

A functional coefficient model view of the saving-investment relation

als Inaugural-Dissertation
zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

M.Sc. Fang Xu
geb. in JiangXi / China

Kiel, 2008

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan: Professor Dr. Helmut Herwartz
Erstberichterstattender: Professor Dr. Helmut Herwartz
Zweitberichterstattender: Professor Dr. Thomas Lux

Tag der Abgabe der Arbeit: 24. September 2007
Tag der mündlichen Prüfung: 19. Dezember 2007

Acknowledgements

This thesis cannot come into being without the guidance of my supervisors, help of friends and support from my family. I would like to express my deepest gratitude to my first supervisor, Prof. Dr. Helmut Herwartz, for his excellent direction, cooperation, and providing me a pleasant atmosphere for doing research. I am also very grateful to my second supervisor, Prof. Dr. Thomas Lux, for introducing me to dynamic macroeconomics and applications of physics methods in economics.

While writing up this dissertation, I have also benefited from discussions with colleagues at the Institute of Statistic and Econometrics at Christian-Albrechts-University of Kiel. Thanks especially go to Prof. Dr. Gerd Hansen, Prof. Dr. Roman Liesenfeld and PD Dr. Uwe Jensen for helpful comments and suggestions. Moreover, I would like to thank Vasyl Golosnoy, Christian Aßmann, Jens Högrefe, Guilherme Valle Moura and Albrecht Mengel for their constant supports.

To my great-granduncle, 徐梵澄 (Xu, FanCheng), I am greatly indebted. Although he has passed away shortly after I began my study in Germany, it is his scientific attitude and principle of life which have accompanied me on my way. This thesis is dedicated to him.

Kiel, April 2008

徐芳 (Xu, Fang)

Contents

| | |
|---|------------|
| List of Tables | V |
| List of Figures | VII |
| 1 Introduction | 1 |
| 1.1 Object of the study | 1 |
| 1.2 Outline of the study | 6 |
| 2 Literature review | 9 |
| 2.1 Economic models explaining a high SI relation | 13 |
| 2.2 Econometric approaches to measure the SI relation | 15 |
| 2.2.1 Sample selection | 15 |
| 2.2.2 Competing panel data models | 17 |
| 3 Data description | 23 |
| 3.1 Eight cross-sections | 24 |
| 3.2 Variables (factors) influencing SI relations | 25 |
| 3.A List of countries | 27 |
| 3.B List of factors | 28 |
| 4 Testing the bounded unit root | 31 |
| 4.1 Methodology | 34 |
| 4.2 Generating critical values | 36 |

| | | |
|----------|--|-----------|
| 4.3 | Results | 39 |
| 4.4 | Summary | 47 |
| 5 | Model selection | 49 |
| 5.1 | Model specifications | 50 |
| 5.2 | Cross validation criteria | 52 |
| 5.3 | Results from model comparisons | 53 |
| 5.3.1 | Static panel models | 55 |
| 5.3.2 | Static vs. weakly dynamic models | 55 |
| 5.3.3 | Weakly dynamic vs. error correction dynamics | 56 |
| 5.4 | Summary | 57 |
| 6 | Preliminary analysis | 59 |
| 6.1 | Between regressions | 59 |
| 6.2 | Pooled regressions | 64 |
| 6.3 | Time dependent SI relations | 66 |
| 6.4 | Cross-section dependent SI relations | 68 |
| 6.5 | Factor dependence - profiles of estimated SI relations | 70 |
| 6.6 | Summary | 73 |
| 7 | Functional coefficient model | 75 |
| 7.1 | Representation and assumptions for the parametric case | 77 |
| 7.2 | Semiparametric case | 81 |
| 7.3 | Bootstrap procedures | 82 |
| 7.3.1 | Residual based, pairs and wild bootstrap | 82 |
| 7.3.2 | Factor based bootstrap | 85 |
| 7.4 | Monte-Carlo analysis | 91 |
| 7.4.1 | Simulation design | 91 |
| 7.4.2 | Size features | 92 |
| 7.4.3 | Power features | 94 |

| | | |
|----------|---|------------|
| 7.5 | An empirical example | 96 |
| 7.6 | Summary | 100 |
| 8 | Application of FCM to SI relations | 101 |
| 8.1 | Time specific regression | 102 |
| 8.2 | Cross-section specific regression | 103 |
| 8.3 | Bivariate state dependent model | 105 |
| 8.4 | Factor dependent model comparison | 107 |
| 8.5 | Implementation | 109 |
| 8.6 | Inference | 112 |
| 8.7 | Results for functional estimation | 114 |
| 8.7.1 | Factors impacting on saving or investment | 114 |
| 8.7.2 | Factors measuring integration of markets | 123 |
| 8.7.3 | Large country effect | 124 |
| 8.7.4 | Implications | 125 |
| 8.8 | Summary | 126 |
| 9 | Concluding remarks | 127 |
| | Bibliography | 131 |
| | Index of Notation | 145 |

List of Tables

| | | |
|-----|--|-----|
| 4.1 | Testing current account for (bounded) nonstationarity | 40 |
| 4.2 | Panel unit root tests for (bounded) integration | 43 |
| 4.3 | (Bounded) unit root tests for saving and investment | 45 |
| 4.4 | Further (bounded) unit root tests | 46 |
| 5.1 | Panel model comparison | 54 |
| 6.1 | Between regression | 60 |
| 6.2 | Pooled regression | 65 |
| 6.3 | Cross-section specific regression | 69 |
| 6.4 | Factor dependence of the SI relation | 71 |
| 7.1 | Size estimates | 93 |
| 7.2 | Size adjusted power estimates | 95 |
| 7.3 | Tests for the factor independent SI relation in between regres- sions | 98 |
| 8.1 | Factor dependent model comparison | 108 |

List of Figures

| | | |
|-----|---|-----|
| 2.1 | Feldstein-Horioka puzzle and it's provoked research | 12 |
| 4.1 | Current account imbalances | 41 |
| 6.1 | Average saving vs. investment ratios | 61 |
| 6.2 | Average saving vs. investment ratios for 16 OECD countries . | 62 |
| 6.3 | Estimated time varying SI relations | 67 |
| 6.4 | Estimated cross-section specific SI relations | 68 |
| 7.1 | Functional estimates of the SI relation in between regressions . | 99 |
| 8.1 | Dynamics of openness ratio | 111 |
| 8.2 | Estimated SI relations conditional on age dependency ratio . | 115 |
| 8.3 | Estimated SI relations conditional on per capita income, pop- ulation growth and real interest rate differential | 118 |
| 8.4 | Estimated SI relations conditional on fiscal variables | 121 |
| 8.5 | Estimated SI relations conditional on openness ratio | 122 |
| 8.6 | Estimated SI relations conditional on logarithm of GDP | 124 |

Chapter 1

Introduction

1.1 Object of the study

Although international macroeconomics has undergone an intensive progress in the last half century, it appears to be a field which is embedded by perplexing puzzles. As main industrial economies have become open to foreign trade since 1950s, and removed capital controls since 1970s, there were believes that models with highly integrated goods and capital markets might be then a better approximation to reality, as can be seen in McKinnon (1981). However, there are many evidences on segmentations, which are likely related to the so called six major puzzles in international macroeconomics (Obstfeld and Rogoff 2000). If international goods markets are highly integrated, why do we observe home bias in trade (the home-bias-in-trade puzzle)? If international capital markets are highly integrated, why do we observe a high association between saving and investment (the Feldstein-Horioka puzzle), the home bias in equity portfolios (the home-bias portfolio puzzle), and not highly correlated consumption growth across OECD countries (the consumption correlations puzzle)? The last two of the six puzzles might be connected with imperfections in both domestic and

international goods markets. Precisely, why are real exchange rate movements highly persistent (the purchasing power parity puzzle), and why do exchange rates appear to be disconnected with the real economy (the exchange rate disconnect puzzle)?

This study is motivated by the Feldstein-Horioka puzzle (FH puzzle). By means of a between regression for OECD countries, Feldstein and Horioka (1980), henceforth FH (1980), document a strong correlation linking the domestic investment and saving measured as a share of the gross domestic product (GDP). This evidence is argued to be at odds with capital mobility. If capital is mobile, domestic savings could seek out the highest return internationally and the domestic investment can be financed also by foreign savings. Following FH (1980) one would expect that under perfect capital mobility the correlation between a country's saving and investment ratio should be small. The FH puzzle has provoked a lively discussion in both theoretical and empirical literature. The initial responses from replicating between regressions with different OECD countries for time period up to early 1990s confirm that the result in FH (1980) is quite robust. Two of the most investigated questions in the analysis which follows are: What does the saving-investment relation really signal and how should the saving-investment (SI) relation be measured? To answer the first question, various influences have been suggested in the theoretical literature such as population growth (Obstfeld 1986), the intertemporal budget constraint (Coakley, Kulasi and Smith 1996, Taylor 2002), output fluctuations in non-traded goods (Tesar 1993), or current account targeting (Artis and Bayoumi 1992). To answer the second question, numerous empirical specifications have been employed. Fundamentally they can be classified into three categories. Static models formalized to explain domestic investment ratios conditional on saving ratios constitute the first category. This framework comprises basic panel specifications such as time-dependent SI relations in-

vestigated by Sinn (1992) and country specific SI relations considered by Obstfeld (1986). A second class of models is given in terms of first differences of domestic saving and investment ratios (Feldstein 1983), which may be regarded as 'weakly dynamic'. More general dynamic patterns are formalized in a third set of empirical contributions comprising error correction models (ECMs), see Jansen and Schulze (1996) for example. By means of ECMs, the stationarity of current account (CA) suggested by the intertemporal budget constraint can be investigated. Since the CA balance can be approximately decomposed as domestic saving minus investment, a stationary CA might reflect a cointegration relation between saving and investment. Although most recent empirical investigations for the SI relation concentrate on this potential cointegration relation (e.g. Abbott and Vita 2003 and Ho 2002a), no unique evidence for a cointegration relation is found. Compared to the abundant empirical investigations for the intertemporal budget constraint, empirical contributions linking all other determinants of the SI relation are rather scarce. A few exceptions and identified factors are Summers (1988) (government budget balance), AmirKhalkhali, Dar and AmirKhalkhali (2003) (government size), Kasuga (2004) (financial structures) and Sachsida and Caetano (2000) (the variability between external and domestic saving). Nevertheless, in these contributions there exists no uniform model to measure the SI relation along with potential determinants.

The first contribution of this study is to develop a new semiparametric approach to analyse a wide range of determinants of the SI relation. This approach is derived as a bivariate generalization of functional coefficient models (Cai, Fan and Yao 2000, henceforth CF&Y), in which model parameters may change with the value of other variables (factors). To identify potential determinants of the SI relation the formalized semiparametric model is suitable to cope with cross-sectional heterogeneity, time and factor de-

pendence. It allows to separate deterministic from measurable economic conditions characterizing the empirical SI relation over time. And it allows to distinguish long and short run factor impacts on the SI relation.

Concerning the model classes, the static panel version of functional coefficient models is adopted. This choice is supported by the following three findings. Firstly, contrasting static and weakly dynamic model formalizations via cross validation (CV) techniques studied for example in Allen (1974), no hint at the necessity of a weakly dynamic model specification is found. Secondly, adding an error correction term in dynamic models does not improve model performance. Evidence for a cointegration relation between domestic saving and investment ratios is not found. A further supporting evidence on this view is related to the second contribution of this study, i.e. considering the 'bounded' nature of the CA imbalance measured as a fraction of GDP in unit root tests. Traditional unit root tests are formalized to distinguish between stationary processes and processes driven by stochastic trends that can grow or decrease to any level. Time-series as the CA imbalance to GDP are bounded not only by construction but also via potential policy controls. However, these bounds have not yet been taken into account when testing the unit root hypothesis in the literature. In this thesis, results from Phillips-Perron tests (Phillips 1987 and Phillips and Perron 1988) for unit roots and bounded unit roots are contrasted for 26 OECD members. While the former hint at stationarity of current account imbalances for 12 economies with 10% significance, the latter indicate bounded stationarity for only 6 economies. As such, neglecting bounds in unit root testing might induce misleading results. Moreover, as empirical findings in favor of nonstationary imbalances are often explained to mirror power deficiencies of unit root tests applied to short time spans of data, panel based test statistics are adopted. The corresponding results hint at bounded nonstationarity of current account imbalance at the aggregate level, and imply

that saving and investment are likely not cointegrated.

With regard to the inference on a significant influence of potential factors on the SI relation, the hypothesis that the functional coefficient formulating the SI relation is constant, i.e. factor independent, is considered. If the functional coefficient is modeled parametrically, the parametric form (constant function) is compared with another parametric alternative. Modeling the functional coefficient nonparametrically, the alternative is the semiparametric regression. In this area of hypothesis testing, F -type tests with bootstrap based inference is broadly used. For a goodness-of-fit test statistic based on comparing the residual sum of squares from parametric and semiparametric functional regressions, CF&Y advocate a residual based bootstrap approach. However, owing to possibly heterogenous error terms, residual based bootstrap inference could lack robustness. To deal with heteroskedasticity, pairs bootstrap (Freedman 1981) and wild bootstrap (Wu 1986, Liu 1988) can be applied. Härdle and Mammen (1993) consider the case of a test statistic that is based on the integrated squared difference between a parametric and a nonparametric estimate. They show that under the null hypothesis the pairs bootstrap based approximation does not consistently estimate the distribution of the test statistic while the wild bootstrap approximation does.

As the third contribution of this thesis, a new, factor based bootstrap approach is proposed to obtain critical values for contrasting the model estimates obtained under the null and the alternative hypothesis. Similar to pairs bootstrap, the new scheme does not require estimating model residuals that are subsequently used for resampling. Different from pairs bootstrap, the factors governing coefficient variation only under the alternative hypothesis are drawn with replacement. This approach is shown to cope with heteroskedasticity since the relationship between the variance of error terms and the corresponding regressors is retained. Modeling the functional coefficient parametrically, the factor based bootstrap approximation

of a F -type test statistic is shown to hold asymptotically. In simulation studies, for small samples factor based inference outperforms the wild bootstrap and the pairs bootstrap counterpart that are also known to cope with heteroskedasticity. Modeling the functional coefficient nonparametrically, only tests with factor based bootstrap provide correct empirical size estimates.

This study investigates various (partly overlapping) cross-sections characterizing the world economy, developing countries, the OECD, the EU and the Euro area. Applying annual data from 1971 to 2002 for these cross-sections, time decreasing and country specific estimates of the SI relation is found by means of static panel data models as between, pooled, cross-section, and time-series regressions. From functional coefficient models, an economy's degree of openness to foreign trade, its age dependency ratio and government current and consumption expenditures are identified to have a significantly negative influence on the SI relation in the long run. Besides, countries with high GDP (measuring the effect of country size) are more likely to have a high SI relation. According to these results, the interpretation of a high SI relation as a signal for low capital mobility in FH (1980) has to be treated with care. Empirically high SI relations could reflect goods market friction, demographic development or fiscal consolidation rather than being puzzling.

1.2 Outline of the study

This thesis is structured as follows. In Chapter 2 some core theoretical and empirical contributions provoked by FH (1980) are sketched briefly. The considered eight cross-sections and potential variables (factors) influencing the SI relation are introduced in Chapter 3.

Chapter 4 tests the (bounded) unit root hypothesis of the current account imbalances. Results from Phillips-Perron tests for unit roots and bounded

unit roots are contrasted for 26 OECD countries. For testing at an aggregate level, panel based test statistics are constructed as Fisher tests (Maddala and Wu 1999) and by means of the modified inverse normal method (Hartung 1999, Demetrescu, Hassler and Tarcolea 2006) with the latter allowing for dependence among single country test statistics.

Given heterogenous econometric models on the SI relation in the literature, Chapter 5 compares the relative merits of competing model classes. Firstly, between, pooled, time dependent and country dependent specifications of the SI relation in static models are compared systematically. Secondly, the scope of static and weakly dynamic models addressing the SI relation is distinguished. Thirdly, weakly dynamic models are contrasted against ECMs.

Heterogenous estimates of the SI relation along both time-series and cross-sectional dimension given our data sets are provided in Chapter 6. A preliminary view on the factor dependence of the SI relation is shown.

Chapter 7 introduces the functional coefficient model (FCM) with the factor based bootstrap inference. The properties of the residual based (RB), wild (WB), pairs (PB) and factor based bootstrap (FB) inference in FCMs with finite samples are compared by means of a simulation study. The Monte Carlo exercises cover both parametric and semiparametric alternative hypotheses. Presuming a parametric pattern for the functional coefficient, the asymptotic behavior of the test statistic and the corresponding factor based bootstrap approximation are investigated analytically.

In Chapter 8 the proposed bivariate state dependent FCM are introduced and the results are discussed. Finally, Chapter 9 summarizes the main findings and concludes.

Chapter 2

Literature review

FH (1980) argue that if capital were mobile, domestic investment would not depend on domestic saving but the world saving, and domestic saving could seek out globally the highest return. Under perfect mobility of capital one would expect a small if any correlation between domestic saving and investment. By contrast, immobility of capital would imply a one-to-one relationship between domestic saving and investment.

To investigate the saving-investment relation FH (1980) make use of a between regression

$$\bar{I}_i = a + b\bar{S}_i + u_i, \quad i = 1 \dots N, \quad (2.1)$$

where $\bar{I}_i = 1/T \sum_{t=1}^T I_{it}$ and $\bar{S}_i = 1/T \sum_{t=1}^T S_{it}$, with I_{it} and S_{it} , $t = 1, \dots, T$, denoting gross domestic investment to GDP and gross domestic saving to GDP in time period t and country i , respectively. Estimating regression (2.1) for 16 OECD countries with annual data from 1960 to 1974, FH (1980) obtain

$$\bar{I}_i = 0.035 + 0.887\bar{S}_i + \hat{u}_i, \quad (2.2)$$

(0.018) (0.074)

with standard errors given in parentheses underneath the parameter estimates. The estimated impact of the average saving ratio on the average investment ratio, 0.887, is not significantly different from unity. The degree of

explanation offered by (2.2) is 0.91. FH (1980) interpret the high SI relation as evidence for capital market segmentation among the considered OECD economies.

A potential methodological problem related to the estimation of equation (2.1) is the endogeneity of the saving ratio. However, several authors (for instance, Feldstein and Horioka 1980 and Frankel 1985) show that the high SI relation stands by using instrument variables for the estimation. Furthermore, analyses replicating between regressions with different samples of OECD countries have shown that the result in FH (1980) is quite robust. Although the estimated SI relation has declined over time, it remained large and significantly different from zero.

The academic debates triggered by the FH puzzle are enormous. They can be divided into four categories roughly. Figure 2.1 illustrates the coherence between the FH puzzle and these four categories of discussions. Firstly, the FH puzzle is based on an 'empirical' investigation of the SI relation. As such, different sample information and econometric models can be applied. The former involves members of the EU, the Euro area, large economies, and the less developed countries. The latter includes static panel data models as between, pooled, time, and cross-section specific regressions (e.g. Sinn 1992 and Obstfeld 1986); weakly dynamic models via regressing the first difference of investment on the first difference of saving (for example Feldstein 1983 and Feldstein and Bacchetta 1991); and ECMs (e.g. Abbott and Vita 2003 and Jansen 1996). Secondly, the interpretation of the observed high SI relation as a signal for low capital mobility is debatable. The high SI relation can be caused by common factors as productivity shocks and population growth (Obstfeld 1986), intertemporal budget constraints (e.g. Coakley, Kulas and Smith 1996 and Taylor 2002), government targeting current account imbalances (for example Artis and Bayoumi 1992 and Summers 1988), and the low integration of international goods markets (e.g. Obstfeld and Ro-

goff 2000 and Wong 1990). Thirdly, by accepting the FH's interpretation for the high SI relation, some authors try to explain the low international capital mobility. Transaction costs, information constraints and exchange rate risk are proposed by Feldstein and Horioka (1980) and Feldstein and Bachetta (1991) as possible grounds. The low capital mobility is also thought to be in line with the not binding real interest rate parity (Frankel 1992), the consumption correlations puzzle (Backus, Kehoe and Kydland 1992), and the home-bias portfolio puzzle (French and Poterba 1991). At last, since FH (1980) intend to test the degree of the capital market integration, other measures for it are considered. These consist of interest rate parities (e.g. Frankel 1992 and Lemmen and Eijffinger 1995), the variance of the current account (Ghosh 1995), the degree of short-run divergence in the impulse responses of saving and investment to shocks (Moreno 1997), the adjustment speed of the current account to its equilibrium value (Taylor 2002), the error correction coefficient in VECMs of saving and investment (e.g. Jansen 1996 and Jansen and Schulze 1996), and a measure of the speed of error correction in VECMs (Hoffmann 2004).

This chapter¹ is devoted to sketch some of the leading research related to the FH puzzle in the first two categories. For a detailed and relatively recent review over the literature on the FH puzzle the reader may consult Coakley, Kulasi and Smith (1998).

¹For a different version of this chapter see "Panel data model comparison for empirical saving-investment relations", with Helmut Herwartz, forthcoming in *Applied Economic Letters*, and "A functional coefficient model view of the Feldstein-Horioka puzzle", with Helmut Herwartz, University Kiel, Economics Working Paper 2007-14.

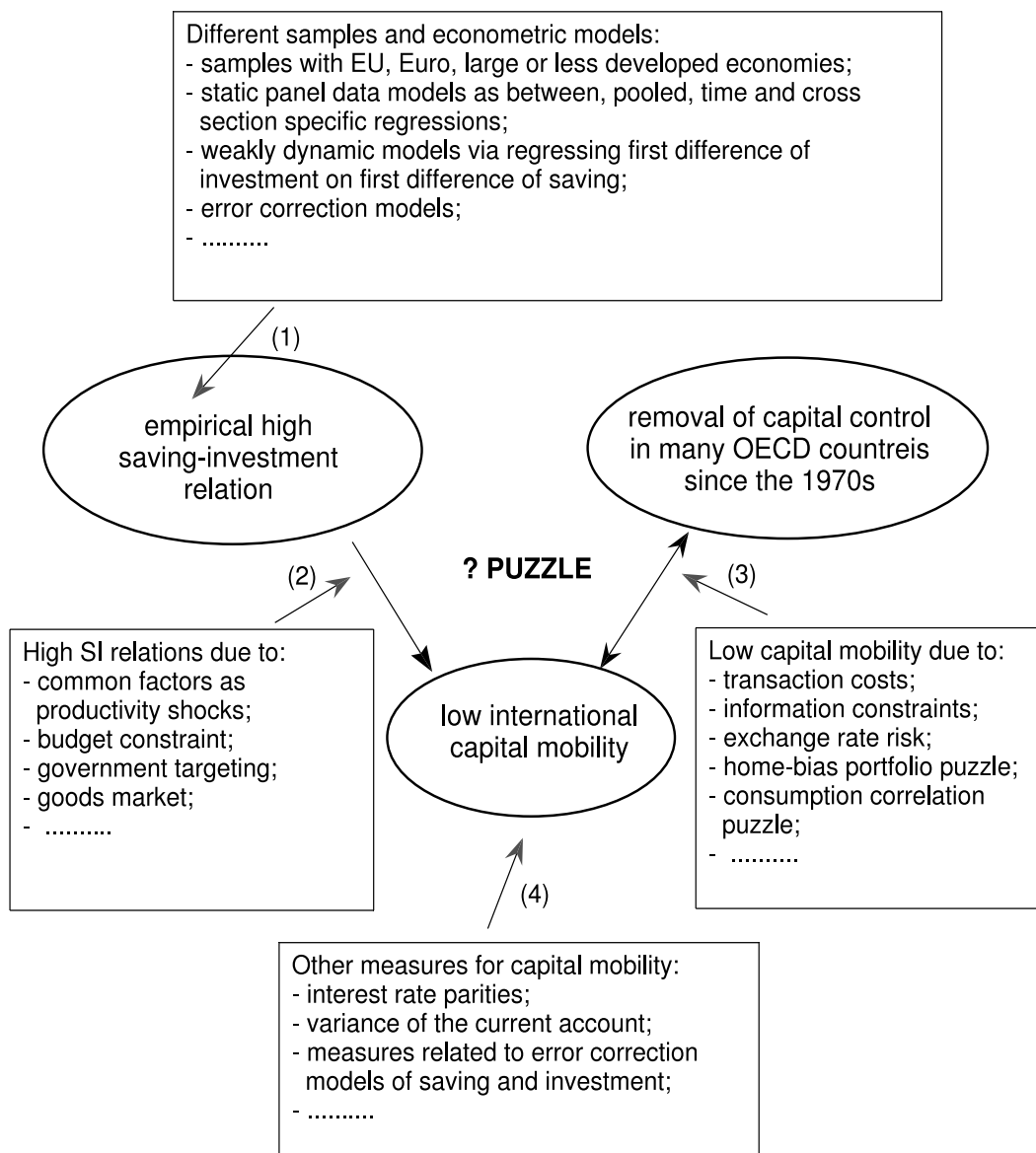


Figure 2.1: Feldstein-Horioka puzzle and it's provoked research.

