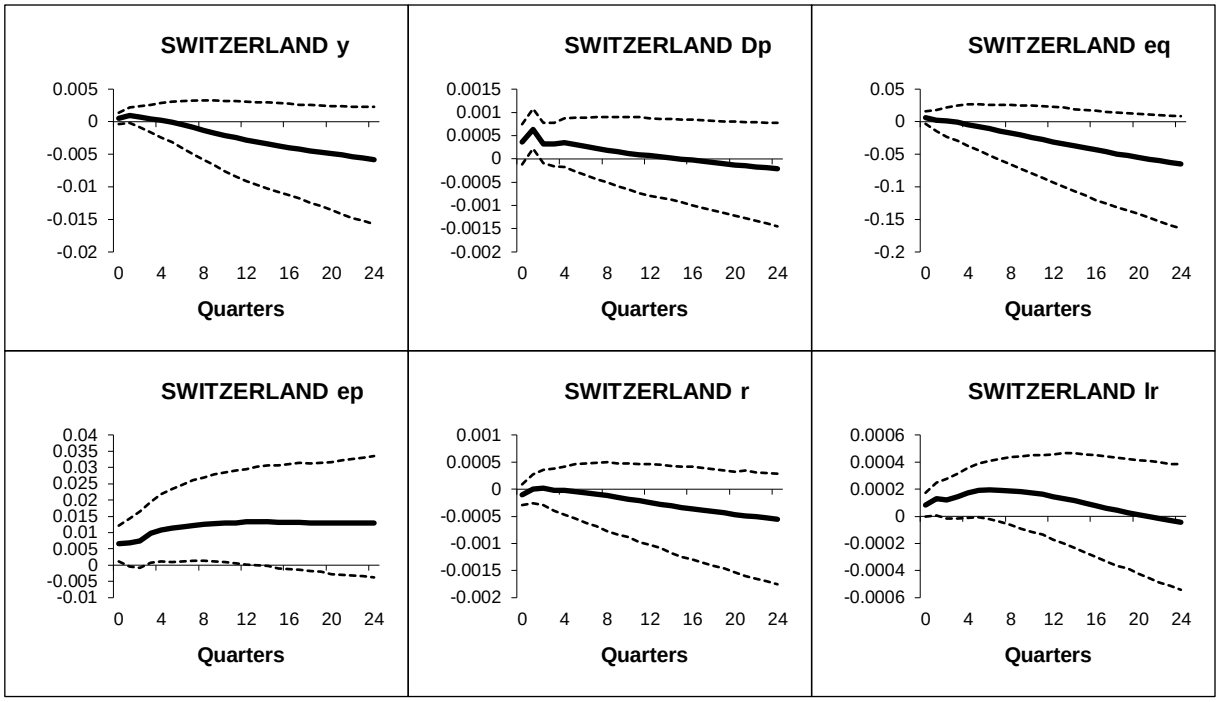

Figure B.2.65: Switzerland - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

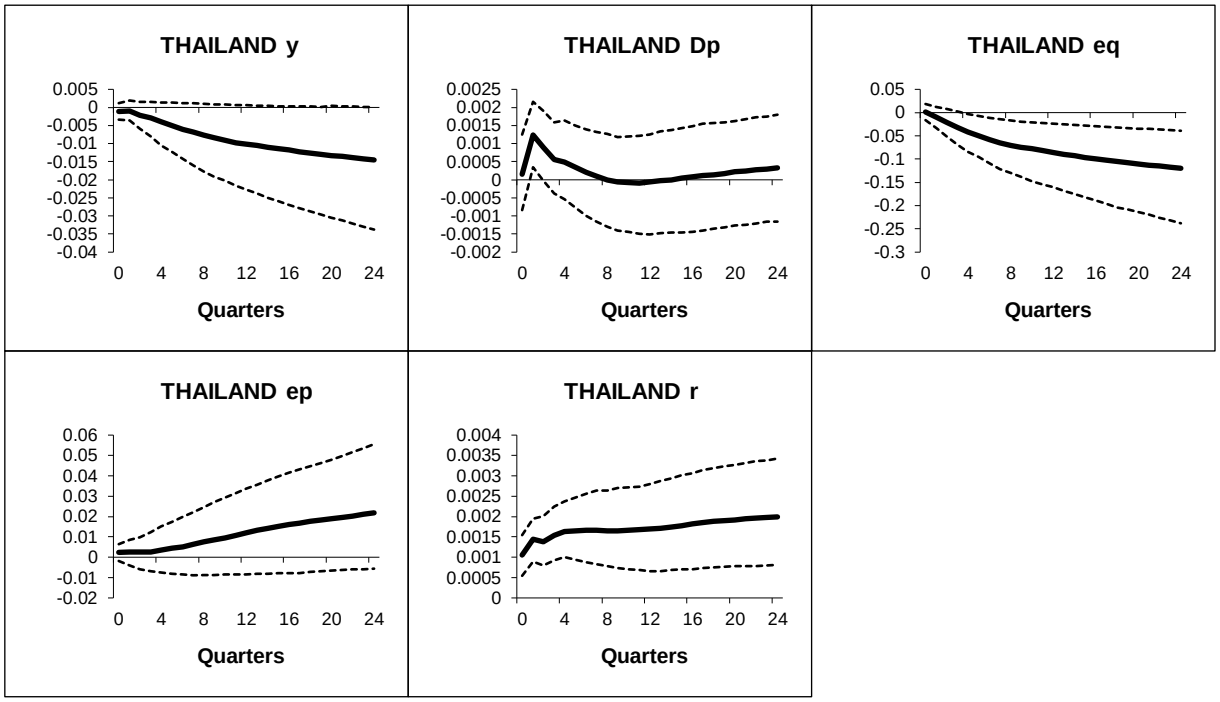

Figure B.2.66: Thailand - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

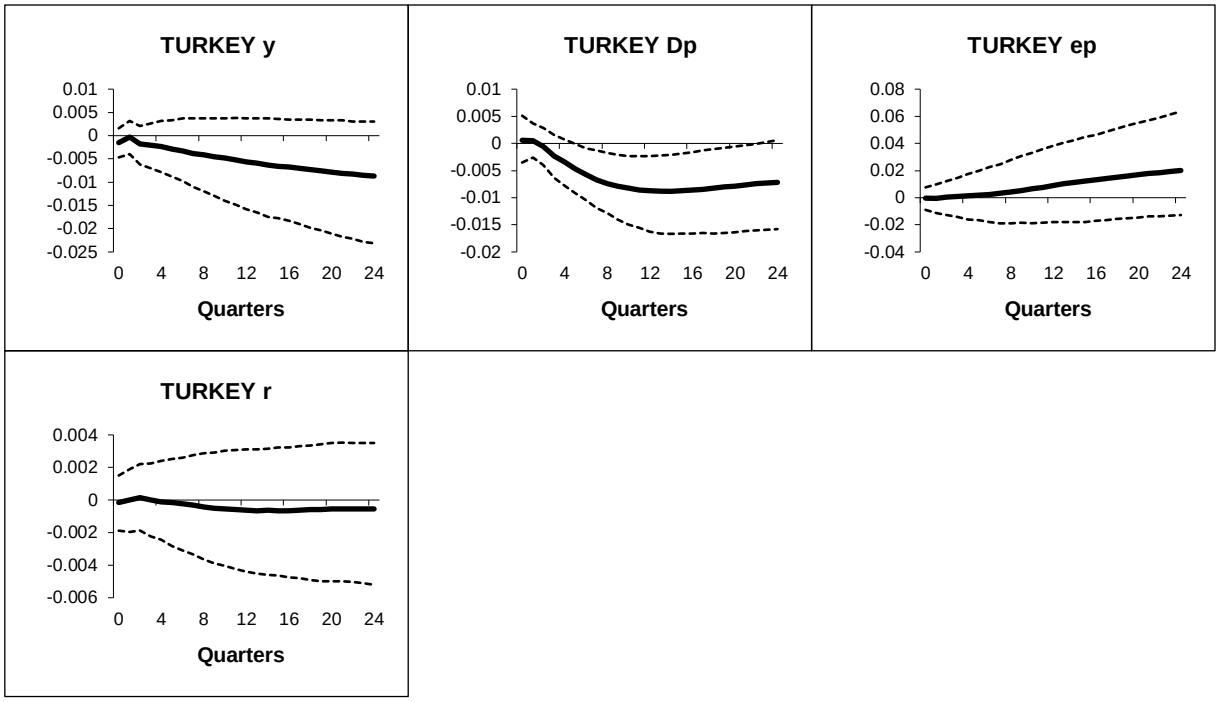

Figure B.2.67: Turkey - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

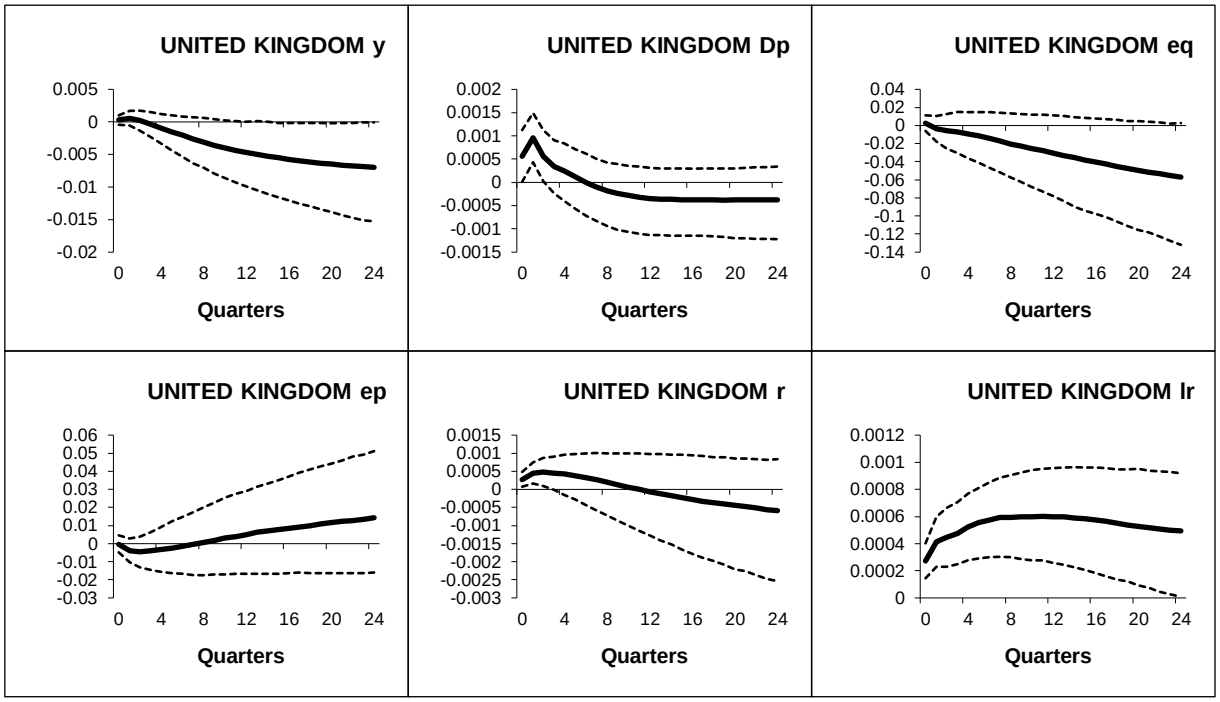

Figure B.2.68: UK - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

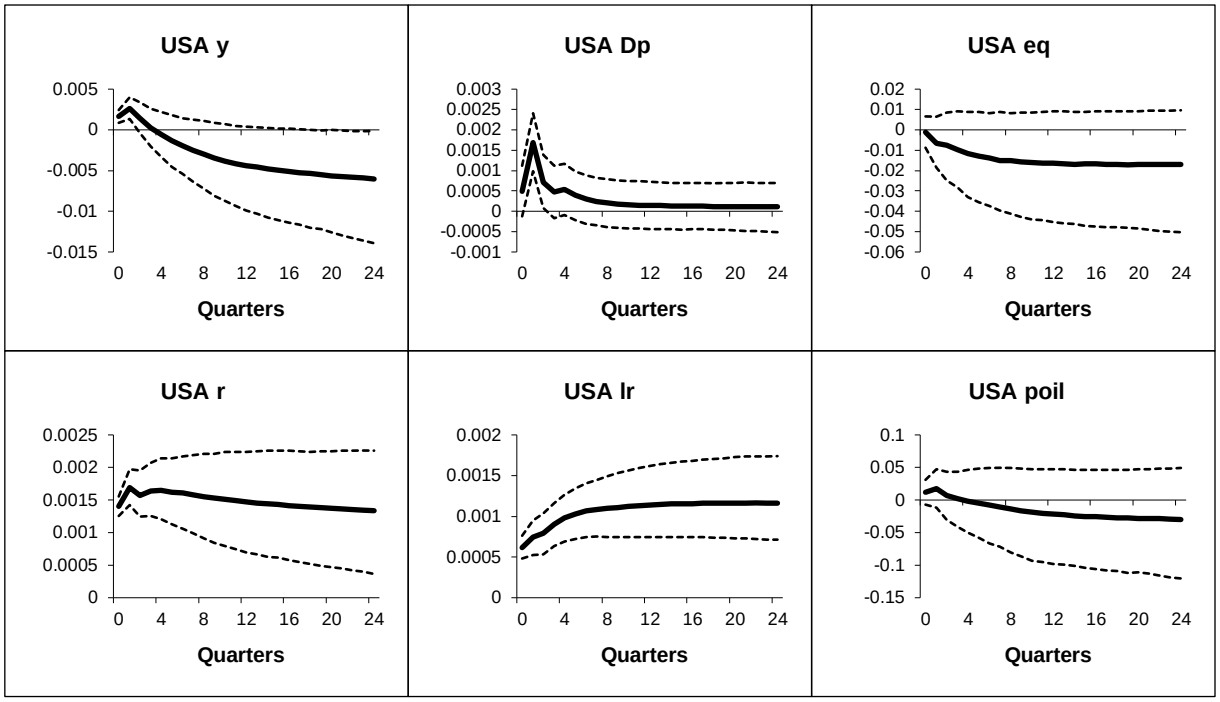

Figure B.2.69: USA - Generalized impulse responses of a positive (1 s.e.) shock to US short-rates (bootstrap mean estimates with 90% bootstrap error bounds). Trade weighted model IRFs.

BIBLIOGRAPHY

Ait-Sahalia, Y., Andritzky, J., Jobst, A., Nowak, S., and Tamirisa, N. (2012). Market response to policy initiatives during the global financial crisis. *Journal of International Economics*, 87(1):162-177.