2.1 Economic models explaining a high SI relation

From the viewpoint of economic theory, firstly, general equilibrium models have been constructed to allow a high SI relation in response to exogenous shocks under high or perfect capital mobility. By means of a life-cycle model Obstfeld (1986) demonstrates that, given a rise in the population growth rate, both the saving and the investment ratio increase. Mendoza (1991) constructs a real-business-cycle model of a small open economy with moderate adjustment costs and small variability and persistence of technological shocks. This model turns out to be consistent with a positive correlation between domestic saving and investment although financial capital is perfectly mobile. Besides, since a low negative relation between CA and investment is in line with a high SI relation, studies about the former provide alternative views for the latter. Glick and Rogoff (1995) find that a linear quadratic intertemporal model with costs of adjustment in investment performs well in explaining CA and investment changes for major industrialized countries for the period 1961 to 1990. They argue that a low negative CA-investment relation may be caused by global productivity shocks and the under-response of consumption to country-specific productivity shocks.

Secondly, the stationarity of the current account balance implied by the long run intertemporal budget constraint (solvency constraint) could induce a high SI relation. By construction, saving minus investment equals the current account balance.² Thus, a close SI relation might reflect stationarity of the current account. Introducing a market determined risk premium on borrowing, Coakley, Kulasi and Smith (1996) show that the long run sol-

²Precisely, the current account balance is equal to the national saving minus the domestic investment. Although it is the domestic saving considered by FH (1980), the difference between these two saving measures is mostly ignored in the related literature.

vency constraint implies a stationary current account. For the case of a simple Solovian economy with stochastic growth, Taylor (2002) demonstrates that stationarity of the current account is a sufficient condition for the long run intertemporal budget constraint to hold. In this vein of economic models a high SI association reflects the solvency constraint, but not necessarily capital immobility. However, the empirical evidence on the stationarity of current account balances is weak or mixed.

In the third place a high SI relation may be due to the government targeting the current account balance. Summers (1988) suggests that the government may offset the difference between private saving and investment by the budget deficit so that large sustained capital flows are not observed. Supporting evidence is found in a between regression of budget deficits on private saving-investment gap for 14 industrialized countries for the period 1973 to 1980. Artis and Bayoumi (1992) also argue that the current account balance was an important target for monetary policy in the 1970s, but not in the 1980s. This policy change appears to correspond to a reduction in the SI relation among OECD countries in the 1980s.

Furthermore, the goods market, not the capital market, may be seen as the binding constraint linking domestic saving and investment. From this perspective, Wong (1990) illustrates that an increase in the saving ratio is accompanied by an increase in the investment ratio as long as non-traded goods are consumed by domestic residents. Tesar (1993) also demonstrates that a model with stochastic fluctuations in the output of non-traded goods is consistent with a high SI association. In case non-traded goods account for a significant share of total output, consumer preferences over traded and non-traded goods and over the intertemporal allocation of consumption may introduce low cross-country correlations of aggregate consumption and an optimal portfolio biased towards claims on domestic output. Describing the so-called consumption correlations puzzle (Backus et al.

1992) and the home-bias portfolio puzzle (French and Poterba 1991), the latter effects are in line with a high SI relation. Comparably, Obstfeld and Rogoff (2000) demonstrate that moderate transactions costs of international trade may cause a substantial difference in real interest rates in spite of full financial market integration. In turn, real interest rate differentials might give rise to a high SI relation.

Finally, government size (AmirKhalkhali et al. 2003), financial structures (Kasuga 2004), and the variability between external and domestic savings (Sachsida and Caetano 2000) are also proposed as possible factors influencing the SI relation empirically.

2.2 Econometric approaches to measure the SI relation

Econometric attempts to solve the FH puzzle might be divided in two groups, namely the use of different sample information and of alternative econometric model specifications. In the following these two groups are sketched briefly.

2.2.1 Sample selection

Contrasting large economies with the rest in OECD countries or developing with OECD countries, the so called large country effect can be observed. Harberger (1980), Murphy (1984) and Obstfeld (1986) show empirically that large countries are likely to have high SI relations. Using a quantitatively restricted equilibrium model with perfect capital mobility, Baxter and Crucini (1993) demonstrate the large country effect theoretically. For a large economy, the world interest rate and many goods prices are more likely endogenous. Then, a shortfall in domestic saving may drive up both the world's

as well as the domestic interest rate. As a result, a large countries' domestic investment decreases. Thus, although capital flows are mobile for the large country, it is likely to show a high SI relation. In contrast, most developing countries are small and cannot influence the world interest rate. Therefore, the corresponding SI relation is lower for developing countries. Murphy (1984) demonstrates that between regression estimates reduce to 0.59 for 10 small OECD countries, and remain as high as 0.98 for 7 large OECD economies. It turns out that particularly the US, Japan and the UK have a dominant impact on the between estimate. By means of time series models for 7 OECD countries Obstfeld (1986) also demonstrates that the measured SI relation is increasing in country size. Focussing on the difference between the saving and investment ratio, Harberger (1980) shows that this difference as a fraction of the investment ratio has a lower absolute value and less variability for OECD countries in comparison with developing economies. As the opposite to the large country effect, Dooley, Frankel and Mathieson (1987) and Mamingi (1994) have found that the SI coefficient is smaller for developing economies in comparison with OECD countries. Dooley et al. (1987) show that between regression estimates are smaller for 48 developing economies than for 14 OECD countries. Using time series data for 58 developing countries, Mamingi (1994) obtains an estimated SI relation which is smaller than the corresponding OECD based measure.

Moreover, the SI relation is found to be lower among members of the EU or the Euro area. Owing to informational and institutional links, financial flows should be larger within the EU than among OECD countries. Feldstein and Bachetta (1991) show that 9 EU countries experienced a sharp decline in the SI relation in the 1980s, while 14 non-EU OECD countries did not. Similarly, Artis and Bayoumi (1992) find for the 6 core economies of the European Monetary System an insignificant SI relation over the period 1981 to 1988. Blanchard and Giavazzi (2002) document that the SI relation

estimated from pooled regression models declines in case the investigated cross-section changes from OECD to the EU or the Euro area. In addition, it is diagnosed to decline over time. According to Blanchard and Giavazzi (2002) the SI relation for the Euro area diminishes to 0.14 when using annual data over the period 1991 to 2001.

2.2.2 Competing panel data models

Basically empirical models can be classified into three categories. Static models comprising basic panel specifications formalized to explain domestic investment ratios conditional on saving ratios constitute the first category. A second class of models is given in terms of first differences of domestic saving and investment ratios, which may be regarded as 'weakly dynamic'. More general dynamic patterns are formalized in a third set of empirical contributions comprising ECMs.

Static panel models

Static panel data models consist of between, pooled, time and cross-section specific regressions. Proceeding from an equilibrium model of saving, investment, net foreign investment and the real domestic interest rate, Feldstein (1983) argues that estimates of the SI coefficient from between regressions provide a reliable basis to evaluate the hypothesis of perfect international capital mobility. Murphy (1984), Obstfeld (1986), Feldstein and Bachetta (1991) and Tesar (1991) estimate the SI relation via between regressions. Between regressions are typically understood to address the long run SI relation, which is not affected by the business cycle. However, between regressions might deliver biased results against capital mobility observing that the long run SI relation could be determined by the intertemporal budget constraint, as pointed out by Sinn (1992). For this reason Sinn (1992) es-

timates time dependent SI relations from cross-sectional regressions. Nevertheless, the evidence offered by time varying SI relations for 23 OECD countries over a sample period from 1960 to 1988 does not overcome the finding of a puzzling high SI relation. Furthermore, Miller (1988), Afxentiou and Serletis (1993) and Alexakis and Apergis (1994) have argued for cross-section specific regressions which are to be preferred in the light of potential cointegration linking domestic saving and investment. In case saving or investment ratios were nonstationary it is unclear what cross-sectional averages entering a between regression actually measure. Another common argument for a cross-section specification is that the SI relation is likely heterogenous across economies. In case of cross-sectional heterogeneity between regressions have attached the risk of providing biased results owing to model misspecification. Obstfeld (1986), Miller (1988), Afxentiou and Serletis (1993), Tesar (1993) and Alexakis and Apergis (1994) also evaluate country specific SI relations.

In addition, several different approaches are suggested to consider the cross-sectional heterogeneity. Corbin (2001) shows that the fixed effect and random effect estimator of the SI relation are smaller in comparison with the pooled and between estimator. Panel data models with fixed effects for OECD countries are also considered by Krol (1996) and Jansen (2000). Comparing their results shows that the inclusion of Luxembourg in the sample provides a low estimate for SI relation. Furthermore, Amirkhalkhali and Dar (1993) permit inter-country variation in both the constant and the slope parameter in panel regressions, which are estimated by means of error component models (Swamy 1970, Swamy and Mehta 1975). Moreover, using mean group estimates (Pesaran and Smith 1995) in a nonstationary and heterogeneous panel, Coakley, Fuertes and Spagnolo (2004) obtain an estimated SI relation which is insignificantly different from zero for 12 OECD countries over the period from 1980 to 2001. Following them, Payne and Ku-

mazawa (2006) obtain a mean group estimator of the SI relation being lower than the corresponding between estimator for 47 developing countries.

Summarizing the panel based responses to the initial contribution by FH (1980) it turns out that the FH puzzle is quite robust over a substantial portfolio of applied static panel data models. Comparisons of alternative panel data modeling frameworks, however, are rare and if available, not very comprehensive or systematic and based on in-sample fitting criteria.

Weakly dynamic models

Motivated by accounting for potentially country specific constant terms in a pooled regression of investment on saving, Feldstein (1983) suggests to regress the first difference of investment ratios on the first difference of saving ratios. As such, the country specific constant term is eliminated in this weakly dynamic regression. To explore the time-series correlations between saving and investment, Bayoumi (1989) adopts also the weakly dynamic regression of investment on saving, but for each country separately.

Error correction models

Due to the intertemporal budget constraint the current account shall be stationary, as argued by Coakley et al. (1996) and Taylor (2002). Most recent empirical investigations for the SI relation are concentrating on a potential cointegration relation between saving and investment.³ Thereby, error correction models are suggested as the suitable framework to measure the SI relation.⁴ However, no consensus on a potential cointegration relation be-

³Abbott and Vita (2003), Coakley and Kulasi (1997), Gulley (1992), Ho (2002a), Ho (2002b), Leachman (1991), Lemmen and Eijffinger (1995), Haan and Siermann (1994), Miller (1988), Vita and Abbott (2002).

⁴Bajo-Rubio (1998), Coiteux and Olivier (2000), Jansen (1996), Jansen and Schulze (1996), Jansen (1997), Jansen (1998), Moreno (1997), Özmen and Parmaksiz (2003b), Özmen and

tween saving and investment has been achieved yet.

It is firstly suggested by Feldstein and Bachetta (1991) to consider a simple adjustment process by which the change in investment varies with the difference between investment and saving in the previous period. Since Jansen (1996) or Jansen and Schulze (1996) has introduced a rather complete ECM to measure the capital mobility, this vein of modeling becomes popular. Conditional on ΔS_{it} , a single-equation ECM for the SI relation can be defined as:

$$\Delta I_{it} = \delta_i + \alpha_i(I_{i,t-1} - \eta_i S_{i,t-1}) + \Gamma_i \Delta S_{it} + e_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad (2.3)$$

where Δ denotes the first difference operator, e.g. $\Delta I_{it} = I_{it} - I_{i,t-1}$. Three ways of identifying capital mobility are proposed by Jansen (1996): not cointegrated saving and investment ratios, a non-stationary current account, and a low value of the short-run SI relation. These three measures can be reflected by the parameters in (2.3). Firstly, evidence for capital mobility can be constituted if $\alpha_i = 0$ in equation (2.3). In this case the saving and investment ratio are not cointegrated, and thus not correlated in the long run. When $\alpha_i \neq 0$, the second way to investigate the capital mobility is to inspect the value of η_i . If $\alpha_i \neq 0$ but $\eta_i \neq 1$, then the saving and investment ratio are cointegrated with vector $(1, \eta_i)'$. In other words, $I_{i,t-1} - \eta_i S_{i,t-1}$ but not $I_{i,t-1} - S_{i,t-1}$ is stationary. As such, the current account $(-I_{i,t-1} + S_{i,t-1})$ may not be stationary, and supporting evidence for capital mobility is found. Thirdly, in case that $\alpha_i \neq 0$ and $\eta_i = 1$ but Γ_i has a low value, existence of capital mobility is also identified. A low short run correlation between the saving and investment ratio is only possible if capital is sufficiently mobile.

Similarly, Pelgrin and Schich (2004) interpret the error correction coefficient, α_i , as an indicator of capital mobility. They view capital mobility as

Parmaksiz (2003a), Pelagidis and Mastrogiannis (2003), Pelgrin and Schich (2004), Schmidt (2001), Sinha and Sinha (2004), Taylor (1996), Taylor (1998).

the ease with which a country can borrow or lend to run prolonged current account imbalances in the short to medium term. Thus, the higher the capital mobility, the lower is the adjustment speed of investment to its long run equilibrium level implied by the one-to-one cointegrating relation linking S_{it} and I_{it} . Implementing a panel ECM for 20 OECD countries over the sample period 1960 to 1999 with three alternative specifications of cross-sectional heterogeneity (dynamic fixed effects, mean group and pooled mean group estimation) Pelgrin and Schich (2004) find that the estimated error correction coefficient, $\hat{\alpha}_i$, is negative and significantly different from zero. A time dependent evaluation reveals that $\hat{\alpha}_i$ comes closer to zero over time, which is consistent with a presumption of increasing capital mobility. Furthermore, the estimated cointegration parameter, $\hat{\eta}_i$, is found to differ only insignificantly from unity, thereby implying a binding long run solvency constraint.

Regarding the ECM specification in (2.3) it is worthwhile to point out that the conditional single equation ECM only offers efficient estimation or inference in case domestic saving is weakly exogenous, i.e. it does not respond to lagged current account imbalances (Johansen 1992). Weak exogeneity of S_{it} is, however, neither tested by Jansen (1996) nor by Pelgrin and Schich (2004). As a more fundamental caveat of cointegration modeling, it is important to mention that standard unit root tests are not constructed for variables which are bounded by construction, as e.g. $S_{i,t-1}$ or $I_{i,t-1}$. Unit root tests are formalized to distinguish between stationary processes and processes driven by stochastic trends. Since the latter can grow or decrease to any level, the notion of nonstationary saving and investment ratios is to some extent counterintuitive. This issue will be addressed in Chapter 4.

Chapter 3

Data description

In the empirical literature of the SI relation, most authors concentrate on one or two specific cross-sections such as OECD members, EU countries, the Euro area, large or less developed economies. This study investigates a set of specific cross-sections, and a general cross-section sampled from all over the world and containing as many economies as possible conditional on data availability. The latter is one of the largest cross-sections that has been considered to analyse the SI relation. Distinguishing numerous specific cross-sections will be useful to reconsider former analyses relating the SI relation e.g. to the degree of market integration or the state of development. The large cross-section promises a global view on descriptive features of the correlation between domestic saving and investment and its underlying determinants. The SI relation is investigated with eight alternative (partly overlapping) cross-sections using annual data from 1971 to 2002 drawn from the World Development Indicators CD-Rom 2004 published by the World Bank¹. Section 3.1 introduces the eight cross-sections. A list of considered countries is provided in the Appendix 3.A. The investigated economic factors is presented in Section 3.2, and their definitions are provided

¹Only for Chapter 4 the sample comprises yearly observations over the period 1971 to 2004

in Appendix 3.B.

3.1 Eight cross-sections

The first and most comprehensive sample covers 97 countries from all over the world (**W97**), for which most observations of the saving and investment ratio from 1971 to 2002 are available. For 6 countries (Canada, Fiji, Ireland, Japan, Myanmar, and Portugal) data for 2002 are not available. These missing values are estimated by means of univariate autoregressive models of order 1 with intercept. Although data for Sao Tome and Principe and Lesotho are published, these two countries are not included owing to an outstandingly high negative saving ratio prevailing over quite a long period. The second cross-section consists of 29 developed countries and is denoted as **D29**. It contains Hong Kong, Singapore and all OECD countries except Czech Republic, Poland, Slovak Republic. The latter three countries are not included due to data nonavailability. The remaining economies contained in W97 but not in D29 is then collected in a third cross-section describing the less developed economies and is denoted as **L68**.

Since FH puzzle is observed mostly with OECD countries, a fourth cross-section (**O26**) consisting of 26 OECD countries is considered. They are D29 minus Luxembourg, Hong Kong and Singapore. Luxembourg is often excluded in empirical analyses of the SI relation owing to presumably peculiar determinants of its savings. As the fifth cross-section 11 Euro area economies excluding Luxembourg (**E11**) are investigated. In the Euro area, there is no exchange rate risk and financial markets should be highly integrated. To offer a 'complementary' view at the link between market integration and the SI relation, a sixth cross-section **O15** defined as O26 minus E11 is considered. Here we focus on weaker forms of market integration and try to isolate their impact on the SI relation.

In addition, a sample covers 14 major countries of the European Union (**E14**) is considered, which are the O26 countries without Australia, Canada, Hungary, Iceland, Japan, Korea, Mexico, New Zealand, Norway, Switzerland, Turkey, and the US. Contrasting this subgroup with O26 may reflect the EU effect on the SI relation. Finally, for completeness and to improve the comparability of our results to FH (1980), the cross-section employed in their initial contribution (**F16**) is also studied. The latter comprises 16 OECD countries namely O26 excluding France, Hungary, Korea, Mexico, Norway, Portugal, Spain, Switzerland and Turkey.

3.2 Variables (factors) influencing SI relations

The choice of potential factor variables is relied on recent studies by Edwards (1995), Debelle and Faruquee (1996), Milesi-Ferretti and Razin (1997), Milesi-Ferretti and Razin (1998), Masson, Bayoumi and Samiei (1998), and Chinn and Prasad (2003).

The employed factor variables allow a classification into three groups. The first group includes long-run variables presenting the the state of development, demographics, and macroeconomic policies. The choice of variables in this group is guided by economic theories of saving and investment. The state of development is measured as GDP per capita (**GDPC**). Demographic variables include the ratio of the dependent population to the working-age population (**AGE**) and the population growth rate (**POPG**). The role of fiscal policies is measured via government budget balances as a share of GDP (**GVBB**), government total expenditures as a share of GDP (**GVTT**), government capital expenditures as a share of GDP (**GVIVM**), government current expenditures as a share of GDP (**GVCE**), and government consumption expenditures as a share of GDP (**GVCON**). Variables describing the integration of goods and financial markets compose the sec-

ond group. It includes the openness ratio measured as the ratio of export plus import to GDP (**OPN**), the ratio of exports to GDP (**EXPT**), the ratio of imports to GDP (**IMPT**), and the absolute real interest rate differential (**INTD**). At last, the scope of variables measuring the dependence of the SI relation on country size is addressed by natural logarithm of GDP (**LGDP**).

Regarding **INTD**, it is measured for a particular economy towards a world real interest rate index. Country specific real interest rates are the lending rates charged by banks on loans to prime customers adjusted for inflation. A GDP weighted average of real interest rates among the US, Germany and Japan is constructed to approximate the real world interest rate. Instead of using the interest rate differential directly we presume that positive and negative realizations are equally informative for the prevalence of capital market frictions. Therefore the impact of the absolute real interest differential on the SI relation is investigated.

3.A List of countries

- W97: Algeria; Argentina; Australia; Austria; Bangladesh; Barbados; Belgium; Benin; Botswana; Brazil; Burkina Faso; Burundi; Cameroon; Canada; Central African Republic; Chile; China; Colombia; Congo Dem. Rep.; Congo Rep.; Costa Rica; Ivory Coast; Denmark; Dominican Republic; Ecuador; Egypt Arab Rep.; El Salvador; Fiji; Finland; France; Gabon; Gambia; Germany; Ghana; Greece; Guatemala; Guyana; Haiti; Honduras; Hong Kong, China; Hungary; Iceland; India; Indonesia; Ireland; Israel; Italy; Jamaica; Japan; Kenya; Korea, Rep.; Kuwait; Luxembourg; Madagascar; Malawi; Malaysia; Mali; Malta; Mauritania; Mexico; Morocco; Myanmar; Nepal; Netherlands; New Zealand; Niger; Nigeria; Norway; Pakistan; Paraguay; Peru; Philippines; Portugal; Rwanda; Saudi Arabia; Senegal; Singapore; South Africa; Spain; Sri Lanka; Suriname; Swaziland; Sweden; Switzerland; Syrian Arab Republic; Thailand; Togo; Trinidad and Tobago; Tunisia; Turkey; Uganda; United Kingdom; United States; Uruguay; Bolivarian Republic of Venezuela; Zambia; Zimbabwe.
- L68: W97 minus D29.
- D29: Australia; Austria Belgium; Canada; Denmark; Finland; France; Germany; Greece; Hong Kong, China; Hungary; Ireland; Italy; Japan; Korea; Luxembourg; Mexico; Netherlands; New Zealand; Norway; Portugal; Singapore; Spain; Sweden; Switzerland; Turkey; U.K.; U.S.
- O26: D29 minus Luxembourg; Hong Kong, China; Singapore.
- O15: O26 minus E11.
- E11: Austria; Belgium; Finland; France; Germany; Greece; Ireland; Italy; Netherlands; Portugal; Spain.

- F16: Australia; Austria; Belgium; Canada; Denmark; Finland; Germany; Greece; Ireland; Italy; Japan; Netherlands; New Zealand; Sweden; U.K.; U.S.
- E14: Austria; Belgium; Denmark; Finland; France; Germany; Greece; Ireland; Italy; Netherlands; Portugal; Spain; Sweden; U.K.

3.B List of factors

Group 1:

AGE: Ratio of the dependent population (younger than 15 and older than 64) to the working-age population (between 15 and 64) (%)

GDPC: Natural logarithm of GDP per capita

POPG: Growth rate of the population (%)

GVBB: Ratio of government overall budget balance (including grants) to GDP (%)

GVTT: Ratio of government total expenditure to GDP (%)

GVIVM: Ratio of government capital expenditure to GDP (%)

GVCE: Ratio of government current expenditure to GDP (%)

GVCON: Ratio of government consumption expenditure to GDP (%)

Group 2:

OPN: Ratio of export plus import to GDP (%)

EXPT: Ratio of exports of goods and services to GDP (%)

IMPT: Ratio of imports of goods and services to GDP (%)

INTD: Absolute real interest rate differential measured for a particular economy towards a world real interest rate index (%)

Group 3:

LGDP: Natural logarithm of GDP

Chapter 4

Testing the bounded unit root

Among economic explanations for a high SI relation introduced in Chapter 2, two of them appear to have the most supporting econometric evidence. One is the stationary CA implied by the solvency constraint. The other is the government targeting the CA imbalances via monetary or fiscal policy. The former is investigated via unit root tests for the CA or cointegration tests for saving and investment in ECMs. The latter is studied by regressing the government budget deficits on the private saving-investment gap. All variables in these studies are measured as a share of GDP. In the recent empirical literature, most studies are performed by applying unit root tests and ECMs to investigate the CA and SI relation .