Andrews, D. W. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica: Journal of the Econometric Society*, pages 1383-1414.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, volume 4. Springer Science & Business Media.

Anselin, L. et al. (1980). Estimation methods for spatial autoregressive structures. *Estimation methods for spatial autoregressive structures.*, (8).

Baxter, M. and Kouparitsas, M. A. (2005). Determinants of business cycle comovement: a robust analysis. *Journal of Monetary Economics*, 52(1):113-157.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261-304.

Bussiere, M., Chudik, A., and Mehl, A. (2011). Does the euro make a di erence? spatio-temporal transmission of global shocks to real effective exchange rates in an infinite var. ECB working paper 1292, European Central Bank.

Choi, C.-Y. and Chudik, A. (2017). Geographic inequality of economic well-being among us cities: Evidence from micro panel data. Working paper, Federal Reserve Bank of Dallas Globalization and Monetary Policy Institute.

Chudik, A. and Fratzscher, M. (2011). Identifying the global transmission of the 2007-2009 financial crisis in a gvar model. European Economic Review, 55(3):325-339.

Chudik, A. and Pesaran, M. H. (2013). Econometric analysis of high dimensional vars featuring a dominant unit. *Econometric Reviews*, 32(5-6):592-649.

Chudik, A. and Pesaran, M. H. (2016). Theory and practice of gvar modelling. *Journal of Economic Surveys*, 30(1):165-197.