However, there are three disputed points related to this vein of modeling. Firstly, evidence about the stationarity of the CA or cointegration relation between saving and investment is still quite mixed. Empirical findings in favor of nonstationary imbalances are often explained to mirror power deficiencies of unit root tests applied to short time spans of data. To address the potential of power loss, authors as Coakley and Kulasi (1997) have applied panel unit root tests (Im, Pesaran and Shin 2002). It appears that the evidence from panel unit root approaches is supportive for the view that the CA is stationary. Nevertheless, panel cointegration tests (Kao and Chiang

2001) applied by Ho (2002a) show mixed evidence on a cointegration relation between saving and investment. And evidence for a unit cointegration relation is even weaker. Secondly, one may argue that the bounded nature of CA series measured as a share of GDP has not yet been taken into account when testing the unit root. Traditional unit root tests are formalized to distinguish between stationary processes (I(0)) and processes driven by stochastic trends (I(1)) that can grow or decrease to any level. Time series as the CA imbalance to GDP are bounded by construction. Furthermore, the CA imbalance could be bounded via policy controls or economic crises. This argument can be lightened up by the distinction between *solvency* and *sustainability*. A CA imbalance lacks sustainability if a drastic policy shift is needed or a crisis is likely to emerge under continuation of the current policy stance or of private sector behavior. Sustainability is a more restrictive concept than solvency as argued by Milesi-Ferretti and Razin (1996). Thus, diagnosing stationarity of CA imbalances by rejection of the I(1) hypothesis can reflect the solvency constraint but does not necessarily imply sustainability. The impact of policy controls or economic crises on the series shall be taken into account. From an econometric perspective bounded stationarity is the more natural counterpart of sustainability as the mere rejection of an unbounded I(1) model.

Recently, the issue of testing for stochastic trends governing bounded processes has been addressed by Cavaliere (2005) and Cavaliere and Xu (2007). While Cavaliere (2005) derives the asymptotic distribution of the Phillips-Perron test statistic for a unit root in bounded time series, Cavaliere and Xu (2007) focus on the corresponding distribution of the Augmented Dickey-Fuller test statistic. An important result from these studies is that bounds alter the asymptotic distribution of the unit root statistic, such that inferential conclusions drawn from standard unit root testing suffer from invalid significance levels. Through testing for bounded integration, mean

reverting behavior of stationary series can be separated from (spurious) reverting dynamics of nonstationary but bounded processes.

This chapter¹ analyses the CA imbalances via the bounded unit root test (Phillips-Perron test) introduced in Cavaliere (2005). Annual data for 26 OECD countries (O26) spanning the period 1971 to 2004 are investigated. At the 10% significance level the common unit root hypothesis is rejected for 12 economies. Among these economies bounded stationarity is confirmed for only 6 CA processes. In other words, for 20 out of 26 OECD economies the CA is classified as bounded nonstationary. For testing at an aggregate level, panel based test statistics are constructed as Fisher tests (Maddala and Wu 1999) and by means of the modified inverse normal method (Hartung 1999, Demetrescu, Hassler and Tarcolea 2006) with the latter allowing for dependence among single country test statistics. Our findings support the view that the CA balance is a bounded nonstationary process and rejections of the I(1) model might be due to the existence of bounds in the sense of policy controls or crises. Thus, high CA deficits as observed for some OECD members could be in line with solvency while violating sustainability features. These results also imply that saving and investment are likely not cointegrated. The high association between saving and investment may be partially due to the bounded support of CA imbalances, and, thus, might not imply a low capital mobility. As the CA can also be approximated as exports minus imports, these pair of series fails the cointegration feature too.

The remainder of this chapter is organized as follows: in the next Section the bounded unit root test, the Fisher test, and the modified inverse normal method are sketched briefly. Section 4.2 discusses the generation of critical values for tests on bounded integration. Empirical results are provided in

¹For a different version of this chapter see "Reviewing the sustainability/stationarity of current account imbalances with tests for bounded integration", with Helmut Herwartz, forthcoming in *The Manchester School*.

Section 4.3. Section 4.4 summarizes the main findings and concludes.

4.1 Methodology

Let $\{\mathcal{X}_t\}_0^T$ denote a partial sum (I(1)) process,

$$\mathcal{X}_t = \mathcal{X}_{t-1} + \varepsilon_t, \quad t = 1, \dots, T. \quad (4.1)$$

By assumption, $\{\varepsilon_t\}_1^T$ in (4.1) has mean zero and satisfies regularity conditions. With $[g]$ denoting the integer part of g , it is particularly ensured that $(\lambda^2 T)^{-1/2} \sum_{t=1}^{[T]} \varepsilon_t \xrightarrow{d} B(\cdot)$ where B and $\lambda^2 = \lim_{T \rightarrow \infty} E(T^{-1} (\sum_{t=1}^{[T]} \varepsilon_t)^2)$ are a standard Brownian motion and the so-called long run variance, respectively. A stochastic process $\{X_t\}_0^T$ obtained from mapping the sample paths of $\{\mathcal{X}_t\}_0^T$ on the interval $[\underline{b}, \bar{b}]$ is called bounded integrated of order 1, briefly BI(1). To obtain $\{X_t\}_0^T$, methods as truncation, censoring or reflection can be applied. For instance, in this study the BI(1) process $\{X_t\}_0^T$ is obtained by reflecting $X_{t-1} + \varepsilon_t$ at \underline{b} and \bar{b} . By means of an indicator unfnction, denoted as $\mathbb{I}(\cdot)$, reflection is formalized as

$$X_t = X_{t-1} + \varepsilon_t + \underline{\xi}_t - \bar{\xi}_t, \quad (4.2)$$

where $\varepsilon_t = \Delta \mathcal{X}_t$, $\underline{\xi}_t = [2\underline{b} - (X_{t-1} + \varepsilon_t)]\mathbb{I}(X_{t-1} + \varepsilon_t < \underline{b})$ and $\bar{\xi}_t = [2\bar{b} - (X_{t-1} + \varepsilon_t)]\mathbb{I}(X_{t-1} + \varepsilon_t > \bar{b})$. Once the bounds are reached, the reflected process (X_t) mirrors the original process (\mathcal{X}_t) around the bounds within the interior state. Defining $u_t = \varepsilon_t + \underline{\xi}_t - \bar{\xi}_t$, with $\underline{\xi}_t$ and $\bar{\xi}_t$ satisfying some regularity conditions, Cavaliere (2005) shows that

$$(\lambda^2 T)^{-1/2} \sum_{t=1}^T u_t - (\lambda^2 T)^{-1/2} X_0 \xrightarrow{d} B_{\underline{c}-c_0}^{\bar{c}-c_0}(\cdot),$$

where $B_{\underline{c}-c_0}^{\bar{c}-c_0}$ is a regulated Brownian motion with bounds at

$$\underline{c} - c_0 = \underline{b}(\lambda^2 T)^{-1/2} - X_0(\lambda^2 T)^{-1/2}, \quad \bar{c} - c_0 = \bar{b}(\lambda^2 T)^{-1/2} - X_0(\lambda^2 T)^{-1/2}.$$

Note that the interval $[\underline{b}, \bar{b}]$ describes the nominal bounds of the series $\{X_t\}_0^T$, while $[\underline{c} - c_0, \bar{c} - c_0]$ are the corresponding standardized bounds of the limiting regulated Brownian motion. A regulated Brownian motion can be obtained by adding a standard Brownian motion and some regulators such that the regulated process is forced to stay within a specified interval. For further details on this class of processes, the reader may refer to Harrison (1985).

Now consider the unit root test regression

$$X_t = \rho X_{t-1} + u_t. \quad (4.3)$$

The null hypothesis ($H_0 : \rho = 1$) corresponds to a BI(1) process, and under the alternative ($|\rho| < 1$) X_t is bounded stationary. According to Cavaliere (2005), the Phillips-Perron unit root test statistic (Phillips 1987, Phillips and Perron 1988),

$$J = \left(\sum_{t=1}^T X_{t-1}^2 \right)^{1/2} (\hat{\rho} - 1) / \hat{\lambda} - (\hat{\lambda}^2 - \hat{\sigma}^2) / \left(4 \hat{\lambda}^2 T^{-2} \sum_{t=1}^T X_{t-1}^2 \right)^{1/2},$$

converges to the so-called ‘bounded unit root distribution’ under H_0 , i.e.

$$J \xrightarrow{d} \frac{B_{\underline{c}-c_0}^{\bar{c}-c_0}(1)^2 - 1}{\left(4 \int_0^1 B_{\underline{c}-c_0}^{\bar{c}-c_0}(s)^2 ds \right)^{1/2}}. \quad (4.4)$$

Compared with the common Dickey-Fuller distribution, the probability density function (pdf) of the bounded unit root distribution is shifted to the left. The smaller the bounds are, the more skewed to negative values is the corresponding pdf. As a consequence, for asymmetric intervals the bound which is smaller in absolute value exerts a dominating impact on the left hand side quantiles of the distribution in (4.4).

Apart from testing the BI(1) hypothesis for single economies, a panel perspective is also adopted. To obtain a test statistic at the pooled level from single country test results, the Fisher test (Maddala and Wu 1999) and the

modified inverse normal method (Hartung 1999, Demetrescu, Hassler and Tarcolea 2006) are considered. While the former approach relies on cross-sectional independence of test statistics, the latter method allows cross sectional dependence that is likely in applied macroeconomic modeling. The Fisher test is derived from test statistics J_i , $i = 1, \dots, N$, that are obtained from a set of a cross sectional test regressions as (4.3). Under the null hypothesis $H_0 : \rho_i = 1 \forall i$ a Fisher statistic is obtained as

$$\Gamma = -2 \sum_{i=1}^N \ln(p_i(J_i)) \xrightarrow{d} \chi^2(2N), \quad (4.5)$$

where $p_i(J_i)$ is the p -value associated with J_i under the BI(1) null hypothesis. Accounting for potential dependence across panel units, the modified inverse normal method is based on probits, $t_i = \Phi^{-1}(p_i(J_i))$, where Φ^{-1} denoting the inverse of the standard normal cumulative distribution function. For the case of a constant correlation among probits i.e. $Cov(t_i, t_j) = \rho$ for $i \neq j$, Hartung (1999) suggests the weighted test statistic

$$t(\hat{\rho}^*, \kappa) = \frac{\sum_{i=1}^N \nu_i t_i}{\sqrt{\sum_{i=1}^N \nu_i^2 + \left[\left(\sum_{i=1}^N \nu_i \right)^2 - \sum_{i=1}^N \nu_i^2 \right] \left[\hat{\rho}^* + \kappa \sqrt{\frac{2}{N+1}} (1 - \hat{\rho}^*) \right]}}, \quad (4.6)$$

which is approximately Gaussian under H_0 as $T \rightarrow \infty$. In (4.6) ν_i denote the weights, $\hat{\rho}^* = \max\left(-\frac{1}{N-1}, \hat{\rho}\right)$, $\hat{\rho} = 1 - \frac{1}{N-1} \sum_{i=1}^N \left(t_i - \frac{1}{N} \sum_{i=1}^N t_i\right)^2$, and $\kappa > 0$ regulates the actual significance level in small samples. In this study, $\nu_i = 1$ and $\kappa = 0.2$ as in Hartung (1999).

4.2 Generating critical values

Bounded random walks are simulated to obtain critical values of the unit root test statistics J_i . To determine the relevant standardized bounds $[\underline{c}_i -$

$c_{i0}, \bar{c}_i - c_{i0}]$, the long run variance λ^2 is firstly estimated by means of the sum-of-covariances estimator (Andrews 1991),

$$\hat{\lambda}_i^2 = \frac{1}{T} \sum_{j=-T+1}^{T-1} K\left(\frac{j}{h}\right) \sum_{t=|j|+1}^T \hat{u}_{ti} \hat{u}_{ti-|j|},$$

where $K(\cdot)$ is the quadratic spectral kernel implemented with bandwidth h . In this study h is estimated according to the automatic bandwidth selection with a first order autoregressive model for \hat{u}_{ti} (Andrews 1991). Secondly, the nominal bounds $[\underline{b}_i, \bar{b}_i]$ are selected. It might be argued that CA measured as percentages of GDP have natural bounds $[-100, 100]$. However, as CA dynamics are likely subjected to policy control, there might be hidden bounds that are considerably smaller in absolute value. Since the true bounds are not observable, we consider a set of potential bounds ranging from the observed country specific maximum and minimum CA measures to respective counterparts that are up to 30% larger in absolute value. Define $\underline{b}_i = \min_t(X_{it})$ and $\bar{b}_i = \max_t(X_{it})$, where X_{it} represents the CA account imbalance observed in country i and period t . The nuisance parameters $\underline{c}_i, \bar{c}_i, c_{i0}$ are estimated for each country i as $\hat{\underline{c}}_i = \underline{b}_i(\hat{\lambda}_i^2 T)^{-1/2}$, $\hat{\bar{c}}_i = \bar{b}_i(\hat{\lambda}_i^2 T)^{-1/2}$ and $\hat{c}_{i0} = X_{i0}(\hat{\lambda}_i^2 T)^{-1/2}$. The considered set of alternative bounds for the regulated Brownian motion is expressed as $[(\hat{\underline{c}}_i - \hat{c}_{i0}) - \delta\omega_i/2, (\hat{\underline{c}}_i - \hat{c}_{i0}) + \delta\omega_i/2]$, $\omega_i = |\hat{\bar{c}}_i - \hat{\underline{c}}_i|$, where $\delta = 0\%, 10\%, \dots, 30\%$. Note that the common Dickey-Fuller distribution corresponds to a case where the implicit bounds approach infinity, i.e. $\delta = \infty$. Given the value of the bounds, bounded random walks of length $T = 1000$ starting with $X_0 = 0$ are simulated with 10000 replications. The (cross section specific) distribution in (4.4) is approximated by the distribution of Dickey-Fuller test statistic obtained from the bounded random walks². It is worth noting that considering bounds for the CA being 30% larger than the observed interval is rather conserva-

²The empirical results with simulated sample size $T = 2000$ are similar as those from $T = 1000$. Moreover, all results documented in this study are qualitatively identical if

tive. According to the present-value representation of the CA in Sheffrin and Woo (1990), a 15% increase in CA surplus or deficit corresponds to a 15% decrease or increase in the change of total net output at all future time horizons, *ceteris paribus*.

The selection of (latent) bounds is crucial for the inferential conclusion drawn from testing the BI(1) hypothesis. Therefore, two additional perspectives are provided to assess potential pitfalls of testing the I(1) hypothesis for bounded processes. Firstly, some ‘break-even’ bounds are determined at which distributional features in terms of p -values obtained from testing the BI(1) and I(1) model cannot be distinguished from each other. For this purpose, the initial estimates $[\hat{c}_i - \hat{c}_{i0}, \hat{c}_i - \hat{c}_{i0}]$ are enlarged proportionally until the bounded model obtains the same p -value for the conventional Dickey-Fuller distribution and the bounded unit root distribution in (4.4). The corresponding nominal break-even bounds $[\underline{b}_i^*, \bar{b}_i^*]$ then indicate a minimum range of the CA under which the standard unit root test does not suffer from oversizing. In other words, rather unreasonably large break-even bounds in comparison with the observed CA ranges $[\underline{b}_i, \bar{b}_i]$ signifies that neglecting the bounded support in the unit root tests might be misleading. Secondly, based on the observed CA ranges, the empirical size violation of common unit root inference implied by the bounded unit root distribution is provided. To be precise, the empirical rejection frequencies ($p_{I(1)}$) for the distribution in (4.4) compared with the 5% quantile of the Dickey-Fuller distribution are reported. Adverse effects of ignoring bounds in unit testing become stronger as the empirical size differs markedly from 5%.

(instead of reflection) censoring is used as an alternative mapping scheme, or the variance parameter σ_i^2 is estimated by imposing $\rho = 1$.

4.3 Results

The considered data set consists of the CA, domestic investment, domestic saving, exports and imports as a share to GDP for the cross section O26 introduced in Chapter 2. The CA balance are approximated as domestic saving minus domestic investment rather than national saving minus domestic investment in order to be consistent with the main empirical discussions of the Feldstein-Horioka puzzle. For this chapter, the data are drawn from the World Development Indicators CD-Rom 2006 published by the World Bank. The sample comprises yearly observations over the period 1971 to 2004. For six countries quotes in 2004 are not available. These missing values are estimated by means of univariate autoregressive models of order 1 with intercept. When only the sample period from 1971 to 2003 is considered, inferential results are qualitatively identical to those reported in this section. Throughout, the discussion of test results is conditional on the 10% significance level.

The results from (bounded) unit root tests for the CA balance are shown in Table 4.1. Among the observed rescaled minimum ($\hat{c}_i - \hat{c}_{i0}$) and maximum ($\hat{c}_i - \hat{c}_{i0}$) CA quotes, the US and Ireland have the lowest (-1.18) and highest (1.43) measure, respectively. When BI(1) tests are applied with increasing bounds, it can be observed that the p -values for the test statistic J_i are monotonously decreasing for almost all economies³. As such, overestimating the bounds biases the results to over-reject the (bounded) unit root hypothesis. In the extreme case of ignoring the limited support of CA measures, diagnosing stationarity of the CA might be strongly misleading. For instance, as can be seen from the p -values in Table 4.1, the I(1) hypothesis is rejected for 12 out of 26 countries ($\delta = \infty$). If the underlying bounds are up

³For economies as Korea and UK with extremely asymmetric bounds, increasing bounds symmetrically may rise the corresponding p -values.

Table 4.1: Testing current account for (bounded) nonstationarity

| Country | J_i | \underline{b}_i | \bar{b}_i | $\hat{c}_i - \hat{c}_{i0}$ | $\hat{c}_i - \hat{c}_{i0}$ | BI(1) p -values with $\delta =$ | | | | | $p_{I(1)}$ | \underline{b}_i^* | \bar{b}_i^* | Δ |
|---------|-------|-------------------|-------------|----------------------------|----------------------------|-----------------------------------|------------|------------|------------|------------|------------|---------------------|---------------|----------|
| | | | | | | 0% | 10% | 20% | 30% | ∞ | | | | |
| AUS | -2.25 | -3.55 | 3.17 | -0.57 | 0.32 | .23 | .16 | .09 | .06 | .02 | .36 | -7.02 | 6.64 | 103% |
| AUT | -1.00 | -3.85 | 4.99 | -0.65 | 0.68 | .72 | .58 | .46 | .39 | .28 | .06 | -8.36 | 9.51 | 102% |
| BEL | -0.87 | -3.36 | 4.44 | -0.98 | 0.41 | .64 | .64 | .64 | .62 | .34 | .21 | -8.90 | 9.98 | 142% |
| CAN | -1.16 | -1.82 | 5.77 | -0.49 | 0.58 | .87 | .72 | .58 | .47 | .23 | .13 | -7.05 | 11.01 | 138% |
| DEN | -0.75 | -4.46 | 6.81 | -0.32 | 1.00 | .67 | .65 | .65 | .66 | .40 | .28 | -14.54 | 16.89 | 179% |
| FIN | -1.03 | -6.22 | 9.25 | -0.32 | 0.96 | .55 | .55 | .57 | .57 | .27 | .28 | -17.59 | 20.61 | 147% |
| FRA | -2.27 | -3.26 | 2.90 | -0.72 | 0.40 | .09 | .05 | .03 | .03 | .02 | .22 | -4.91 | 4.55 | 53% |
| GER | -0.34 | -5.03 | 4.93 | -0.58 | 1.24 | .79 | .79 | .78 | .75 | .55 | .07 | -12.71 | 12.60 | 154% |
| GRC | -0.22 | -10.46 | -2.18 | -0.52 | 0.63 | 1.0 | 1.0 | 1.0 | 1.0 | .61 | .10 | -22.61 | 9.97 | 293% |
| HUN | -3.29 | -9.12 | 4.57 | -0.15 | 0.56 | .04 | .05 | .06 | .03 | .00 | .31 | -14.89 | 10.33 | 84% |
| ISL | -3.42 | -9.78 | 5.33 | -0.17 | 0.62 | .04 | .05 | .03 | .01 | .00 | .26 | -14.80 | 10.35 | 66% |
| IRL | -0.29 | -16.88 | 16.56 | -0.53 | 1.43 | .79 | .78 | .78 | .78 | .58 | .09 | -41.39 | 41.07 | 147% |
| ITA | -2.27 | -4.17 | 4.93 | -0.57 | 0.50 | .04 | .03 | .02 | .02 | .02 | .13 | -5.88 | 6.64 | 38% |
| JPN | -1.92 | -0.92 | 3.92 | -0.62 | 0.21 | .29 | .28 | .28 | .29 | .05 | .28 | -4.40 | 7.40 | 144% |
| KOR | -2.96 | -11.16 | 12.87 | -0.02 | 1.05 | .01 | .03 | .06 | .08 | .00 | .07 | -22.81 | 24.52 | 97% |
| MEX | -2.43 | -5.03 | 9.58 | -0.26 | 0.69 | .17 | .16 | .10 | .06 | .02 | .27 | -12.72 | 17.27 | 105% |
| NLD | -0.36 | -1.11 | 6.11 | -0.08 | 0.94 | .84 | .81 | .79 | .78 | .56 | .13 | -14.78 | 19.79 | 379% |
| NOR | -1.10 | -9.16 | 17.33 | -0.27 | 0.80 | .60 | .56 | .56 | .56 | .25 | .27 | -33.06 | 41.24 | 180% |
| NZL | -3.20 | -12.08 | 3.46 | -0.82 | 0.17 | .06 | .06 | .04 | .01 | .00 | .20 | -17.75 | 9.12 | 73% |
| PRT | -0.81 | -15.79 | -1.46 | -0.65 | 0.24 | .96 | .91 | .87 | .82 | .37 | .28 | -35.27 | 18.02 | 272% |
| SPA | -1.79 | -5.13 | 1.27 | -0.55 | 0.10 | .37 | .35 | .33 | .33 | .07 | .27 | -13.20 | 9.35 | 252% |
| SWE | 0.37 | -1.59 | 8.03 | -0.43 | 1.00 | 1.0 | 1.0 | 1.0 | .99 | .79 | .18 | -16.21 | 22.65 | 304% |
| SWI | -0.42 | -1.98 | 7.06 | -0.19 | 0.88 | .89 | .85 | .83 | .81 | .54 | .22 | -17.82 | 22.90 | 351% |
| TUR | -1.75 | -7.49 | 2.40 | -0.33 | 0.36 | .99 | .94 | .78 | .57 | .08 | .94 | -15.44 | 10.36 | 161% |
| UK | -1.93 | -4.92 | 2.84 | -0.74 | 0.17 | .21 | .24 | .26 | .28 | .05 | .21 | -10.70 | 8.62 | 149% |
| US | 0.38 | -4.78 | 0.99 | -1.18 | 0.22 | .99 | .98 | .96 | .95 | .79 | .24 | -14.22 | 10.43 | 328% |

Phillips-Perron statistics J_i for the 26 OECD countries are shown with the observed minimum and maximum bounds $[\underline{b}_i, \bar{b}_i]$ and the corresponding estimated bounds $[\hat{c}_i - \hat{c}_{i0}, \hat{c}_i - \hat{c}_{i0}]$ for the regulated Brownian motion. For BI(1) tests bounds are increased proportionally (0%, 10%, 20%, 30%) to $[\hat{c}_i - \hat{c}_{i0}, \hat{c}_i - \hat{c}_{i0}]$. The p -values from I(1) tests are denoted as the case when the increased proportion of bounds is ∞ . p -values smaller than 0.10 are highlighted in bold face. The column $p_{I(1)}$ documents empirical rejection frequencies obtained from sampling BI(1) test statistics and comparing them with the 5% quantile of the Dickey-Fuller distribution. $[\underline{b}_i^*, \bar{b}_i^*]$ are the break-even bounds equating the p -values from testing the I(1) and BI(I) model. $\Delta = \frac{|\bar{b}_i^* - \underline{b}_i^*|}{|\bar{b}_i - \underline{b}_i|} - 1$ is the increased proportion of break-even bounds compared to the observed ranges. The considered countries are AUS-Australia, AUT-Austria, BEL-Belgium, CAN-Canada, DEN-Denmark, FIN-Finland, FRA-France, GER-Germany, GRC-Greece, HUN-Hungary, ISL-Iceland, IRL-Ireland, ITA-Italy, JPN-Japan, KOR-Korea (Rep.), MEX-Mexico, NLD-Netherlands, NOR-Norway, NZL-New Zealand, PRT-Portugal, SPA-Spain, SWE-Sweden, SWI-Switzerland, TUR-Turkey, UK-United Kingdom and US-United States.

to 10% larger than the observed range the BI(1) hypothesis is rejected for only 6 countries. Thus, for 6 significant I(1) test statistics the mean reverting behavior of the CA might be addressed to the bounded nature of the data (Australia, Japan, Mexico, Spain, Turkey, and UK).