Cipollini, A., Parla, F., et al. (2018). Housing market shocks in italy: a gvar approach. Technical report, Universita di Modena e Reggio Emilia, Dipartimento di Economia "Marco Biagi".

Dees, S., Mauro, F. d., Pesaran, M. H., and Smith, L. V. (2007). Exploring the international linkages of the euro area: a global var analysis. *Journal of applied econometrics*, 22(1):1-38.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134-144.

Eickmeier, S. and Ng, T. (2015). How do us credit supply shocks propagate internationally? a gvar approach. *European Economic Review*, 74:128-145.

Elhorst, J. P., Gross, M., and Tereanu, E. (2018). Spillovers in space and time: Where spatial econometrics and global var models meet. unpublished.

Feldkircher, M. and Huber, F. (2016). The international transmission of US shocks - Evidence from bayesian global vector autoregressions. European Economic Review, 81:167-188.

Feldkircher, M. and Korhonen, I. (2014). The rise of china and its implications for emerging markets-evidence from a gvar model. *Pacific Economic Review*, 19(1):61-89.

Friedman, M. (1953). *The methodology of positive economics*. University of Chicago Press.

Green, K. C. and Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence. *Journal of Business Research*, 68(8):1678-1685.

Gross, M. (2018). Estimating gvar weight matrices. *Spatial Economic Analysis*, pages 1-22.

Harbo, I., Johansen, S., Nielsen, B., and Rahbek, A. (1998). Asymptotic inference on cointegrating rank in partial systems. *Journal of Business & Economic Statistics*, 16(4):388-399.

Imbs, J. (2004). Trade, finance, specialization, and synchronization. *Review of Economics and Statistics*, 86(3):723-734.

Jannsen, N. (2010). National and international business cycle effects of housing crises. *Applied Economics Quarterly*, 56(2):175-206.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica: *Journal of the Econometric Society*, pages 1551-1580.

Johansen, S. (1992). Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of econometrics*, 52(3):389-402.

Kiguel, M. A. and Liviatan, N. (1995). Stopping three big inflations: Argentina, Brazil, and Peru. In *Reform, recovery, and growth: Latin America and the Middle East*, pages 369-414. University of Chicago Press.

Koehler, A. B. and Murphree, E. S. (1988). A comparison of the akaike and schwarz criteria for selecting model order. *Applied Statistics*, pages 187-195.

Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119-147.

Martin, F. and Cuaresma, J. C. (2017). Weighting schemes in global var modelling: a forecasting exercise. *Letters in Spatial and Resource Sciences*, 10(1):45-56.

Mohaddes, K. and Raissi, M. (2018). *Compilation, Revision and Updating of the Global VAR (GVAR) Database, 1979Q2-2016Q4*. University of Cambridge: Faculty of Economics (mimeo).

Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69(6):1519-1554.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association,* 84(405):223-230.

Park, H. J. and Fuller, W. A. (1995). Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis*, 16(4):415-429.

Pesaran, M. H. (2015). *Time series and panel data econometrics*. Oxford University Press.

Pesaran, M. H., Schuermann, T., and Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business & Economic Statistics*, 22(2):129-162.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2000). Structural analysis of vector error correction models with exogenous i (1) variables. *Journal of Econometrics*, 97(2):293-343.

Ploberger, W. and Kramer, W. (1992). The cusum test with ols residuals. *Econometrica: Journal of the Econometric Society*, pages 271-285.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American statistical Association*, 55(290):324-330.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, pages 1-48.

Sun, M. Y., Heinz, M. F. F., and Ho, G. (2013). Cross-country linkages in Europe: A global var analysis. Working Paper Series 13-194, International Monetary Fund.

Vansteenkiste, I. (2007). Regional housing market spillovers in the US: lessons from regional divergences in a common monetary policy setting. Working Paper Series 708, European Central Bank.

Vansteenkiste, I. and Hiebert, P. (2011). Do house price developments spillover across euro area countries? evidence from a global var. *Journal of Housing Economics*, 20(4):299-314.

Verick, S. and Islam, I. (2010). The great recession of 2008-2009: causes, consequences and policy responses. Discussion Paper Series 4934, IZA Institute of Labor Economics.