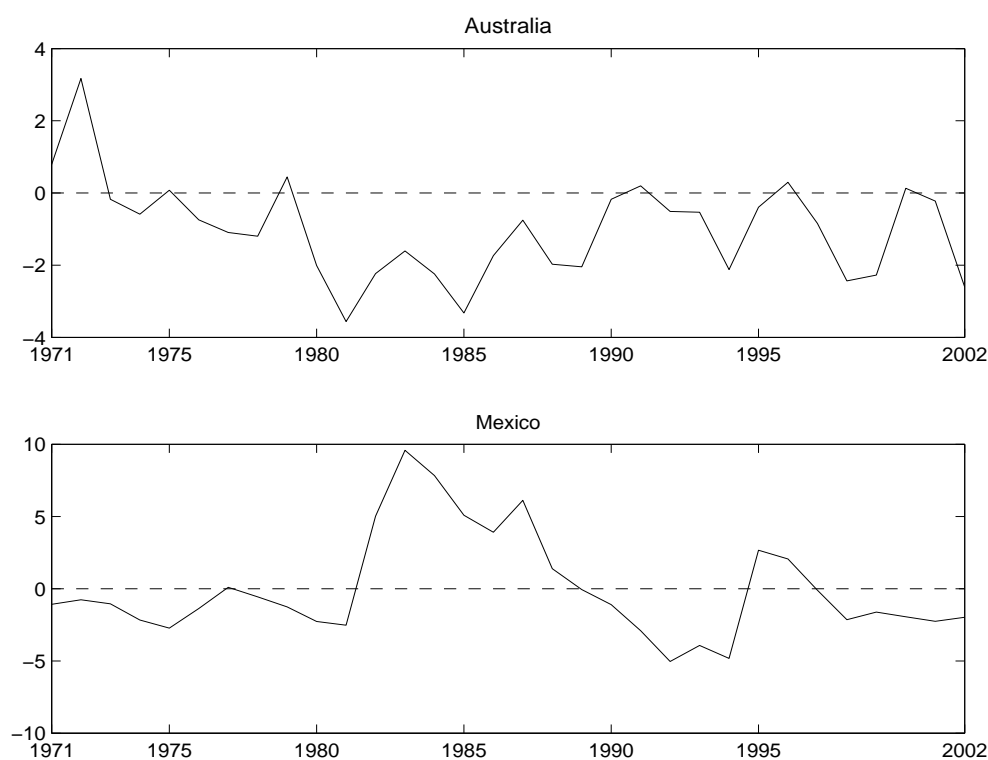


Figure 4.1: Current account imbalances measured as percentages of GDP.

For example, while CA movements of Australia might be subject to policy control, those of Mexico are likely subject to an economic crisis. As can be seen in Figure 4.1, Australia has experienced relatively high CA deficits since 1980. However, a deficit peak has never been persistent. Both policy statements and empirical work have supported the view that Australia's government or central bank is targeting the CA imbalances. Paul Keating, treasurer in the government of Bob Hawke, stated many times that the economic policy in Australia had been directed to reduce the CA deficits. Empirical investigations as Karfakis and Kim (1995) support the view that the

CA deficit news signals a future monetary contraction by the Reserve Bank of Australia. In contrast to Australia, Mexico has experienced financial crisis. Since 1985 the Mexican government has began to liberalize the trade sector and introduce market-oriented institutions. As the reform became consolidated, Mexico was subject to huge capital inflows which helped to finance large CA deficits. As can be seen in Figure 4.1, the deficits were around 5% between 1992 and 1994. These deficits turned out to be unsustainable and led to economic crisis in 1994.

Moreover, the obtained break-even bounds $[\underline{b}_i^*, \bar{b}_i^*]$ are provided in the right hand side panel of Table 4.1. To have the same p -values of the J_i statistics for the 12 economies diagnosed as unbounded stationary, the implied bounds are between 38% (Italy) to about 252% (Spain) times larger in comparison with their observed ranges. Most of these bounds likely reach above a sustainable CA level. Considering Mexico, for instance, the latter implicit interval equalizing p -values of testing the I(1) and BI(1) model is $[-12.72, 17.27]$. As argued, however, a CA deficit of about 5% of GDP in Mexico has already induced a financial crisis at the end of 1994.

Empirical size violations involved with common unit root testing further illustrate how inferential results are biased towards the rejection of stochastic trends. It can be seen that I(1) tests have throughout empirical sizes much larger than the presumed nominal significance level of 5%. For an extreme case, namely Turkey, the common I(1) test diagnoses stationarity of the underlying bounded nonstationary series with $p_{I(1)} = 94\%$ significance.

At last, combining p -values of the common Phillips-Perron statistic to an aggregate measure, the unit root hypothesis is robustly rejected via Fisher tests and implementations of the modified inverse normal method with a cross correlation coefficient up to 0.5 ($\delta = \infty$, last column in Table 4.2). When testing the BI(1) model via Fisher tests, the CA is diagnosed as bounded nonstationary for bounds that are up to 20% larger than the observed range.

Table 4.2: Panel unit root tests for (bounded) integration

| Increased proportion of bounds (δ) | 0% | 10% | 20% | 30% | 40% | 50% | ∞ |
|---|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|
| Fisher | 60.09 [0.21] | 60.24 [0.20] | 65.72 [0.10] | 74.87 [0.02] | 85.68 [0.00] | 97.23 [0.00] | 131.24 [0.00] |
| $t(\hat{\rho}^* = 0.1)$ | 0.81 [0.79] | 0.53 [0.70] | 0.09 [0.53] | -0.35 [0.36] | -0.81 [0.21] | -1.18 [0.12] | -2.49 [0.01] |
| $t(\hat{\rho}^* = 0.3)$ | 0.57 [0.72] | 0.38 [0.65] | 0.06 [0.52] | -0.25 [0.40] | -0.58 [0.28] | -0.83 [0.20] | -1.76 [0.04] |
| $t(\hat{\rho}^* = 0.5)$ | 0.47 [0.68] | 0.31 [0.62] | 0.05 [0.52] | -0.20 [0.42] | -0.47 [0.32] | -0.68 [0.25] | -1.44 [0.07] |

Results from applications of the modified inverse normal method. p -values appear in parentheses below the corresponding test statistics. The bounds are increased proportionally (0%, ..., 50%) to $[\hat{c}_i - \hat{c}_{i0}, \hat{c}_i - \hat{c}_{i0}]$.

It might be argued that CA account imbalances in the OECD area likely show contemporaneous correlation of underlying innovations. This feature is ignored when constructing Fisher statistics (Banerjee and Zanghieri 2003). Allowing for potential contemporaneous error correlations, the modified inverse normal method shows that the CA is bounded nonstationary for bounds that are up to 50% larger than the observed range.⁴ Moreover, it can be observed that the higher the cross correlation, the stronger the evidence supporting the null hypothesis of bounded integration. Keeping the null hypothesis of overall bounded nonstationarity, the panel perspective does not suffer from inconclusiveness under the heterogeneous alternative ($H_1 : |\rho_i| < 1$, for some i).

⁴The estimated cross correlation coefficient $\hat{\rho}$ from the data is smaller than $-\frac{1}{N-1} = -0.04$. Thus, according to Hartung (1999) $\hat{\rho}^* = -0.04$ should be chosen. Alternatively, $\hat{\rho}^* = 0.1, 0.3, 0.5$ is selected to demonstrate the effect of the cross correlation. As a competing avenue to account for cross sectional error correlations we simulate bounded processes from innovations that exhibit the empirical covariance features of autoregressions fitted to the CA series. Empirical results on this issue are not shown but are qualitatively identical to those obtained from the inverse normal method.

As a direct implication of these results, saving and investment measured as a ratio of GDP fail the cointegration feature. Testing the (bounded) unit root hypothesis for domestic saving and investment strongly supports the view that these series are (bounded) nonstationary. As can be seen in Table 4.3, for all 26 OECD economies except Japan both the I(1) and BI(1) null hypothesis cannot be rejected. For the latter case, the observed ranges of the series are used as bounds. Since increasing bounds for the BI(1) model provides decreasing p -values at most to the level from the I(1) model which is still larger than 5%, applying bounds larger than the observed range delivers the same results. Besides, neglecting the net capital gains on foreign assets, the CA balance can be also (approximately) decomposed as exports minus imports⁵. Results of BI(1) and I(1) tests for the latter series are shown in Table 4.4. The time series of exports and imports to GDP are found nonstationary and bounded nonstationary, respectively, for all economies considered. Thus, according to the bounded nonsationarity of CA imbalances, supporting evidence for conintegration relation between exports and imports is also not found.

⁵Time series of the CA balances constructed by exports minus import are fast exactly the same as those obtained by domestic saving minus investment. The only difference lies in the missing values in 2004, which are caculated by autoregresive models.

Table 4.3: (Bounded) unit root tests for saving and investment

| Country | Saving | | | Investment | | |
|---------|--------|-------------|-------------|------------|-------|-------------|
| | J_i | p_b | p_{ub} | J_i | p_b | p_{ub} |
| AUS | -1.28 | 0.21 | 0.19 | -0.67 | 0.22 | 0.19 |
| AUT | -0.56 | 0.64 | 0.48 | -1.15 | 0.72 | 0.48 |
| BEL | -1.01 | 0.27 | 0.28 | -1.01 | 0.46 | 0.28 |
| CAN | -0.22 | 1.00 | 0.61 | -0.46 | 1.00 | 0.61 |
| DEN | 0.04 | 1.00 | 0.71 | -0.83 | 1.00 | 0.71 |
| FIN | -0.53 | 0.98 | 0.49 | -1.03 | 0.86 | 0.49 |
| FRA | -1.42 | 0.32 | 0.15 | -1.27 | 0.42 | 0.15 |
| GER | -1.00 | 0.27 | 0.28 | -1.80 | 0.27 | 0.28 |
| GRC | -1.16 | 0.34 | 0.22 | -0.80 | 0.43 | 0.22 |
| HUN | -0.95 | 0.46 | 0.30 | -1.14 | 0.48 | 0.30 |
| ISL | -1.48 | 0.39 | 0.13 | -1.31 | 0.31 | 0.13 |
| IRL | 2.19 | 1.00 | 0.99 | -0.13 | 1.00 | 0.99 |
| ITA | -1.09 | 0.53 | 0.25 | -0.77 | 0.36 | 0.25 |
| JPN | -2.28 | 0.02 | 0.02 | -1.54 | 0.19 | 0.02 |
| KOR | 0.84 | 0.98 | 0.89 | -0.09 | 1.00 | 0.89 |
| MEX | -0.24 | 0.99 | 0.60 | -0.15 | 1.00 | 0.60 |
| NLD | -0.37 | 0.81 | 0.56 | -1.36 | 0.59 | 0.56 |
| NOR | -0.07 | 1.00 | 0.66 | -1.20 | 0.95 | 0.66 |
| NZL | -0.51 | 1.00 | 0.50 | -0.46 | 1.00 | 0.50 |
| PRT | -0.72 | 0.96 | 0.40 | -0.49 | 1.00 | 0.40 |
| SPA | -0.52 | 0.74 | 0.49 | -0.06 | 1.00 | 0.49 |
| SWE | -0.39 | 0.94 | 0.55 | -1.07 | 0.90 | 0.55 |
| SWI | -1.35 | 0.15 | 0.16 | -1.49 | 0.15 | 0.16 |
| TUR | 0.25 | 0.96 | 0.76 | 0.51 | 1.00 | 0.76 |
| UK | -1.45 | 0.22 | 0.14 | -0.55 | 0.96 | 0.14 |
| US | -1.07 | 0.55 | 0.26 | -0.27 | 1.00 | 0.26 |

Phillips-Perron test statistic J_i . Taking the observed range of series as bounds, p_b presents the corresponding p -value from the BI(1) test. p_{ub} describes the p -value from the I(1) test. For further notes see also Table 4.1.

Table 4.4: Further (bounded) unit root tests

| Country | Export | | | Import | | |
|---------|--------|-------|----------|--------|-------|----------|
| | J_i | p_b | p_{ub} | J_i | p_b | p_{ub} |
| AUS | 0.39 | 0.90 | 0.80 | 0.99 | 0.99 | 0.92 |
| AUT | 2.36 | 1.00 | 1.00 | 1.68 | 1.00 | 0.98 |
| BEL | 1.42 | 1.00 | 0.96 | 1.23 | 1.00 | 0.95 |
| CAN | 0.92 | 0.97 | 0.90 | 0.94 | 0.97 | 0.91 |
| DEN | 1.18 | 1.00 | 0.94 | 0.68 | 1.00 | 0.86 |
| FIN | 0.69 | 0.99 | 0.87 | 0.34 | 1.00 | 0.78 |
| FRA | 1.23 | 0.99 | 0.94 | 0.88 | 1.00 | 0.90 |
| GER | 2.99 | 1.00 | 1.00 | 1.49 | 1.00 | 0.97 |
| GRC | 0.34 | 0.97 | 0.78 | 0.67 | 0.96 | 0.86 |
| HUN | 0.77 | 0.99 | 0.88 | 0.80 | 1.00 | 0.89 |
| ISL | -0.23 | 1.00 | 0.60 | -0.28 | 0.97 | 0.58 |
| IRL | 1.31 | 0.97 | 0.96 | 0.49 | 0.93 | 0.83 |
| ITA | 0.97 | 1.00 | 0.91 | 0.72 | 1.00 | 0.86 |
| JPN | -0.27 | 1.00 | 0.59 | -0.31 | 1.00 | 0.57 |
| KOR | 0.77 | 1.00 | 0.88 | 0.56 | 1.00 | 0.84 |
| MEX | 0.70 | 1.00 | 0.87 | 1.56 | 1.00 | 0.97 |
| NLD | 0.92 | 1.00 | 0.91 | 0.72 | 1.00 | 0.88 |
| NOR | 0.26 | 1.00 | 0.76 | -0.77 | 0.88 | 0.39 |
| NZL | 0.27 | 1.00 | 0.77 | 0.15 | 1.00 | 0.74 |
| PRT | 0.27 | 1.00 | 0.77 | 0.36 | 1.00 | 0.79 |
| SPA | 0.88 | 0.99 | 0.90 | 1.46 | 0.99 | 0.96 |
| SWE | 1.44 | 1.00 | 0.97 | 0.98 | 1.00 | 0.91 |
| SWI | 1.65 | 1.00 | 0.97 | 0.32 | 1.00 | 0.77 |
| TUR | 1.11 | 0.98 | 0.93 | 1.96 | 1.00 | 0.99 |
| UK | 0.06 | 1.00 | 0.70 | 0.31 | 1.00 | 0.77 |
| US | 0.50 | 0.98 | 0.83 | 1.93 | 0.99 | 0.99 |

Phillips-Perron test statistic J_i . Taking the observed range of series as bounds, p_b presents the corresponding p -value from the BI(1) test. p_{ub} describes the p -value from the I(1) test. For further notes see also Table 4.1.

4.4 Summary

This chapter contrasts tests on nonstationarity and bounded nonstationarity of current account imbalances for 26 OECD countries over the period 1971 to 2002. The standard approach of unit root testing by means of the Phillips-Perron statistic ignores the limited support of current account imbalances expressed as percentages of the gross domestic product. Taking account of the bounded nature of actual current account processes weakens the evidence against the prevalence of stochastic trends throughout. At the aggregate level the null hypothesis of bounded integration cannot be rejected for the current account imbalances. Therefore, policy controls or economic crises might cause mean reverting patterns of CA imbalances on the one hand. On the other hand, however, they could hide important stochastic features pointing to unsustainability of the CA for numerous developed economies.

A direct implication of these results is that saving and investment are likely not cointegrated. The high association between saving and investment may be partially due to the existence of the bounds on the CA imbalances. Given this evidence, ECMs for saving and investment are not treated as derived from a system of cointegrating variables. Rather its empirical performance will be addressed in comparison with static panel based formalizations of the SI relation since error correction dynamics might also be formalized for stationary variables or bounded nonstationary variables. The next chapter is devoted for the model comparison.

Chapter 5

Model selection

A wide portfolio of heterogeneous econometric specifications has been employed to investigate the SI relation. Somewhat surprisingly, the relative merits of competing model classes have not yet been provided in a systematic and comprehensive fashion. The purpose of this chapter¹ is to determine a family of econometric models that is most suitable in explaining actual investment ratios via cross-validation (CV) techniques. The three categories of panel data models discussed in Section 2.2.2 are considered: static panel models, weakly dynamic models, and ECMs.

Firstly, between, pooled, time dependent and country dependent specifications of the SI relation in static models are compared systematically. Secondly, the scope of static and weakly dynamic models addressing the SI relation are distinguished. This comparison is informative to uncover potential mean reverting features of the saving and investment ratio since differencing stationary time series will likely involve a loss in accuracy of fit. In the opposite case of nonstationary ratios, a model in first differences is suitable to guard against spurious regressions. Since taking first differ-

¹For a different version of this chapter see “Panel data model comparison for empirical saving-investment relations”, with Helmut Herwartz, forthcoming in *Applied Economic Letters*.

ences of investment ratios will also remove individual effects, this comparison sheds light on the prevalence of individual effects as a characteristic of investment ratios. Thirdly, weakly dynamic models are contrasted against ECMs to distinguish cointegrating features from scenarios of independent stochastic trends governing the domestic saving and investment ratio.

This chapter is organized as follows: In the first Section the considered panel data models are introduced. Model selection criteria are described in Section 5.2. Empirical results obtained from the model comparison are provided in Section 5.3. Section 5.4 summarizes briefly the main findings.

5.1 Model specifications

An unrestricted static representation of the relationship between domestic investment and saving may be given as

$$I_{it} = a_{it} + b_{it}S_{it} + u_{it}. \quad (5.1)$$

The empirical implementation of the relation in (5.1) will, generally, require some restrictions on the parameters a_{it} and b_{it} which could be formalized in the time-series dimension, the cross-section dimension or both. Following these lines we consider four settings for the choice of these two parameters: In the first two places the model parameters are estimated by means of pooled and between regressions, abbreviated and formalized as

$$\begin{aligned} pol : I_{it} &= a + bS_{it} + u_{it}, \\ bet : \bar{I}_i &= a + b\bar{S}_i + \bar{u}_i, \end{aligned}$$

respectively. As two main competitors of these highly restricted regression designs the parameters of the model in (5.1) can be either time specific or

vary over the cross-section, i.e.

$$tim : I_{it} = a_t + b_t S_{it} + u_{it},$$

$$cro : I_{it} = a_i + b_i S_{it} + u_{it}.$$

Apart from general static models as formalized in (5.1), SI relations can be considered in ‘weakly dynamic’ models as

$$\Delta I_{it} = c_{it} + d_{it} \Delta S_{it} + v_{it}, \quad (5.2)$$

where Δ is the first difference operator, e.g. $\Delta I_{it} = I_{it} - I_{i,t-1}$. As when (5.1) is implemented, CV measures are provided for pooled, between, time and cross-section specific regressions of ΔI_{it} on ΔS_{it} . In a further step the ‘weakly dynamic’ model in (5.2) will be augmented with (alternative representations of) lagged error correction terms. To be explicit the following model versions are compared by means of CV criteria:

$$ecm1 : \Delta I_{it} = \delta_i + \alpha_i (I_{i,t-1} - \eta_i S_{i,t-1}) + \Gamma_i \Delta S_{it} + e_{it},$$

$$ecm2 : \Delta I_{it} = \delta_i + \alpha_i (I_{i,t-1} - S_{i,t-1}) + \Gamma_i \Delta S_{it} + e_{it},$$

$$ecm3 : \Delta I_{it} = \delta_i + \alpha_i (I_{i,t-1} - S_{i,t-1}) + e_{it}.$$

Whereas the first model allows the parameter η_i to enter unrestrictedly, the last two models make use of the restriction $\eta_i = 1$ implying that the current account imbalance impacts on the investment ratio.

Regarding the four settings for the static model (5.1), one may also consider time varying ECMs for the completeness of model comparison. Although time variation may, in principle, also apply for error correction dynamics we refrain from formalizing time dependent ECM models for two reasons: First, CV criteria estimated for the model class in (5.2) will show that time dependence is likely not an important feature of the parametric description of ΔI_{it} . Secondly, in the light of recent work on threshold cointegration (Balke and Fomby 1997) it is likely that time variation in α_i is better

conditioned upon economic states rather than presuming deterministic time shifts of model parameters.

5.2 Cross validation criteria

In principle, model comparison may follow some optimization of in-sample criteria (log-likelihood estimates, model selection criteria, (adjusted) R^2 , etc.) or out-of-sample performance. Since in-sample features of alternative panel data models often only allow more or less trivial rankings according to the number of explanatory variables (pooled regression, between regression, within regression, allowance of cross-section specific or time dependent parameters, etc.), it is a-priori more tempting to base model evaluation on some measure of out-of-sample performance. To obtain criteria for model comparison, CV techniques (Allen 1974, Stone 1974, Geisser 1975) are employed. These techniques are seen as an out-of-sample based means to distinguish the relative merits of competing models that is not trivially affected by outstanding factors as e.g. the number of model parameters. CV techniques are widely used in applied non- and semiparametric modeling. In the following a brief outline of the implementation of cross validation methods in this study is provided.

To discriminate panel based estimators at an aggregated level the following CV criterion is used:

$$cv_{(mod)} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T |y_{it} - \hat{y}_{it}(mod)|. \quad (5.3)$$

In (5.3) 'forecasts' $\hat{y}_{it}(mod)$ for some dependent variable of interest (the investment ratio I_{it} , say) are based on so-called leave one out or jackknife estimators, i.e.

$$\hat{y}_{it}(mod) = z'_{it} \hat{\beta}_{mod,it}, \quad (5.4)$$

with $\hat{\beta}_{-mod,it}$ being an estimated parameter vector that is obtained from a particular model, $y_{it} = z'_{it}\hat{\beta}_{-mod,it} + e_{it}$, after removing the it -th pair of dependent and explanatory variables from the sample. The particular model representations entering CV based comparisons are those introduced in the last section. Apart from model comparison by means of absolute forecast errors CV criteria derived from squared forecast errors are also provided, i.e.

$$cv^2_{(mod)} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{y}_{it(mod)})^2. \quad (5.5)$$

In addition, the determination of CV measures for the representation of changes of the investment ratio may follow the same lines as discussed for the level representation. To allow cross model comparison, however, jackknife forecasts of ΔI_{it} have to be transformed to forecasts for the level variables I_{it} . Since $\hat{I}_{it} = \Delta \hat{I}_{it} + I_{it-1}$, CV estimates comparing \hat{I}_{it} and I_{it} are equal to those obtained from a comparison of $\Delta \hat{I}_{it}$ and ΔI_{it} . For the purpose of informationally equivalent model comparison we compute CV criteria for the level of the investment ratio using the model family in (5.2) and recursive forecasts \hat{I}_{it-1} , $t = 2, \dots, T$, initialized with the first observation I_{i1} . Note that CV estimates for the ‘weakly dynamic’ model family and ECMs are generally obtained over samples covering one observation less in comparison with the level representation in (5.1).

5.3 Results from model comparisons

Cross-validation results for the eight cross-sections introduced in Chapter 3 are documented in Table 5.1. The panels A, B, and C of the Table show the CV estimates for models specified in levels, first differences and ECM model versions, respectively. Apart from giving raw CV measures (cv and cv^2) scale invariant normalized results (\tilde{cv} and \tilde{cv}^2) are also shown. For the

Table 5.1: Panel model comparison

| Model | cv | cv^2 | \tilde{cv} | \tilde{cv}^2 | cv | cv^2 | \tilde{cv} | \tilde{cv}^2 | Model | cv | cv^2 | \tilde{cv} | \tilde{cv}^2 |
|-----------------|------|--------|--------------|------------------|------|--------|--------------|----------------|-------------|------|--------|--------------|----------------|
| Panel A: Static | | | | Panel B: Dynamic | | | | Panel C: ECM | | | | | |
| W97 | | | | | | | | | | | | | |
| <i>bet</i> | 4.46 | 38.06 | 1.40 | 1.85 | 6.18 | 65.32 | 1.45 | 1.68 | <i>ecm1</i> | 5.74 | 74.46 | 1.34 | 1.91 |
| <i>pol</i> | 4.47 | 37.68 | 1.41 | 1.83 | 5.94 | 60.71 | 1.39 | 1.56 | <i>ecm2</i> | 5.90 | 73.60 | 1.38 | 1.89 |
| <i>tim</i> | 4.44 | 38.66 | 1.40 | 1.88 | 6.50 | 79.51 | 1.52 | 2.04 | <i>ecm3</i> | 5.40 | 58.71 | 1.26 | 1.51 |
| <i>cro</i> | 3.17 | 20.59 | 1.00 | 1.00 | 4.27 | 38.89 | 1.00 | 1.00 | | | | | |
| L68 | | | | | | | | | | | | | |
| <i>bet</i> | 4.96 | 45.79 | 1.38 | 1.80 | 6.65 | 76.22 | 1.33 | 1.51 | <i>ecm1</i> | 6.84 | 97.80 | 1.36 | 1.94 |
| <i>pol</i> | 4.98 | 45.31 | 1.39 | 1.78 | 6.39 | 70.36 | 1.27 | 1.40 | <i>ecm2</i> | 6.90 | 94.18 | 1.38 | 1.87 |
| <i>tim</i> | 5.04 | 47.87 | 1.40 | 1.88 | 7.16 | 96.10 | 1.43 | 1.91 | <i>ecm3</i> | 6.06 | 71.55 | 1.21 | 1.42 |
| <i>cro</i> | 3.59 | 25.48 | 1.00 | 1.00 | 5.02 | 50.34 | 1.00 | 1.00 | | | | | |
| D29 | | | | | | | | | | | | | |
| <i>bet</i> | 3.15 | 18.74 | 1.44 | 2.05 | 3.97 | 27.79 | 1.58 | 2.31 | <i>ecm1</i> | 3.15 | 19.73 | 1.25 | 1.64 |
| <i>pol</i> | 3.17 | 18.37 | 1.44 | 2.01 | 3.42 | 21.52 | 1.36 | 1.79 | <i>ecm2</i> | 3.53 | 25.34 | 1.40 | 2.10 |
| <i>tim</i> | 2.90 | 15.49 | 1.32 | 1.69 | 3.76 | 26.82 | 1.50 | 2.23 | <i>ecm3</i> | 3.83 | 28.62 | 1.52 | 2.38 |
| <i>cro</i> | 2.19 | 9.14 | 1.00 | 1.00 | 2.52 | 12.04 | 1.00 | 1.00 | | | | | |
| O26 | | | | | | | | | | | | | |
| <i>bet</i> | 2.87 | 14.45 | 1.38 | 1.83 | 3.33 | 19.26 | 1.41 | 1.80 | <i>ecm1</i> | 3.02 | 18.85 | 1.29 | 1.77 |
| <i>pol</i> | 2.87 | 14.48 | 1.38 | 1.83 | 3.09 | 16.31 | 1.31 | 1.53 | <i>ecm2</i> | 3.53 | 25.98 | 1.50 | 2.43 |
| <i>tim</i> | 2.66 | 12.23 | 1.28 | 1.55 | 3.31 | 20.16 | 1.41 | 1.89 | <i>ecm3</i> | 3.85 | 29.15 | 1.64 | 2.73 |
| <i>cro</i> | 2.08 | 7.90 | 1.00 | 1.00 | 2.35 | 10.67 | 1.00 | 1.00 | | | | | |
| O15 | | | | | | | | | | | | | |
| <i>bet</i> | 2.51 | 11.94 | 1.16 | 1.36 | 3.72 | 22.46 | 1.43 | 1.72 | <i>ecm1</i> | 3.54 | 24.80 | 1.36 | 1.89 |
| <i>pol</i> | 2.52 | 11.93 | 1.17 | 1.35 | 3.80 | 24.27 | 1.46 | 1.85 | <i>ecm2</i> | 4.12 | 34.57 | 1.59 | 2.64 |
| <i>tim</i> | 2.62 | 11.55 | 1.21 | 1.31 | 4.81 | 39.88 | 1.85 | 3.05 | <i>ecm3</i> | 4.27 | 36.64 | 1.64 | 2.80 |
| <i>cro</i> | 2.16 | 8.81 | 1.00 | 1.00 | 2.60 | 13.09 | 1.00 | 1.00 | | | | | |
| E11 | | | | | | | | | | | | | |
| <i>bet</i> | 3.30 | 17.96 | 1.69 | 2.70 | 3.02 | 15.94 | 1.49 | 2.16 | <i>ecm1</i> | 2.32 | 10.75 | 1.15 | 1.46 |
| <i>pol</i> | 2.94 | 14.63 | 1.51 | 2.20 | 2.90 | 16.36 | 1.44 | 2.22 | <i>ecm2</i> | 2.72 | 14.26 | 1.35 | 1.94 |
| <i>tim</i> | 2.72 | 12.17 | 1.39 | 1.83 | 3.52 | 20.09 | 1.74 | 2.73 | <i>ecm3</i> | 3.28 | 18.95 | 1.62 | 2.57 |
| <i>cro</i> | 1.96 | 6.65 | 1.00 | 1.00 | 2.02 | 7.37 | 1.00 | 1.00 | | | | | |
| E14 | | | | | | | | | | | | | |
| <i>bet</i> | 3.05 | 15.26 | 1.55 | 2.32 | 2.82 | 13.82 | 1.46 | 2.08 | <i>ecm1</i> | 2.30 | 10.22 | 1.20 | 1.54 |
| <i>pol</i> | 2.99 | 14.59 | 1.52 | 2.22 | 2.79 | 14.67 | 1.45 | 2.21 | <i>ecm2</i> | 2.68 | 13.37 | 1.39 | 2.02 |
| <i>tim</i> | 2.66 | 12.24 | 1.35 | 1.86 | 2.76 | 12.28 | 1.43 | 1.85 | <i>ecm3</i> | 3.15 | 17.43 | 1.64 | 2.63 |
| <i>cro</i> | 1.97 | 6.58 | 1.00 | 1.00 | 1.92 | 6.63 | 1.00 | 1.00 | | | | | |
| F16 | | | | | | | | | | | | | |
| <i>bet</i> | 2.55 | 11.61 | 1.43 | 1.99 | 2.47 | 9.69 | 1.37 | 1.69 | <i>ecm1</i> | 2.42 | 11.52 | 1.34 | 2.00 |
| <i>pol</i> | 2.61 | 11.53 | 1.46 | 1.98 | 2.55 | 12.02 | 1.42 | 2.09 | <i>ecm2</i> | 2.84 | 16.13 | 1.58 | 2.81 |
| <i>tim</i> | 2.34 | 9.85 | 1.30 | 1.69 | 2.94 | 15.71 | 1.63 | 2.73 | <i>ecm3</i> | 3.13 | 17.50 | 1.74 | 3.04 |
| <i>cro</i> | 1.79 | 5.82 | 1.00 | 1.00 | 1.80 | 5.75 | 1.00 | 1.00 | | | | | |

The table shows absolute and normalized CV criteria. In panels A (models in levels) and B (models in first differences), the considered implementations of panel models are the between (*bet*), pooled (*pol*), time (*tim*) and cross-section specific (*cro*) regression. Smallest CV estimates are normalized to unity. Results obtained in Panel C are for the ECMs where the CV estimates are normalized in the way that the corresponding CV estimates for the cross-section dependent regression in first differences is equal to unity.

purpose of normalization, CV estimates from cross-section specific model formalizations are set to unity. All models describing ΔI_{it} share the same benchmark model for normalization such that an immediate contrasting of ‘weakly dynamic’ models as (5.2) and ECMs is feasible. Cross comparison of the model families given in (5.1) and (5.2) is feasible by regarding (non-normalized) absolute CV estimates obtained from the benchmark (cross-section specific) models.

5.3.1 Static panel models

Concentrating on the model family (5.1) the overall evidence is that country specific panel models provide the most suitable framework to investigate the SI relation. This model class uniformly yields smallest CV estimates over all cross-sections. For the largest cross-section (W97) all remaining modeling approaches perform similarly poor in comparison with cross-section specific modeling given both normalized CV criteria. It turns out that the second best models, time specific regressions ($\tilde{c}v$) and the pooled regression ($\tilde{c}v^2$), are about 40% and 83% in excess of the corresponding estimates obtained from cross-section specific regressions. The CV results are also remarkable in the sense that time dependent regressions which allow a relatively large number of model parameters, namely 64 ($T = 32$), perform similar to the highly restricted pooled regression models encountering only two parameters. With regard to the relative performance of the cross-section specific regression against between regression, mean absolute forecast errors ($\tilde{c}v$) for the latter are between 16% (O15) and 69% (E11) worse.

5.3.2 Static vs. weakly dynamic models

As mentioned comparing the model families in (5.1) and (5.2) sheds light on the potential of mean reversion as a characteristic of domestic invest-

ment and saving ratios. Moreover, such a comparison hints at the prevalence of individual effects in (5.1) which are removed by differencing. For both model families cross-section specific model formalizations uniformly outperform the remaining panel based estimation schemes, i.e. between regression, time specific and pooled modeling. For F16, E14 and E11 both CV measures (cv and cv^2) yield only small numerical differences when comparing the performance of cross-section specific regressions for the levels and first differences of the domestic investment ratio. For all remaining cross-sections, however, CV estimates are clearly in favor of a specification explaining the investment ratio rather than its changes. Concentrating e.g. on mean absolute forecast errors, cross-section specific panel approaches to changes of the investment ratio yield cv estimates that are between 13% (O26) and 39.8% (L68) worse than corresponding statistics obtained for the level representation.

5.3.3 Weakly dynamic vs. error correction dynamics

Although model representations of changes of the investment ratio have been outperformed by level representations it is still interesting to address the issue of potential error correction dynamics. Comparing normalized CV estimates in Panels B and C of Table 5.1, it can be seen that none of the ECM model versions closely approaches the cross-section specific ‘weakly dynamic’ model $\Delta I_{it} = c_i + d_i \Delta S_{it} + e_{it}$. These results are the more surprising when recalling that the first three ECM versions are formalized conditional on the cross-section member. Overall mean absolute forecast errors obtained from cross-section specific ECMs are between 15% (E11, model *ecm1*) and 74% (F16, model *ecm3*) larger than the benchmark presuming absence of error correction dynamics. These results are also at odds with a presumption of cointegration linking the ratios of domestic saving and

investment over GDP. In case of cointegration just regressing ΔI_{it} on ΔS_{it} would suffer from statistical inefficiency owing to the neglect of the long run equilibrium relationship.

5.4 Summary

This chapter investigates the relation between domestic saving and investment for eight cross-sections covering the sample period 1971 to 2002. Cross-validation criteria are applied to compare different specifications of the SI relation. From static model performance it is found that the best performing parametric description of the SI relation is cross-section specific. As such, SI relations might be subject to other country specific economic conditions and policies than global or cross-sectional capital mobility. Contrasting static and weakly dynamic model formalizations no hint at the necessity of a weakly dynamic model specification is found. This evidence might be due to individual effects governing investment ratios. Moreover, adding an error correction term in dynamic models does not improve model performance. Supporting evidence for a cointegration relation between domestic saving and investment ratios is not found.

Chapter 6

Preliminary analysis

The previous analyses have reviewed theoretical and empirical approaches to model the SI relation, introduced the data and detected the country specific feature of the SI relation. Before the functional coefficient model is applied to investigate determinants of the SI relation, it is interesting to obtain some preliminary view of the SI relation from our data sets in this chapter¹. In the next section, standard between regressions as in FH (1980) are considered. Section 6.2 investigates alternatively pooled regressions. Time dependent regressions as adopted by Sinn (1992) and cross-section dependent regression as proposed by Obstfeld (1986) are considered in Section 6.3 and 6.4, respectively. The issue of factor dependence is addressed in Section 6.5. Section 6.6 summarizes the main findings.

6.1 Between regressions

By means of the between regression (2.1) applied by FH (1980) with annual data covering the period 1971 to 2002, the results in Panel A of Table 6.1 are

¹For a different version of this chapter see “A functional coefficient model view of the Feldstein-Horioka puzzle”, with Helmut Herwartz, University Kiel, Economics Working Paper 2007-14.

obtained. Panel A of Figure 6.1 and 6.2 present the corresponding scatter plots of the average saving ratio against the average investment ratio. Except ordinary least square (OLS) regression lines (solid), 45° lines (dashed) are plotted. Using the 45° line to present the zero current account position is firstly suggested by Tesar (1991). An observation above this line mirrors a deficit in the country, where the saving ratio is lower than the investment ratio on average. Analogously, an observation under this line indicates a surplus in the country. As can be seen from the first plot in Panel A of Figure 6.1, CA positions from W97 scattering around the 45° line. The sum of total CA imbalances from a large enough sample of countries all over the world should be near zero. Furthermore, this diagonal line also indicates a one to one SI relation. The smaller the difference between the regression and 45° line, the higher the estimated SI relation.

Table 6.1: Between regression

$$\bar{I}_i = a + b\bar{S}_i + u_i$$

| Samples | W97 | L68 | D29 | O26 | O15 | E11 | E14 | F16 |
|----------------------|------------------------|-----------------------|-----------------------|-----------------------|------------------------|------------------|------------------|-----------------------|
| Panel A: 1971 - 2002 | | | | | | | | |
| \hat{b} | 0.43 (11.24) | 0.40 (8.47) | 0.64 (7.16) | 0.59 (5.11) | 0.77 (7.96) | -0.16 (-0.71) | 0.13 (0.53) | 0.62 (4.44) |
| R^2 | 0.57 | 0.52 | 0.65 | 0.52 | 0.83 | 0.05 | 0.02 | 0.58 |
| Panel B: 1971 - 1986 | | | | | | | | |
| \hat{b} | 0.44 (9.44) | 0.39 (7.13) | 0.80 (6.85) | 0.69 (6.58) | 0.86 (10.43) | 0.18 (0.78) | 0.36 (1.46) | 0.66 (4.50) |
| R^2 | 0.48 | 0.44 | 0.63 | 0.64 | 0.89 | 0.06 | 0.15 | 0.59 |
| Panel C: 1987 - 2002 | | | | | | | | |
| \hat{b} | 0.38 (9.96) | 0.39 (7.80) | 0.42 (5.42) | 0.39 (3.27) | 0.65 (5.23) | -0.14 (-0.90) | -0.02 (-0.10) | 0.30 (2.24) |
| R^2 | 0.51 | 0.48 | 0.52 | 0.31 | 0.68 | 0.08 | 0.00 | 0.26 |

This table reports slope estimates from the between regressions of the average investment ratio on the average saving ratio. t -statistics appear in parentheses below the coefficient estimates. Coefficients which are significant at the 5% level are highlighted in bold face.

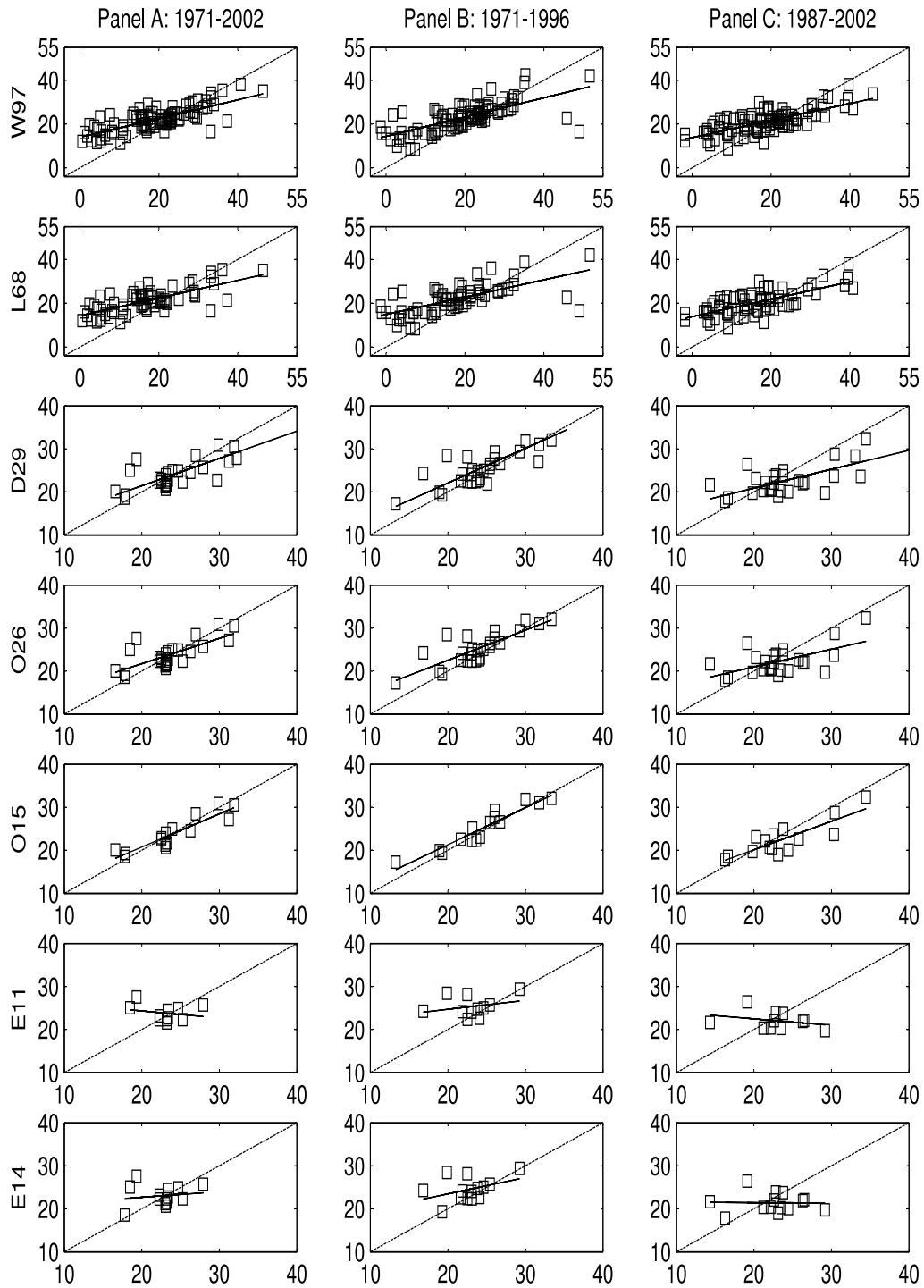


Figure 6.1: Average saving ratios vs. average investment ratios over the period 1971 to 2002 and the two equally sized subperiods for seven cross-sections. The dashed diagonal lines present the zero current account position. The solid lines illustrate OLS between regression lines.

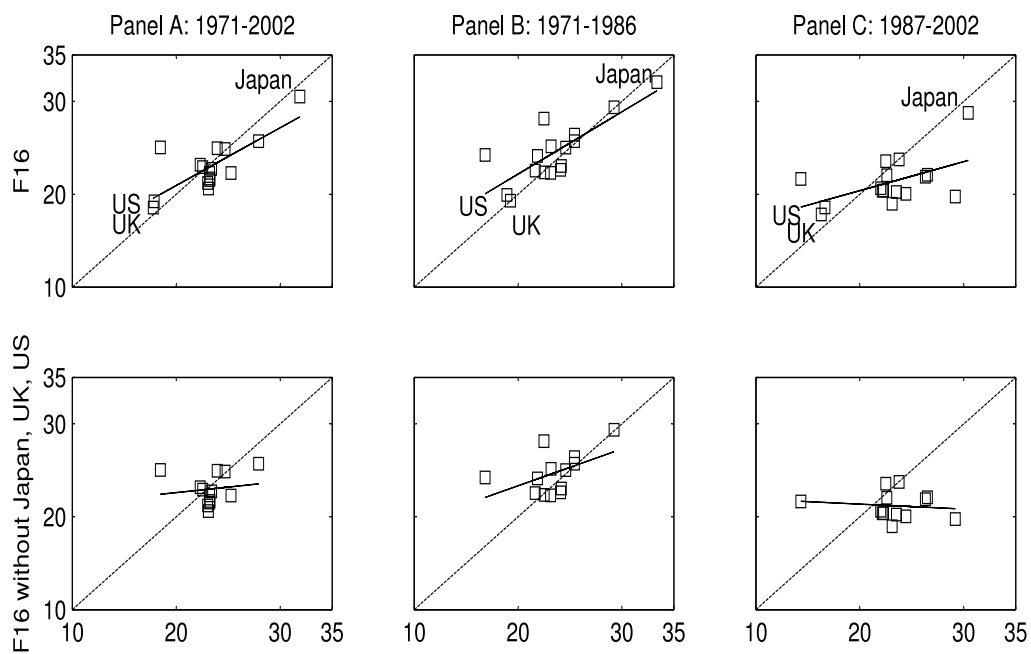


Figure 6.2: Average saving ratios vs. average investment ratios over the period 1971 to 2002 and the two equally sized subperiods. The first row is for FH16, the 16 OECD countries considered by FH(1980), and the second row is for the same sample but without Japan, UK, and US. The dashed diagonal lines present the zero current account position. The solid lines illustrate OLS between regression lines.

At first, it can be seen from Panel A of Figure 6.1 that the regression lines in W97 and L68 have similar slopes. Table 2.1 shows the corresponding estimates of 0.43 and 0.40. Following the arguments in FH (1980) a significantly positive SI relation in W97 and L68 is not surprising. Some patterns of capital market segmentation are likely over a group of 97 economies sampled from all over the world, and are even possible among less developed economies. From a global perspective capital mobility is limited 'on average'. When one moves the view from plots of W97 and L68 to those of D29, O26 and O15, it can be immediately observed that the difference between the regression and diagonal line becomes smaller. On one side, this supports the FH results that the SI relation is high among OECD countries, which seems puzzling given improved capital mobility. On the other side, it can be seen that the highest estimator of the SI relation (0.77) and the highest degrees of explanation (83%) are obtained for O15, where countries belong to OECD but not the Euro area. This evidence supports indirectly the view that capital mobility is higher among countries in the Euro area than those outside. The direct evidence can be found by observing the last two plots from E14 and E11 in Panel A of Figure 6.1. The corresponding regression lines are nearly flat and provide the largest distances to diagonal lines. Both cross-sections have insignificant estimates for the SI relation as can be seen in Table 6.1. As such, the EU and Euro effect are confirmed. Capital mobility is high among EU countries and even higher in the Euro area. This evidence is consistent with the interpretation of FH that SI relation signal the degree of capital mobility. In addition, considering F16, both the estimated SI relation (0.62) and the degree of explanation (0.58) are smaller for the period 1971 to 2002 compared to the corresponding results (0.887 and 0.91) given in FH (1980) for the period 1960 to 1974. This finding is in line with the presumption that capital mobility has increased over time. Moreover, the large country effect can be observed by removing Japan, the UK and the US from

F16. As can be seen in Figure 6.2, the regression lines become flatter when excluding these large countries. It could be shown that the corresponding estimated SI relation becomes insignificant.

Given the weakened evidence in favor of a large or even significant SI relation in more recent samples in comparison with FH (1980), it is sensible to check the robustness of the previous between regression results by means of two equally sized sub-samples, covering the periods 1971 to 1986 and 1987 to 2002, respectively. Between regression results obtained for these two subperiods are documented in Panel B and C of Table 6.1. The corresponding scatter plots are shown in Panel B and C of Figure 6.1 and 6.2. It can be seen that the estimated SI relation has decreased in all cross-sections. This evidence points to some variation of the SI relation and, moreover, is consistent with the generally improving integration of capital and goods markets. Furthermore, the between estimates are insignificantly different from zero in E14 and E11 for both subperiods. Although the SI relation in OECD economies (D29, O26, O15, F16) is still significantly different from zero, it is much smaller for the more recent period. The degree of explanation achieved by between regressions for the second subset is lower than for the first, and varies between 26% (F16) and 68% (O15) when E11 and E14 are excluded.

6.2 Pooled regressions

As an alternative to between regressions, pooled regressions can be applied to estimate SI relation at the aggregate level. Corresponding results provided in Table 6.2 are qualitatively identical to those reported for the between regressions in Table 6.1. The estimated SI relation is significantly positive in W97 and L68. It becomes much greater for OECD economies as D29, O26, O15, and F16. A decreasing estimator of the SI relation over time can

be observed in all eight considered cross-sections. The only difference in the results from pooled regressions as those from between regressions lies in the estimates for E14 and E11. For these two cross-sections significantly positive pooled estimates of SI relations are found although their values are small compared to estimates from other cross-sections.

Table 6.2: Pooled regression

$$I_{it} = a + bS_{it} + u_{it}$$

| Samples | W97 | L68 | D29 | O26 | O15 | E11 | E14 | F16 |
|----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|-----------------------|-----------------------|------------------------|
| Panel A: 1971 - 2002 | | | | | | | | |
| \hat{b} | 0.37 (38.51) | 0.34 (29.23) | 0.54 (24.62) | 0.58 (22.70) | 0.74 (26.33) | 0.25 (5.41) | 0.33 (7.81) | 0.54 (17.13) |
| R^2 | 0.32 | 0.28 | 0.40 | 0.38 | 0.59 | 0.08 | 0.12 | 0.37 |
| Panel B: 1971 - 1986 | | | | | | | | |
| \hat{b} | 0.39 (27.29) | 0.35 (20.35) | 0.73 (21.94) | 0.70 (21.05) | 0.81 (22.42) | 0.46 (7.34) | 0.56 (9.52) | 0.69 (16.86) |
| R^2 | 0.32 | 0.28 | 0.51 | 0.52 | 0.68 | 0.24 | 0.29 | 0.53 |
| Panel C: 1987 - 2002 | | | | | | | | |
| \hat{b} | 0.33 (27.51) | 0.33 (20.93) | 0.42 (17.41) | 0.42 (12.86) | 0.63 (16.13) | 0.05 (0.94) | 0.13 (2.81) | 0.36 (10.06) |
| R^2 | 0.33 | 0.29 | 0.40 | 0.29 | 0.52 | 0.01 | 0.03 | 0.28 |

This table reports slope estimates from the pooled regressions of the investment ratio on the saving ratio. t -statistics appear in parentheses below the coefficient estimates. Coefficients which are significant at the 5% level are highlighted in bold face.

In the light of the results from between and pooled regressions one may conjecture that the FH puzzle is not such a big puzzle anymore when concentrating on more recent time windows. In a similar vein, using data for 12 OECD countries from 1980 to 2001, Coakley et al. (2004) show the insignificance of the SI relation by means of nonstationary panel models. Blanchard and Giavazzi (2002) also document a small SI relation in a pooled regression for the EU area using the sample period 1991 to 2001. From a statistical as well as economic viewpoint, however, potential time variation of the SI relation provokes some subsequent issues. With regard to statistical aspects

it is not clear in how far conclusions offered by (misspecified) time homogeneous econometric models are spurious or robust under respecification of the model. From an economic perspective it is tempting to disentangle the economic forces behind the observed decreasing trend in the SI relation. With regard to this aspect it is of particular interest to separate deterministic time features from measurable economic factors driving the SI relation. The next subsection will underscore that the SI relation is likely not homogenous within the two subsamples considered so far but time varying throughout.

6.3 Time dependent SI relations

To address the time variation of the SI relation in some more detail, a sequence of (time specific) cross-sectional OLS regressions as proposed by Sinn (1992) is employed:

$$I_{it} = a_t + b_t S_{it} + u_{it}, \quad i = 1 \dots N. \quad (6.1)$$

Time varying slope estimates, \hat{b}_t , $t = 1971, \dots, 2002$, obtained from model (6.1) for the eight cross-sections are shown in Figure 6.3 jointly with corresponding 95% confidence intervals. Eyeball inspection confirms that the SI relation has decreased over time in all cross-sections. It has declined from 0.39 and 0.29 to 0.18 and 0.20 in W97 and L68 respectively; from a range between 0.72 and 0.86 to about 0.2 in O26, O15, and F16; from 0.63 to 0.01 in D29, from 0.54 to 0.1 in E14, and from 0.43 to zero in E11. Regarding E14 and E11 our results are in line with Blanchard and Giavazzi (2002). Besides, a sharp reduction of \hat{b}_t between 1975 and 1980 is found for F16, E14 and E11. This evidence might be due to the fact that many industrialized economies experienced current account deficits in this period. The latter might mirror the effects of oil price shocks in the late 1970s. According to Sachs (1981), however, changes in investment opportunities rather than oil price changes

could dominate the medium run behavior of current accounts in the 1970s.

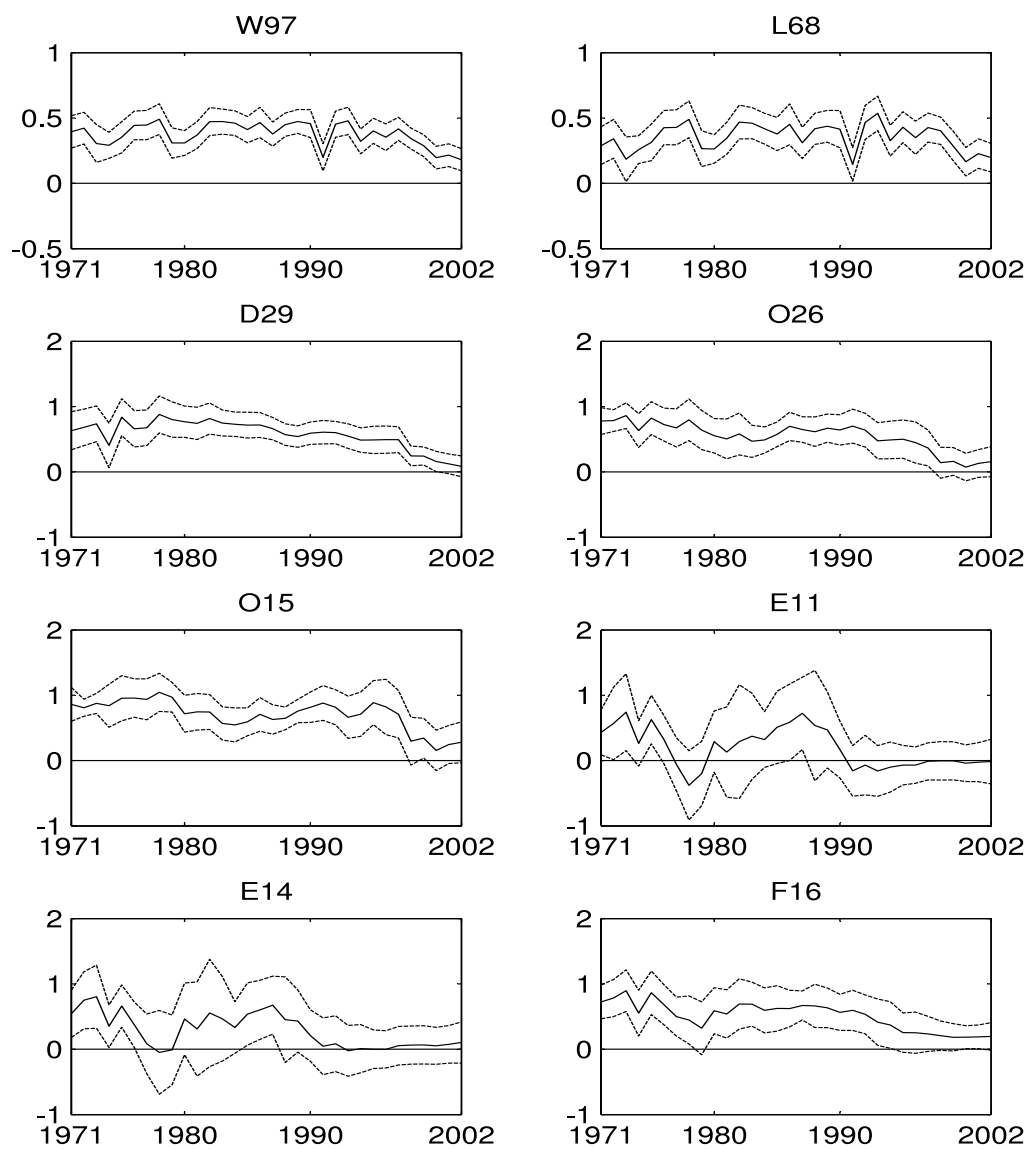


Figure 6.3: Estimated time varying SI relations obtained from model (6.1) for the eight cross-sections. The solid line with stars shows the point estimates and the two dashed lines are the corresponding 95% confidence intervals.

6.4 Cross-section dependent SI relations

Except time variation of the SI relation, cross-sectional variation of the SI relation has also been observed in the literature. Furthermore, the highly significant and quantitatively substantial country heterogeneity of SI relation have been diagnosed for the eight considered cross-sections in Chapter 5 via CV criteria. This feature is addressed via (cross-section specific) time dimensional OLS regressions as proposed by Obstfeld (1986),

$$I_{it} = a_i + b_i S_{it} + u_{it}, \quad t = 1 \dots T. \quad (6.2)$$

Figure 6.4 shows the slope estimates of this regression for the considered 97 countries with period from 1971 to 2002. They are quite heterogenous and varies approximately from -0.6 to 1.5. These estimates as well as their t -statistics can be seen in Table 6.3. When the two subperiods are used, no uniformly trending behaviour of SI relations can be observed.

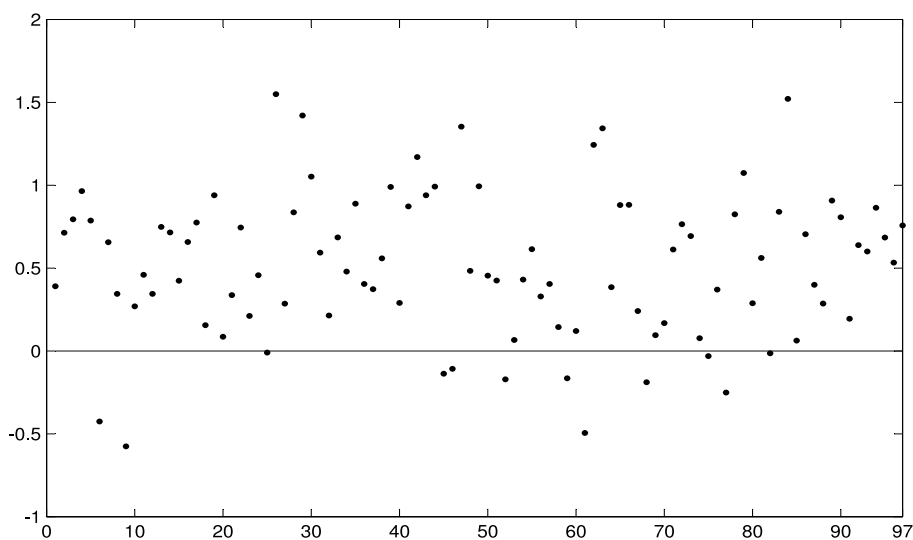


Figure 6.4: Point estimates of cross-section specific SI relations obtained from model (6.2) for the 97 countries in W97.

Table 6.3: Cross-section specific regression

$$I_{it} = a_i + b_i S_{it} + u_{it}$$

| Country | 71-02 | 71-86 | 87-02 | Country | 71-02 | 71-86 | 87-02 | Country | 71-02 | 71-86 | 87-02 |
|---------|------------------|------------------|------------------|---------|------------------|------------------|------------------|---------|------------------|------------------|------------------|
| DZA | 0.39 (2.29) | 0.40 (1.86) | -0.10 (-1.07) | GHA | 0.48 (1.34) | 0.63 (3.57) | 0.66 (1.77) | NGA | 0.24 (1.79) | 0.61 (3.53) | -0.27 (-1.55) |
| ARG | 0.71 (6.35) | 0.92 (7.92) | -0.30 (-1.54) | GRC | 0.89 (18.13) | 1.12 (13.43) | 0.64 (3.28) | NOR | -0.19 (-0.52) | -1.21 (-2.70) | -0.39 (-1.47) |
| AUS | 0.79 (8.88) | 0.49 (4.32) | 1.18 (7.02) | GTM | 0.40 (2.37) | 1.14 (4.64) | 0.06 (0.09) | PAK | 0.09 (1.03) | -0.26 (-0.75) | 0.08 (0.39) |
| AUT | 0.96 (11.84) | 0.98 (8.79) | 0.49 (2.77) | GUY | 0.37 (2.34) | 0.24 (1.27) | 0.84 (2.73) | PRY | 0.17 (1.60) | 1.04 (2.97) | 0.07 (1.27) |
| BGD | 0.79 (13.84) | 0.89 (5.32) | 0.78 (12.12) | HTI | 0.56 (3.00) | -0.21 (-0.90) | 0.83 (3.18) | PER | 0.61 (5.32) | 0.51 (3.04) | 1.04 (6.61) |
| BRB | -0.43 (-2.57) | -0.28 (-1.65) | -0.51 (-2.28) | HND | 0.99 (6.28) | 0.80 (3.65) | 0.74 (3.18) | PHL | 0.76 (4.72) | 1.54 (5.44) | -0.36 (-1.56) |
| BEL | 0.66 (5.34) | 0.77 (6.77) | 0.55 (3.12) | HKG | 0.29 (1.14) | 0.84 (1.87) | -0.21 (-0.71) | PRT | 0.69 (3.89) | 0.65 (2.40) | 0.72 (3.67) |
| BEN | 0.34 (2.59) | 0.36 (1.36) | 0.57 (4.86) | HUN | 0.87 (7.20) | 1.37 (3.10) | 0.58 (5.25) | RWA | 0.08 (1.70) | 0.37 (1.77) | 0.12 (3.34) |
| BWA | -0.58 (-4.50) | -0.60 (-2.38) | 0.16 (0.39) | ISL | 1.17 (7.61) | 1.37 (4.57) | -0.86 (-2.03) | SAU | -0.03 (-0.36) | -0.14 (-0.90) | 0.01 (0.16) |
| BRA | 0.27 (1.60) | 0.27 (0.51) | 0.29 (3.03) | IND | 0.94 (15.35) | 1.14 (8.64) | 0.92 (7.53) | SEN | 0.37 (5.73) | 0.25 (4.03) | 1.01 (5.24) |
| BFA | 0.46 (5.97) | 0.55 (4.23) | 0.81 (4.94) | IDN | 0.99 (6.54) | 0.40 (2.27) | 1.57 (10.94) | SGP | -0.25 (-1.91) | 0.30 (2.81) | 0.04 (0.13) |
| BDI | 0.34 (2.02) | 0.61 (1.93) | 0.43 (1.78) | IRL | -0.14 (-1.55) | -0.82 (-2.74) | 0.46 (7.43) | ZAF | 0.82 (7.91) | 0.25 (0.88) | 0.48 (4.87) |
| CMR | 0.75 (5.27) | 0.60 (4.44) | 0.77 (2.06) | ISR | -0.11 (-0.72) | 0.03 (0.11) | 0.33 (1.69) | SPA | 1.07 (6.25) | 1.11 (4.65) | 1.11 (3.14) |
| CAN | 0.72 (4.86) | 1.18 (3.18) | 0.40 (2.70) | ITA | 1.35 (6.49) | 0.90 (2.64) | -0.14 (-0.29) | LKA | 0.29 (0.97) | -0.07 (-0.13) | 0.92 (3.45) |
| CAF | 0.42 (4.89) | 0.58 (6.19) | 0.32 (2.31) | JAM | 0.48 (2.14) | 0.67 (2.13) | -0.29 (-1.27) | SUR | 0.56 (6.67) | 0.69 (6.20) | 0.12 (0.84) |
| CHL | 0.66 (7.73) | 0.66 (3.62) | 0.04 (0.17) | JPN | 0.99 (14.70) | 1.02 (7.25) | 0.88 (14.45) | SWZ | -0.01 (-0.14) | -0.16 (-1.24) | -0.06 (-0.53) |
| CHN | 0.77 (12.60) | 1.09 (9.86) | 0.65 (4.27) | KEN | 0.45 (4.18) | 0.35 (1.53) | 0.24 (2.16) | SWE | 0.84 (3.51) | 0.77 (4.36) | 0.96 (2.47) |
| COL | 0.16 (1.00) | -0.17 (-1.40) | 0.28 (1.13) | KOR | 0.42 (3.67) | 0.36 (2.53) | 0.91 (2.18) | SWI | 1.52 (9.69) | 1.24 (9.00) | 2.31 (11.99) |
| ZAR | 0.94 (7.81) | 0.61 (2.03) | 0.91 (7.83) | KWT | -0.17 (-4.77) | -0.30 (-4.89) | -0.30 (-7.93) | SYR | 0.06 (0.37) | 1.02 (2.11) | 0.13 (0.92) |
| COG | 0.09 (0.76) | 0.34 (2.55) | 0.07 (0.60) | LUX | 0.07 (1.18) | 0.08 (0.81) | -0.17 (-2.38) | THA | 0.70 (3.91) | 0.38 (1.39) | 1.93 (2.52) |
| CRI | 0.34 (2.55) | 0.09 (0.96) | 0.63 (2.93) | MDG | 0.43 (3.05) | -0.53 (-2.28) | 0.62 (5.67) | TGO | 0.40 (3.19) | 0.29 (1.18) | 0.45 (1.78) |
| CIV | 0.74 (4.55) | 0.23 (0.65) | 0.36 (4.09) | MWI | 0.61 (4.74) | 1.02 (2.49) | 0.47 (2.40) | TTO | 0.29 (2.25) | 0.05 (0.45) | 0.49 (2.07) |
| DNK | 0.21 (1.02) | 1.06 (5.46) | 0.35 (1.18) | MYS | 0.33 (2.05) | 0.20 (0.69) | -0.11 (-0.26) | TUN | 0.91 (3.32) | 0.46 (1.03) | 1.41 (5.59) |
| DOM | 0.46 (3.38) | 0.47 (3.90) | 0.52 (1.55) | MLI | 0.40 (4.39) | 0.09 (1.06) | 0.06 (0.28) | TUR | 0.81 (9.01) | 0.80 (2.60) | 0.59 (2.04) |
| ECU | -0.01 (-0.07) | -0.03 (-0.17) | -0.38 (-1.51) | MLT | 0.14 (1.34) | 0.05 (0.48) | 1.56 (3.15) | UGA | 0.19 (0.94) | 0.35 (2.78) | 0.80 (4.59) |
| EGY | 1.55 (5.95) | 1.35 (4.56) | 2.18 (5.98) | MRT | -0.16 (-1.39) | -0.24 (-1.56) | 0.40 (1.23) | UK | 0.64 (4.49) | 0.51 (1.40) | 1.25 (5.64) |
| SLV | 0.29 (3.09) | 0.66 (4.94) | -0.15 (-0.42) | MEX | 0.12 (1.00) | 0.08 (0.40) | 0.18 (1.16) | US | 0.60 (5.60) | 0.58 (3.04) | 0.65 (2.47) |
| FJI | 0.84 (5.00) | 0.60 (2.35) | 0.10 (1.01) | MAR | -0.49 (-2.26) | -0.83 (-1.93) | 0.38 (1.85) | URY | 0.86 (6.01) | 0.88 (4.58) | -0.07 (-0.36) |
| FIN | 1.42 (5.51) | 1.72 (5.03) | 0.77 (2.08) | MMR | 1.24 (8.15) | 1.36 (8.25) | 0.76 (4.42) | VEN | 0.68 (3.78) | 0.66 (2.51) | -0.57 (-2.04) |
| FRA | 1.05 (9.07) | 0.82 (9.47) | 1.47 (3.70) | NPL | 1.34 (7.76) | 1.19 (6.00) | 0.76 (4.76) | ZMB | 0.53 (6.43) | 0.54 (3.77) | -0.03 (-0.18) |
| GAB | 0.59 (3.80) | 0.50 (1.78) | -0.02 (-0.16) | NLD | 0.38 (2.37) | 0.83 (3.75) | 0.21 (0.85) | ZWE | 0.76 (6.24) | 0.76 (5.47) | 0.73 (3.30) |
| GMB | 0.21 (0.78) | 0.09 (0.15) | 0.18 (1.44) | NZL | 0.88 (3.09) | 0.42 (1.12) | 0.76 (3.55) | | | | |
| GER | 0.68 (4.37) | 0.77 (7.59) | 1.25 (3.53) | NER | 0.88 (6.34) | 0.81 (4.15) | 0.70 (4.07) | | | | |

This table reports slope estimates from the cross-section specific regressions of the investment ratio on the saving ratio. Period 1971 to 2002 as well as the two subperiods of 1971 to 1986 and 1987 to 2002 is considered. t -statistics appear in parentheses below the coefficient estimates. The abbreviation of countries can be found in the index of notation.

6.5 Factor dependence - profiles of estimated SI relations

Given the apparently decreasing trend and cross-sectional heterogeneity of the SI relation documented in the previous subsections, it is tempting to investigate if this heterogenous pattern is a purely deterministic feature or could be explained by (measurable) economic conditions. To provide a first view at the determinants of empirical SI relations the following cross-sectional regressions are considered:

$$\hat{b}_i = \gamma_0 + \gamma_1 \bar{w}_i + e_i, \quad i = 1, \dots, N. \quad (6.3)$$

In equation (6.3) $\bar{w}_i = 1/T \sum_{t=1}^T \tilde{w}_{it}$ with \tilde{w}_{it} being a measure of some factor characterizing the i -th member of the cross-section at time period t . The dependent variable \hat{b}_i in (6.3) is the slope estimate obtained from cross-section specific regressions (6.2). In case $\hat{\gamma}_1$ differs significantly from zero, this is regarded as an evidence to support the view that cross-sectional heterogeneity of the SI relation can be explained by the underlying factor. The regression in (6.3) takes a cross-sectional view at the determinants of the SI relation. This is justified in the light of country specific features of the SI relation diagnosed in Chapter 5.

As a particular caveat of the regression (6.3) one may point out that the dependent variables are not observed but (unbiased) estimators from some first step regression. As a more direct variant to detect factor dependence using observable regressands, one may regard the following model

$$I_{it} = a_i + \gamma_0 S_{it} + \gamma_1 \bar{w}_i S_{it} + u_{it}, \quad (6.4)$$

which is obtained from substituting $b_i = \gamma_0 + \gamma_1 \bar{w}_i$ in (6.2). The point estimates for the SI relation from the regression (6.4) are rather close to those obtained from (6.3). However, owing to potential cross-sectional heterogeneity and autocorrelation of model disturbances u_{it} in (6.4) estimation

Table 6.4: Factor dependence of the SI relation

$$\hat{\gamma}_1 \text{ in } \hat{b}_i = \gamma_0 + \gamma_1 \bar{w}_i + e_i$$

| Factor | W97 | L68 | D29 | O26 | O15 | E11 | F16 | E14 |
|-----------|--------------------------|--------------------------|--------------------------|--------------------------|-------------------|--------------------------|--------------------------|--------------------------|
| Group 1 | | | | | | | | |
| AGE | -0.006 (-2.51) | -0.003 (-0.77) | -0.013 (-1.11) | -0.026 (-2.41) | -0.018 (-1.48) | -0.068 (-2.57) | -0.052 (-2.52) | -0.060 (-2.26) |
| POPG | -0.131 (-3.12) | -0.067 (-1.07) | -0.300 (-2.28) | -0.183 (-1.23) | -0.086 (-0.49) | -1.109 (-2.07) | -0.186 (-0.75) | -0.755 (-1.44) |
| GDPC | 0.025 (0.92) | -0.089 (-2.18) | 0.088 (0.69) | 0.119 (1.01) | 0.105 (0.83) | 0.143 (0.38) | 0.033 (0.10) | -0.030 (-0.09) |
| GVBB | -0.020 (-1.68) | -0.012 (-0.99) | -0.056 (-1.95) | -0.013 (-0.37) | -0.017 (-0.21) | 0.009 (0.16) | -0.027 (-0.69) | -0.021 (-0.48) |
| GVTT | -0.003 (-1.10) | -0.007 (-2.59) | -0.000 (-0.07) | -0.004 (-0.56) | -0.005 (-0.58) | -0.013 (-0.93) | -0.008 (-1.00) | -0.015 (-1.18) |
| GVIVM | -0.035 (-2.03) | -0.018 (-1.05) | -0.076 (-1.06) | 0.009 (0.11) | 0.032 (0.34) | -0.126 (-0.79) | 0.012 (0.12) | -0.022 (-0.17) |
| GVCE | 0.001 (0.20) | -0.007 (-1.39) | -0.003 (-0.28) | -0.007 (-0.79) | -0.007 (-0.59) | -0.028 (-1.47) | -0.015 (-1.35) | -0.030 (-1.66) |
| GVCON | -0.016 (-1.98) | -0.032 (-3.99) | 0.010 (0.55) | -0.007 (-0.38) | -0.009 (-0.41) | -0.016 (-0.35) | -0.026 (-0.98) | -0.024 (-0.80) |
| Group 2 | | | | | | | | |
| OPN | -0.004 (-3.84) | -0.005 (-3.52) | -0.004 (-2.45) | -0.005 (-1.53) | 0.000 (0.06) | -0.009 (-2.93) | -0.006 (-2.10) | -0.009 (-2.68) |
| EXPT | -0.007 (-3.53) | -0.009 (-3.32) | -0.008 (-2.46) | -0.009 (-1.54) | -0.000 (-0.03) | -0.017 (-2.68) | -0.011 (-2.04) | -0.016 (-2.55) |
| IMPT | -0.009 (-4.05) | -0.009 (-3.55) | -0.008 (-2.43) | -0.010 (-1.50) | 0.002 (0.16) | -0.020 (-3.18) | -0.012 (-2.14) | -0.019 (-2.79) |
| INTD | -0.003 (-0.64) | 0.001 (0.25) | -0.022 (-0.32) | -0.021 (-0.32) | 0.015 (0.18) | -0.092 (-0.74) | -0.010 (-0.13) | -0.071 (-0.62) |
| Group 3 | | | | | | | | |
| LGDP | 0.063 (3.30) | 0.046 (1.66) | 0.075 (1.34) | 0.021 (0.36) | -0.024 (-0.34) | 0.137 (1.20) | 0.046 (0.65) | 0.130 (1.27) |
| <i>cr</i> | 1.985 | 1.996 | 2.045 | 2.056 | 2.131 | 2.201 | 2.120 | 2.145 |

OLS slope estimates for profile regressions (6.3). t -statistics appear in the parentheses below the coefficient estimates. Coefficient estimates which are significant at the 5% level are highlighted in bold face. cr denotes the t -critical value at 5%.

uncertainty might be easier to control in the regression (6.3). A preliminary view on the factor dependence of the SI relations can be obtained by the latter regressions.

Results from the profile regressions are reported in Table 6.4. Coefficient estimates which are significant at the 5% level are highlighted in bold face. For only two classes of factors, evidence for significant influence on the SI relation is found for at least five cross-sections. The first class comprises factors describing the goods market integration: the openness ratio, the export and import ratio. They have a significantly negative influence on the SI relation for both developing and developed economies. Compared to other countries, economies with a more integrated goods market might have a weaker SI relation. In contrast, significant influence from real interest rate differential measuring the financial market integration on the SI relation is not found. The second class composes the age dependency ratio. A negative impact of the age dependency ratio on the cross-sectional SI relation is diagnosed over 4 OECD samples and W97. The higher the ratio of dependents to the working-age population, the less is, *ceteris paribus*, the domestic saving. This might lead to the disconnection of domestic saving and investment. When performing a surface regression by regressing \hat{b}_i simultaneously on the age dependency ratio and the three factors denoting the goods market integration, it turns out that owing to multicollinearity only two factors remain to have significant explanatory power. Such surface regressions detect the age dependency ratio and one of the trade related measures (openness, exports, imports) to explain estimates \hat{b}_i significantly.

Apart from parameter significance discussed, the detected cross-section patterns are mostly uniform in the sense that the diverse profile regressions indicate the same direction of potential state variables affecting the SI relation. As further results it is worthwhile to point out that factors as population growth, country size or fiscal variables fail to describe the cross-

sectional pattern of SI relations significantly over more than two samples. This failure of significance, however, may also be addressed to a false presumption of time homogeneity of the SI relation or the factor state (\bar{w}_i) or both.

6.6 Summary

By means of between, pooled, time and cross-section specific regressions, time decreasing and country heterogeneous SI relations are found. To provide a preliminary view on the determinants of empirical SI relations, the estimated cross-section specific SI relation is regressed on potential economic factors. The age dependency ratio and the openness ratio are diagnosed to influence the cross-section specific SI relation significantly. Therefore, the panel-heterogeneous SI relation might be due to economic state variables rather than be a purely deterministic feature or some random processes. The next chapter introduces the functional coefficient model adopted to analyse factor dependence of the SI relation.

Chapter 7

Functional coefficient model

A useful extension of the classical linear regression model are functional (varying) coefficient models where model parameters may change with the value of other variables (factors). Functional specifications can be traced back to generalized linear models (Nelder and Wedderburn 1972). To formalize the functional coefficient, parametric representations such as finite order polynomials or Fourier expansions, or otherwise nonparametric approaches can be employed (Cleveland, Grosse and Shyu 1992, Hastie and Tibshirani 1993, Chen and Tsay 1993). A natural specification issue in these models is to infer whether the functional coefficient is actually constant. Modeling the functional coefficient parametrically, the constant form is compared with another parametric alternative. Modeling the functional coefficient nonparametrically, the alternative is the semiparametric regression. In this area of hypothesis testing, F -type tests are widely applied since the model of interest can easily be evaluated under both the null and the alternative hypothesis.

Furthermore, bootstrap based inference is broadly used if the asymptotic distribution of a test statistic is difficult or impossible to derive analytically. This may arise as a consequence of nuisance parameters affecting a test statistic, for instance, due to neglected heteroskedasticity. For the

case of (asymptotically) pivotal test statistics, moreover, particular bootstrap schemes might outperform first order asymptotic approximations by faster convergence to the nominal significance levels (Beran 1988, Hall and Titterington 1989). For a test statistic comparing the residual sum of squares from parametric and semiparametric functional regressions, Cai, Fan and Yao (2000) advocate a residual based bootstrap approach. Owing to possibly heterogenous error terms, residual based bootstrap inference could lack robustness. To deal with heteroskedasticity, pairs bootstrap (Freedman 1981) and wild bootstrap (Wu 1986, Liu 1988) can be applied.

In this chapter¹, a new, factor based bootstrap approach is proposed for the considered hypothesis that the functional coefficient is constant, i.e. factor independent. The underlying factors, governing parameter variation only under the alternative hypothesis, are drawn with replacement. The approach can cope with heteroskedasticity as it preserves the relationship between the error term variance and the corresponding regressors. Furthermore, in the framework of semiparametric regressions the factor based bootstrap could be more advantageous than wild, pairs or residual based bootstrap inference. This might be due to the fact that the former is likely better immunized against adverse effects of under- or oversmoothing in nonparametric regression than the latter approaches.

The properties of the residual based (RB), wild (WB), pairs (PB) and factor based bootstrap (FB) inference in functional coefficient models with finite samples are compared by means of a simulation study. The Monte Carlo exercises cover both parametric and semiparametric alternative hypotheses. Moreover, presuming a parametric pattern for the functional coefficient, the asymptotic behavior of the test statistic and the corresponding

¹For a different version of this chapter see "A new approach to bootstrap inference in functional coefficient models", with Helmut Herwartz, University Kiel, Economics Working Paper 2007-15.

factor based bootstrap approximation are investigated. At last, functional coefficient models are applied to the between regression of domestic investment on domestic saving considered by Feldstein and Horioka (1980).

The remainder of this chapter is organized as follows: The next section introduces the functional coefficient model and the considered test statistic when modeling the functional coefficient parametrically. In Section 7.2 the semiparametric case is introduced. The four alternative bootstrap approaches (RB, WB, PB and FB) are sketched in Section 7.3. Asymptotic results for the FB approximation are also shown in this Section. These approaches are compared by means of Monte-Carlo simulations in Section 7.4. In Section 7.5 the saving retention coefficient in between regressions is investigated via functional coefficient models. Section 7.6 summarizes briefly main findings and concludes.

7.1 Representation and assumptions for the parametric case

A functional coefficient regression can be given as

$$y_i = \sum_{k=1}^K \underline{\beta}_k(\omega_i) x_{ik} + e_i, \quad i = 1, \dots, N, \quad (7.1)$$

where K and N denote the number of regressors and available observations, respectively, with K assumed to be fixed. In model (7.1) y_i depends in a possibly nonlinear fashion on explanatory variables x_{i1} to x_{iK} , including a constant. A single coefficient $\underline{\beta}_k$ is a function of a random variable ω_i , called the factor henceforth. Let the functional coefficient have the form $\underline{\beta}_k(\omega_i) = \beta_{1k} + \beta_{2k} f_k(\omega_i)$, where $f_k(\omega_i)$ is assumed to be known. Then, given the parametric representation, (7.1) can be estimated by applying OLS to the

regression

$$y_i = \sum_{k=1}^K (\beta_{1k}x_{ik} + \beta_{2k}x_{ik}f_k(\omega_i)) + e_i.$$

Thus, model (7.1) can be written as

$$y_i = \mathbf{z}'_{i1}\boldsymbol{\beta}_1 + \mathbf{z}'_{i2}\boldsymbol{\beta}_2 + e_i, \quad (7.2)$$

$$= \mathbf{z}'_i\boldsymbol{\beta} + e_i, \quad (7.3)$$

where $\mathbf{z}'_{i1} = (x_{i1}, \dots, x_{iK})$, $\mathbf{z}'_{i2} = (f_1(\omega_i)x_{i1}, \dots, f_K(\omega_i)x_{iK})$, $\boldsymbol{\beta}_1 = (\beta_{11}, \dots, \beta_{1K})'$, $\boldsymbol{\beta}_2 = (\beta_{21}, \dots, \beta_{2K})'$, $\mathbf{z}'_i = (\mathbf{z}'_{i1}, \mathbf{z}'_{i2})$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2)'$ is a $2K$ -dimensional parameter vector. The corresponding matrix forms are

$$\mathbf{y} = \mathbf{Z}_1\boldsymbol{\beta}_1 + \mathbf{Z}_2\boldsymbol{\beta}_2 + \mathbf{e}, \quad (7.4)$$

$$= \mathbf{Z}\boldsymbol{\beta} + \mathbf{e}, \quad (7.5)$$

where $\mathbf{y} = (y_1, \dots, y_N)'$, $\mathbf{Z}'_1 = (\mathbf{z}_{11}, \dots, \mathbf{z}_{N1})$, $\mathbf{Z}'_2 = (\mathbf{z}_{12}, \dots, \mathbf{z}_{N2})$, $\mathbf{Z}' = (\mathbf{z}_1, \dots, \mathbf{z}_N)$ and $\mathbf{e} = (e_1, \dots, e_N)'$. The following assumptions concerning the functional coefficient model are made:

(A1) Factor variables $\omega_1, \omega_2, \dots, \omega_N$ are independent, identically distributed;

The error terms e_1, \dots, e_N are independent with mean zero and variance σ_i^2 ; $E(e_i^4) < \infty$; ω_i and e_i are independent.

(A2) $f_k(\omega_i), \forall k \in [1, K]$ is a real bounded continuous function;

$E[f_k(\omega_i)] = 0$ and $E[f_k^4(\omega_i)] < \infty, \forall k \in [1, K]$; $E[f_k(\omega_i)f_l(\omega_i)] < \infty, \forall k, l \in [1, K]$;

(A3) The matrix \mathbf{Z}_1 is fixed; $\lim_{N \rightarrow \infty} \mathbf{Z}'_1\mathbf{Z}_1/N = \Sigma_{11}$, a finite, positive definite matrix;

(A4) $\text{plim} \mathbf{Z}'_2\mathbf{Z}_2/N = \Sigma_{22}$, a finite, positive definite matrix;

(A5) $\text{plim} \left(\frac{1}{N} \sum_{i=1}^N e_i^2 \mathbf{z}_{i1} \mathbf{z}'_{i1} \right) = V_1$, a finite, positive definite matrix;

$\text{plim} \left(\frac{1}{N} \sum_{i=1}^N e_i^2 \mathbf{z}_{i2} \mathbf{z}'_{i2} \right) = V_2$, a finite, positive definite matrix;

$$(A6) \text{ plim} \left(\frac{1}{N} \mathbf{e}' \mathbf{e} \right) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sigma_i^2 = \sigma^2.$$

For the later purposes of resampling within the functional coefficient model, (A1) ensures that drawing from $\{\omega_i\}_{i=1}^N$ with replacement generates bootstrap samples of factors sharing their distribution with the factor sample. According to continuity stated in (A2), the random variable $f_k(\omega_i)$ is also independent and identically distributed. Furthermore, assuming $E[f_k(\omega_i)] = 0$ is without loss of generality. To see this, suppose that the functional coefficient can be formulated as $\underline{\beta}_k(\omega_i) = \beta_{1k}'' + \beta_{2k}'' f_k''(\omega_i)$, where $E[f_k''(\omega_i)] = \mu_k \neq 0$. Then, a corresponding mean zero functional form can always be found, i.e.

$$\begin{aligned} \underline{\beta}_k(\omega_i) &= \beta_{1k}'' + \beta_{2k}'' (f_k''(\omega_i) - \mu_k) + \beta_{2k}'' \mu_k \\ &= \beta_{1k} + \beta_{2k} f_k(\omega_i), \end{aligned}$$

with $\beta_{1k} = \beta_{1k}'' + \beta_{2k}'' \mu_k$, $\beta_{2k} = \beta_{2k}''$, $f_k(\omega_i) = f_k''(\omega_i) - \mu_k$ and $E[f_k(\omega_i)] = 0$. In addition, while (A3) defines the limiting behavior for the fixed matrix \mathbf{Z}_1 , (A4) to (A6) state that the weak law of large numbers applies for the corresponding stochastic elements.

To determine the asymptotic properties of the OLS estimator,

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{Z}' \mathbf{e},$$

we consider the limiting distribution of $\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, i.e.

$$\begin{pmatrix} \sqrt{N}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \\ \sqrt{N}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2) \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1' \mathbf{Z}_1 / N & \mathbf{Z}_1' \mathbf{Z}_2 / N \\ \mathbf{Z}_2' \mathbf{Z}_1 / N & \mathbf{Z}_2' \mathbf{Z}_2 / N \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Z}_1' \mathbf{e} / \sqrt{N} \\ \mathbf{Z}_2' \mathbf{e} / \sqrt{N} \end{pmatrix}.$$

Based on Assumption (A2) and the law of large numbers for the independent and heterogeneous random variables, it can be shown that $\text{plim} \mathbf{Z}_1' \mathbf{Z}_2 / N = 0$. Furthermore, by Assumptions (A1), (A5) and the Lindeberg-Feller central limit theorem,

$$\frac{1}{\sqrt{N}} \mathbf{Z}_1' \mathbf{e} \xrightarrow{d} N(0, V_1) \text{ and } \frac{1}{\sqrt{N}} \mathbf{Z}_2' \mathbf{e} \xrightarrow{d} N(0, V_2).$$

Therefore, as $N \rightarrow \infty$,

$$\sqrt{N}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \xrightarrow{d} N(0, \Sigma_{11}^{-1} V_1 \Sigma_{11}^{-1}) \text{ and } \sqrt{N}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2) \xrightarrow{d} N(0, \Sigma_{22}^{-1} V_2 \Sigma_{22}^{-1}),$$

which hold by (A3) and (A4). If the error terms in (7.1) are homoskedastic, $E(e_i^2 | \mathbf{z}_i) = E(e_i^2) = \sigma^2$, then $V_1 = \sigma^2 \Sigma_{11}$, $V_2 = \sigma^2 \Sigma_{22}$ and, thus,

$$\sqrt{N}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1) \xrightarrow{d} N(0, \sigma^2 \Sigma_{11}^{-1}), \quad \sqrt{N}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2) \xrightarrow{d} N(0, \sigma^2 \Sigma_{22}^{-1}).$$

For testing a standard regression model against the functional alternative, the following pair of hypotheses is formalized:

$$H_0 : \boldsymbol{\beta}_2 = 0 \quad \text{vs.} \quad H_1 : \boldsymbol{\beta}_2 \neq 0. \quad (7.6)$$

The considered goodness-of-fit test statistic compares the fitting accuracy of the empirical model under both hypotheses,

$$T_N = (\text{RSS}_0 - \text{RSS}_1) / (\text{RSS}_1 / N). \quad (7.7)$$

The underlying quantities in (7.7), RSS_i , $i = 0, 1$, are obtained from restricted (H_0) and unrestricted (H_1) OLS regressions respectively, i.e.

$$\text{RSS}_0 = \sum_{i=1}^N \left(y_i - \mathbf{z}'_{i1} \hat{\boldsymbol{\beta}}_1^{(0)} \right)^2 \text{ and } \text{RSS}_1 = \sum_{i=1}^N \left(y_i - \mathbf{z}'_{i1} \hat{\boldsymbol{\beta}}_1^{(1)} - \mathbf{z}'_{i2} \hat{\boldsymbol{\beta}}_2^{(1)} \right)^2. \quad (7.8)$$

Under the null hypothesis T_N is asymptotically distributed as $\chi^2(K)$ if disturbances are homoskedastic, but fails asymptotic pivotalness in case of heteroskedasticity. A heteroskedasticity consistent test statistic (White 1980) is

$$T_N^{hc} = \Delta' \left(R(\mathbf{Z}'\mathbf{Z})^{-1} \left(\sum_{i=1}^N (\hat{e}_i^{(1)})^2 \mathbf{z}_i \mathbf{z}'_i \right) (\mathbf{Z}'\mathbf{Z})^{-1} R' \right)^{-1} \Delta, \quad (7.9)$$

$\Delta = (R\hat{\boldsymbol{\beta}}^{(1)} - R\boldsymbol{\beta})$, $R = (0_{K \times K}, I_K)$ with $0_{K \times K}$ and I_K denoting a $K \times K$ matrix of zeros and the K -dimensional identity matrix, respectively. Under the null hypothesis T_N^{hc} converges in distribution to $\chi^2(K)$. Note that

the representation in (7.9) corresponds to an equivalent formulation of (7.7), namely

$$T_N = \Delta'(R(\mathbf{Z}'\mathbf{Z})^{-1}R')^{-1}\Delta/(\hat{\mathbf{e}}^{(1)'}\hat{\mathbf{e}}^{(1)}/N),$$

where all estimates are obtained from unrestricted (H_1) OLS regressions.

7.2 Semiparametric case

In case of an unknown or unspecified functional form $\underline{\beta}_k(\omega_i)$, the model in (7.1) is a semiparametric regression by observation and in matrix notation, respectively,

$$\begin{aligned} y_i &= \mathbf{z}'_{i1}\underline{\beta}(\omega) + e_i, \quad \underline{\beta}(\omega) = (\underline{\beta}_1(\omega), \dots, \underline{\beta}_K(\omega))' \\ \mathbf{y} &= \mathbf{Z}_1\underline{\beta}(\omega) + \mathbf{e}. \end{aligned}$$

A semiparametric estimator similar to the Nadaraya-Watson estimator (Nadaraya 1964, Watson 1964) is

$$\hat{\underline{\beta}}(\omega) = (\mathbf{Z}'_1\mathbf{W}\mathbf{Z}_1)^{-1}\mathbf{Z}'_1\mathbf{W}\mathbf{y}. \quad (7.10)$$

In (7.10) ω is a local point, $\mathbf{W} = \text{diag}\{K_h(\omega_1 - \omega), \dots, K_h(\omega_N - \omega)\}$, $K_h(\cdot) = K(\cdot/h)/h$, with $K(\cdot)$ being a kernel function and h the bandwidth parameter.

To distinguish the functional from a constant coefficient model by means of resampling the T_N statistic in (7.7), the following hypotheses are contrasted,

$$H_0 : \underline{\beta}(\omega) = \underline{\beta}_1 \quad \text{vs.} \quad H_1 : \underline{\beta}(\omega) \neq \underline{\beta}_1.$$

In analogy to (7.8), a semiparametric estimator of RSS_1 is

$$\text{RSS}_1 = \sum_{i=1}^N (y_i - \mathbf{z}'_{i1}\hat{\underline{\beta}}(\omega_i))^2. \quad (7.11)$$

Different as in the parametric case, the asymptotic behavior of T_N is non-standard even for homoskedastic error terms.

It is worthwhile mentioning that in case the form of $f_k(\omega_i)$ is unknown, Fourier series or other parametric approximations might be used instead. Goodness-of-fit tests may offer higher power when testing against specific alternatives. However, if the alternative is misspecified, contrasting two parametric forms might be inferior to testing against a semiparametric alternative owing to potential inconsistency (Horowitz and Härdle 1994).

7.3 Bootstrap procedures

In this section the factor based bootstrap (FB) is introduced to test the pair of hypotheses in (7.6) by means of T_N in (7.7). For a comprehensible comparison between the new procedure and other common bootstrap techniques, residual based (RB), wild (WB) and pairs bootstrap (PB) inference are also briefly sketched.

7.3.1 Residual based, pairs and wild bootstrap

Residual based bootstrap

The idea of random resampling with replacement, the so called “bootstrap”, was originally proposed by Efron (1979) who has also suggested a residual based resampling scheme for regression models. To approximate the distribution of T_N , RB consists of the following steps:

- 1) Bootstrap residuals $\{e_i^*\}_{i=1}^N$ are drawn with replacement from centered residuals $\{\hat{e}_i - \sum_{i=1}^N \hat{e}_i/N\}_{i=1}^N$ with \hat{e}_i obtained under H_1 ,

$$\hat{e}_i = y_i - \mathbf{z}'_{i1} \hat{\boldsymbol{\beta}}_1^{(1)} - \mathbf{z}'_{i2} \hat{\boldsymbol{\beta}}_2^{(1)}.$$

- 2) A bootstrap sample $\{y_i^*\}_{i=1}^N$ consistent with H_0 is $y_i^* = \mathbf{z}'_{i1} \hat{\boldsymbol{\beta}}_1^{(0)} + e_i^*$.

- 3) Given the sample $\{y_i^*, z_{i1}, \omega_i\}_{i=1}^N$, RSS_0^* and RSS_1^* are calculated as those in (7.8) to determine the bootstrap test statistic T_N^* .
- 4) Steps 1) to 3) are performed R times with R chosen sufficiently large. If T_N exceeds the $(1 - \alpha)$ -quantile of $\{T_N^{*(r)}\}_{r=1}^R$ H_0 is rejected with significance level α .

As an alternative to step 1) one may condition the resampling on residuals \hat{e}_i estimated under H_0 . Intuitively, resampling from restricted residuals may offer improved size features in finite samples while it bears the risk of power loss under the alternative hypothesis. For the Monte Carlo analysis provided in Section 7.4 both restricted and unrestricted residual estimates are considered, no systematic differences in finite sample size and power are diagnosed.

Wild bootstrap

Implemented via sampling with replacement RB builds upon an iid assumption to hold for error terms e_i . Consequently, RB fails to mimic distributional features of nonspherical disturbances. For the case of heteroskedasticity, WB is widely used. This bootstrap approach goes back to a proposal of Wu (1986) and has been established by Liu (1988) to evaluate the asymptotic distribution of studentized statistics in static linear regression models under heteroskedasticity of unknown form. The first step of WB generates the bootstrap residuals $\{e_i^*\}_{i=1}^N$ from unrestricted model estimates as

$$e_i^* = \hat{e}_i v_i, \hat{e}_i = y_i - \mathbf{z}'_{i1} \hat{\boldsymbol{\beta}}_1^{(1)} - \mathbf{z}'_{i2} \hat{\boldsymbol{\beta}}_2^{(1)},$$

where the random variable v_i is independent of the data, $E(v_i) = 0$ and $E(v_i^2) = 1$. There are numerous variants to specify the distribution of v_i

that basically differ with regard to the finite order moments of \hat{e}_i mimicked by the WB (Liu 1988, Mammen 1993). We choose one of the simplest, the so-called Rademacher distribution as recommended by Davidson and Flachaire (2001),

$$v_i = \begin{cases} -1 & \text{with probability } 0.5 \\ 1 & \text{with probability } 0.5. \end{cases}$$

Note that owing to $E[v_i^4] = 1$ also fourth order features of \hat{e}_i are retained by this particular implementation of WB. Once $\{e_i^*\}_{i=1}^N$ is obtained, the remaining steps are the same as those for RB.

Pairs bootstrap

Pairs bootstrap as an alternative avenue to address heteroskedastic error distributions has been originally advocated by Freedman (1981). Opposite to RB and WB, PB processes tuples of observed dependent and explanatory variables:

- 1) Bootstrap samples $\{y_i^*, z_{i1}^*, \omega_i^*\}_{i=1}^N$ are drawn with replacement from $\{y_i - z'_{i1}\hat{\beta}_1^{(1)} - z'_{i2}\hat{\beta}_2^{(1)} + z'_{i1}\hat{\beta}_1^{(0)}, z_{i1}, \omega_i\}_{i=1}^N$.
- 2) A bootstrap variant of T_N is then obtained from restricted and unrestricted estimation using the bootstrap sample. Given the sample $\{y_i^*, z_{i1}^*, \omega_i^*\}_{i=1}^N$, RSS_0^* and RSS_1^* are calculated as those in (7.8) to determine the bootstrap test statistic T_N^* .
- 3) Once T_N^* is obtained, the iteration and test decision are as described in step 4) of RB.

Note that dependent variables y_i^* are generated from the estimated dependent variable under the null hypothesis ($z'_{i1}\hat{\beta}_1^{(0)}$) plus the estimated error terms under the alternative ($y_i - z'_{i1}\hat{\beta}_1^{(1)} - z'_{i2}\hat{\beta}_2^{(1)}$). As such, the distribution of the bootstrap test statistic T_N^* can approximate the distribution

of T_N under the null hypothesis even if the alternative model holds. For the F -test statistic, Mammen (1993) shows that its distribution under the null hypothesis coincides asymptotically with the corresponding WB and PB counterparts under some regularity conditions.

7.3.2 Factor based bootstrap

FB proceeds from the perspective that under H_0 the factor variable does not impact on the parameters of interest. Opposite to RB, WB and PB, FB does not process first step residual estimates. The following steps are involved:

- 1) Bootstrap factor variables $\{\omega_i^*\}_{i=1}^N$ are drawn with replacement from $\{\omega_i\}_{i=1}^N$.
- 2) Based on the sample $\{y_i, z_{i1}, \omega_i^*\}_{i=1}^N$, RSS_0^* and RSS_1^* are calculated as those in (7.8) and the bootstrap statistic T_N^* is obtained. Note that in this case, RSS_0^* and RSS_0 are identical since ω_i has no influence on y_i under H_0 .
- 3) Once T_N^* is obtained the iteration and test decision are as described in step 4) of RB.

Opposite to PB, the FB scheme leaves the dependent and explanatory variables unchanged, i.e. the variables in y_i and z_{i1} are fixed throughout. As a further distinction in comparison with PB, FB samples $z_{i2}^* = (f_1(\omega_i^*)x_{i1}, \dots, f_K(\omega_i^*)x_{iK})'$ differ numerically from the population quantities while PB samples $z_{i2}^* = (f_1(\omega_i^*)x_{i1}^*, (f_1(\omega_i^*)x_{i1}^*, \dots, f_K(\omega_i^*)x_{iK}^*))'$ retain the 'numerical support' of $z_{i2} = (f_1(\omega_i)x_{i1}, \dots, f_K(\omega_i)x_{iK})'$.

A unique property of the FB based scheme is that it only works under $H_0 : \beta_2 = 0$. For the original data, the considered model is

$$\mathbf{y} = \mathbf{Z}_1\boldsymbol{\beta}_1 + \mathbf{Z}_2\boldsymbol{\beta}_2 + \mathbf{e}, \quad (7.12)$$

where \mathbf{Z}_1 is fixed, and \mathbf{Z}_2, e are stochastic. For the bootstrap data, the corresponding model is,

$$\mathbf{y} = \mathbf{Z}_1\beta_1 + \mathbf{Z}_2^*\beta_2^* + e, \quad (7.13)$$

where \mathbf{Z}_1 and e are fixed, and \mathbf{Z}_2^* is stochastic. Under the null hypothesis, $\beta_2 = 0$, the data generating process for the bootstrap data is the same as for the original data and $\beta_2^* = 0$. To distinguish the OLS estimator for β_1 in (7.13) from the one in (7.12), the former is denoted as $\hat{\beta}_1^*$. OLS residual estimates are denoted \hat{e}^* . Consider now the limiting behavior of the OLS estimates under the bootstrap space,

$$\begin{pmatrix} \sqrt{N}(\hat{\beta}_1^* - \beta_1) \\ \sqrt{N}(\hat{\beta}_2^* - \beta_2^*) \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1'\mathbf{Z}_1/N & \mathbf{Z}_1'\mathbf{Z}_2^*/N \\ \mathbf{Z}_2^{*'}\mathbf{Z}_1/N & \mathbf{Z}_2^{*'}\mathbf{Z}_2^*/N \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Z}_1'e/\sqrt{N} \\ \mathbf{Z}_2^{*'}e/\sqrt{N} \end{pmatrix}.$$

Before investigating the asymptotic properties of the FB approximations, it is useful to introduce the following Lemmas.

Lemma 1 *Let $\omega_1, \omega_2, \dots$ be independent, with common distribution F , and F_N be the empirical distribution of $\omega_1, \dots, \omega_N$. As $N \rightarrow \infty$, $F_N \rightarrow F$ for all continuity points of F . Given $\{\omega_i\}_1^N, \omega_1^*, \dots, \omega_N^*$ are conditionally independent, with common distribution F_N .*

PROOF. ω_i^* is obtained by drawing with replacement from the sample $\{\omega_i\}_1^N$.

Lemma 2 *As $N \rightarrow \infty$,*

- a) $(\mathbf{Z}_1'\mathbf{Z}_2^*/N) \xrightarrow{p} 0$;
- b) $(\mathbf{Z}_2^{*'}\mathbf{Z}_2^*/N) \xrightarrow{p} \Sigma_{22}$;
- c) $\mathbf{Z}_2^{*'}e/\sqrt{N} \xrightarrow{d} N(0, V_2)$.

PROOF. a) Define an element in $\mathbf{z}_{i1}\mathbf{z}_{i2}^{*'}'$ as $x_{il}x_{ik}f_k(\omega_i^*)$ with $l, k \in [1, K]$. Since $f_k(\cdot), \forall k \in [1, K]$, is a bounded, continuous function,

$$E(x_{il}x_{ik}f_k(\omega_i^*)) = x_{il}x_{ik}E(f_k(\omega_i^*)) \xrightarrow{p} x_{il}x_{ik}E(f_k(\omega_i)) = 0,$$

by Lemma 1 and Assumption (A1) and (A2). Similarly,

$$E(x_{il}^2 x_{ik}^2 f_k^2(\omega_i^*)) = x_{il}^2 x_{ik}^2 E(f_k^2(\omega_i^*)) \xrightarrow{p} x_{il}^2 x_{ik}^2 E(f_k^2(\omega_i)) < \infty.$$

Therefore, according to the law of large numbers for independent and heterogeneous random variables, $\sum_{i=1}^N x_{il} x_{ik} f_k(\omega_i^*)/N \xrightarrow{p} 0$, and a) is proved.

b) Apply analogous arguments by defining an element in $\mathbf{z}_{i2}^* \mathbf{z}_{i2}'$ as $x_{il} x_{ik} f_l(\omega_i^*) f_k(\omega_i^*)$ with $l, k \in [1, K]$. As $N \rightarrow \infty$,

$$\begin{aligned} E(x_{il} x_{ik} f_l(\omega_i^*) f_k(\omega_i^*)) &= x_{il} x_{ik} E(f_l(\omega_i^*) f_k(\omega_i^*)) \xrightarrow{p} x_{il} x_{ik} E(f_l(\omega_i) f_k(\omega_i)); \\ E(x_{il}^2 x_{ik}^2 f_l^2(\omega_i^*) f_k^2(\omega_i^*)) &= x_{il}^2 x_{ik}^2 E(f_l^2(\omega_i^*) f_k^2(\omega_i^*)), \\ &\xrightarrow{p} x_{il}^2 x_{ik}^2 E(f_l^2(\omega_i) f_k^2(\omega_i)), \\ &\leq x_{il}^2 x_{ik}^2 E^{1/2}(f_l^4(\omega_i)) E^{1/2}(f_k^4(\omega_i)) < \infty. \end{aligned}$$

The last two inequalities are obtained via Hölder's inequality and Assumption (A2) respectively. By the law of large numbers for independent and heterogeneous random variables, $\sum_{i=1}^N x_{il} x_{ik} f_l(\omega_i^*) f_k(\omega_i^*)/N \xrightarrow{p} \sum_{i=1}^N x_{il} x_{ik} E(f_l(\omega_i) f_k(\omega_i))/N$. The latter expression is the element in l -th row and k -th column of the matrix Σ_{22} by Assumption (A4), and, thus, b) is proved.

c) Conditional on the sample $\{y_i, \mathbf{z}_{i1}\}_{i=1}^N$, the corresponding $\{e_i\}_{i=1}^N$ for the bootstrap data are fixed. It follows then from Lemma 1 and Assumption (A1) and (A2) that

$$\begin{aligned} E(\mathbf{z}_{i2}^* e_i) &= [x_{i1} e_i E(f_1(\omega_i^*)), \dots, x_{iK} e_i E(f_K(\omega_i^*))]' \\ &\xrightarrow{p} [x_{i1} e_i E(f_1(\omega_i)), \dots, x_{iK} e_i E(f_K(\omega_i))] = 0; \\ E(e_i^2 \mathbf{z}_{i2}^* \mathbf{z}_{i2}' e_i) &= e_i^2 E(\mathbf{z}_{i2}^* \mathbf{z}_{i2}') \xrightarrow{p} e_i^2 Q_i, \text{ with } Q_i = E(\mathbf{z}_{i2} \mathbf{z}_{i2}'). \end{aligned}$$

Defining $V_3 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N e_i^2 Q_i$, the Lindeberg-Feller central limit theorem leads to

$$\mathbf{Z}_2^{*'} \mathbf{e} / \sqrt{N} = \sum_{i=1}^N \mathbf{z}_{i2}^* e_i / \sqrt{N} \xrightarrow{d} N(0, V_3).$$

Unconditionally, $\{e_i\}_{i=1}^N$ is drawn from a distribution with mean zero and variance σ_i^2 . Define an element in $e_i^2 Q_i$ as $e_i^2 Q_i(l, k)$ with $l, k \in [1, K]$. According to (A1) and (A2), $E(e_i^2 Q_i(l, k)) = \sigma_i^2 Q_i(l, k)$ and $E(e_i^4 Q_i(l, k)^2) = E(e_i^4) Q_i(l, k)^2 < \infty$. By the law of large numbers for independent and heterogeneous random variables

$$\frac{1}{N} \sum_{i=1}^N e_i^2 Q_i - \frac{1}{N} \sum_{i=1}^N \sigma_i^2 Q_i \xrightarrow{a.s.} 0, \text{ and, thus, } V_3 - V_2 \xrightarrow{a.s.} 0.$$

Proof for c) is completed.

The following propositions assert that the asymptotic properties of FB based OLS estimates $\hat{\beta}_2^*$ approximate their population counterparts if $\beta_2 = 0$. Moreover, the FB variant of T_N^* is stated to share the asymptotic features of T_N under H_0 .

Proposition 1 *Assume model (7.5) with conditions (A1)-(A5). Under $H_0 : \beta_2 = 0$, as $N \rightarrow \infty$,*

- a) $\mathbf{Z}^{*'} \mathbf{Z}^* / N \xrightarrow{p} \text{diag}(\Sigma_{11}, \Sigma_{22})$, with $\mathbf{Z}^* = (\mathbf{Z}_1, \mathbf{Z}_2^*)$;
- b) $\hat{\beta}_1^* \xrightarrow{p} \beta_1$;
- c) $\sqrt{N} \hat{\beta}_2^* \xrightarrow{d} N(0, \Sigma_{22}^{-1} V_2 \Sigma_{22}^{-1})$.

PROOF. a) Note that

$$\mathbf{Z}^{*'} \mathbf{Z}^* / N = \begin{pmatrix} \mathbf{Z}_1' \mathbf{Z}_1 / N & \mathbf{Z}_1' \mathbf{Z}_2^* / N \\ \mathbf{Z}_2^{*'} \mathbf{Z}_1 / N & \mathbf{Z}_2^{*'} \mathbf{Z}_2^* / N \end{pmatrix}.$$

Combining Lemma 2.a and 2.b, statement a) is proved.

b) Based on statement a) of Proposition 1, it can be obtained that

$$(\hat{\beta}_1^* - \beta_1) - \Sigma_{11}^{-1} \mathbf{Z}_1' e / N \xrightarrow{p} 0.$$

Conditional on the observed sample, e is fixed. Unconditionally, $\mathbf{Z}_1' e / N \xrightarrow{a.s.} 0$ according to the law of large numbers for independent and heterogeneous random variables. Thus, statement b) is shown.

c) Similarly, through statement a) of Proposition 1, it can be shown that

$$\sqrt{N}(\hat{\beta}_2^* - \beta_2^*) - \Sigma_{22}^{-1} \mathbf{Z}_2^{*'} \mathbf{e} / \sqrt{N} \xrightarrow{p} 0.$$

Since $\mathbf{Z}_2^{*'} \mathbf{e} / \sqrt{N} \xrightarrow{d} N(0, V_2)$ according to Lemma 2.c and $\beta_2^* = 0$ under H_0 , statement c) is proved by asymptotic equivalence as defined in Lemma 4.7 of White (2001).

Proposition 2 *Assume model (7.5) with conditions (A1)-(A6). Under $H_0 : \beta_2 = 0$, as $N \rightarrow \infty$,*

$$\sup_{0 < c < \infty} |P(T_N^* \leq c) - P(T_N \leq c)| \xrightarrow{p} 0.$$

PROOF. Denote $H_0 : \beta_2 = 0$ as $R\beta = 0$, with R defined in (7.9). The goodness-of-fit test statistic T_N in (7.7) can be equivalently formulated as

$$T_N = \frac{\hat{\beta}_2' (R(\mathbf{Z}'\mathbf{Z})^{-1}R')^{-1} \hat{\beta}_2}{\tilde{\sigma}^2} = Q'Q,$$

where $Q = (R(\mathbf{Z}'\mathbf{Z})^{-1}R')^{-1/2} \hat{\beta}_2 / \tilde{\sigma}$, $\tilde{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N \hat{e}_i^2$, and $\{\hat{e}_i\}_{i=1}^N$ are OLS residuals from (7.12). The corresponding bootstrap statistic is then,

$$T_N^* = \frac{\hat{\beta}_2^{*'} (R(\mathbf{Z}^{*'}\mathbf{Z}^*)^{-1}R')^{-1} \hat{\beta}_2^*}{\tilde{\sigma}^{*2}} = Q^{*'}Q^*,$$

where $Q^* = (R(\mathbf{Z}^{*'}\mathbf{Z}^*)^{-1}R')^{-1/2} \hat{\beta}_2^* / \tilde{\sigma}^*$, $\tilde{\sigma}^{*2} = \frac{1}{N} \sum_{i=1}^N \hat{e}_i^{*2}$, and OLS residuals \hat{e}_i^* stem from (7.13). According to previous arguments for Proposition 1a and 1c,

$$\begin{aligned} \sup_{0 < c < \infty} |P((R(\mathbf{Z}^{*'}\mathbf{Z}^*/N)^{-1}R')^{-1/2} \sqrt{N}R\hat{\beta}^* \leq c) \\ - P((R(\mathbf{Z}'\mathbf{Z}/N)^{-1}R')^{-1/2} \sqrt{N}R\hat{\beta} \leq c)| \xrightarrow{p} 0. \end{aligned}$$

Since $\text{plim}(\tilde{\sigma}^2) = \sigma^2$, $\sup_{0 < c < \infty} |P(Q^* \leq c) - P(Q \leq c)| \xrightarrow{p} 0$ if $\text{plim}(\tilde{\sigma}^{*2}) = \sigma^2$.

The latter is shown through the following Lemma. Thus, $\sup_{0 < c < \infty} |P(Q^{*'}Q^* \leq c) - P(Q'Q \leq c)| \xrightarrow{p} 0$ and Proposition 2 is proved.

Lemma 3 Assume model (7.5) with conditions (A1)-(A6). Under $H_0 : \beta_2 = 0$, as $N \rightarrow \infty$, $\tilde{\sigma}^{*2} \xrightarrow{p} \sigma^2$.

PROOF. Note $\hat{e}^{*'} \hat{e}^*/N = e'e/N - (e'Z^*/N) (Z^{*'}Z^*/N)^{-1} (Z^{*'}e/N)$. Based on Proposition 1 and Assumption (A6), it can be obtained that

$$\begin{aligned} & \text{plim}[(e'Z^*/N) (Z^{*'}Z^*/N)^{-1} (Z^{*'}e/N)] \\ &= \text{plim}(e'Z^*/N) \text{diag}(\Sigma_{11}^{-1}, \Sigma_{22}^{-1}) \text{plim}(Z^{*'}e/N), \\ &= \text{plim}(e'Z_1/N) \Sigma_{11}^{-1} \text{plim}(Z_1'e/N) - \text{plim}(e'Z_2^*/N) \Sigma_{22}^{-1} \text{plim}(Z_2^{*'}e/N), \\ &= \text{plim}(e'Z_1/N) \Sigma_{11}^{-1} \text{plim}(Z_1'e/N). \end{aligned}$$

Unconditionally, e_i is an independently distributed zero mean random variable. Therefore, $Z_1'e/N \xrightarrow{a.s.} 0$ and, thus, $\tilde{\sigma}^{*2} \xrightarrow{p} \sigma^2$.

It is worth mentioning that assumption (A1) on the independence of factor observations $\omega_1, \omega_2, \dots, \omega_N$ might be relaxed since the factor has no impact on the dependent variable under H_0 . Thus, the asymptotic features of the test statistic under H_0 should not be influenced when allowing some form of dependent, e.g. serially correlated, factor observations. For convenience of proving the asymptotic properties of the test statistic under H_0 , however, this study relies on (A1) stating independence of $\{\omega_i\}_{i=1}^N$.

Throughout, this section has been focussed on contrasting two parametric hypothesis by means of the T_N statistic. The implementation of resampling schemes in case of the nonparametric functional form $\underline{\beta}_k(\omega_i)$ in (7.1) is straightforward. Then, as advocated by CF&Y, \hat{e}_i is obtained as $\hat{e}_i = y_i - z'_{i1} \hat{\underline{\beta}}(\omega_i)$ in step 1) of RB. Correspondingly, WB variants will draw $e_i^* = \hat{e}_i v_i$. PB and FB samples are drawn from $\{y_i - z'_{i1} \hat{\underline{\beta}}(\omega_i) + z'_{i1} \hat{\underline{\beta}}_1^{(0)}, z_{i1}, \omega_i\}_{i=1}^N$ and $\{y_i, z_{i1}, \omega_i^*\}_{i=1}^N$, respectively. Finally, RSS_1^* is calculated analogously to (7.11) instead of (7.8) for all four bootstrap approaches.

7.4 Monte-Carlo analysis

This section provides a comparison of the four bootstrap methods by means of Monte Carlo (MC) experiments. The empirical features of the alternative designs for resampling T_N are investigated by testing the parametric null hypothesis against a parametric or a semiparametric alternative. Throughout, MC exercises cover 2000 replications with the number of bootstrap draws chosen as $R = 299$.

7.4.1 Simulation design

Specifying regression (7.1) the following data generating model is considered

$$y_i = \underline{\beta}(\omega_i)x_i + e_i, \quad \underline{\beta}(\omega_i) = 0.13 + \sin(\omega_i), \quad i = 1, \dots, N, \quad (7.14)$$

where ω_i is sampled from the Gaussian distribution. The regressor variables $\{x_i\}_{i=1}^N$ are drawn once from the Gaussian distribution and then fixed over all MC replications. To evaluate size and power features of competing inferential approaches $\{y_i\}_{i=1}^N$ are generated according to a sequence of alternative models indexed by η ,

$$y_i = \left(\frac{1}{N} \sum_{i=1}^N \underline{\beta}(\omega_i) + \eta \left[\underline{\beta}(\omega_i) - \frac{1}{N} \sum_{i=1}^N \underline{\beta}(\omega_i) \right] \right) x_i + e_i,$$

where $\eta = 0.05q$ with $q = 0, 2, 4$. The constant coefficient model is obtained for $q = 0$ whereas the strongest parameter variation is realized in case $q = 4$. The error terms $\{e_i\}_{i=1}^N$ in (7.14) are simulated as a normally distributed variable with mean zero and variance σ_i^2 , where both homogeneous ($\sigma_i^2 = 1$) and heterogeneous ($\sigma_i^2 = 0.8|x_i|$) variances are considered. MC sample sizes cover the cases $N = 50, 100, 200$ and $N = 400$.

For parametric modeling, the regression

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i \sin(\omega_i) + e_i,$$

is employed such that the constant coefficient model can be formulated as $H_0 : \beta_2 = 0$. For semiparametric modeling, the kernel estimator (7.10) is applied to the regression

$$y_i = \underline{\beta}_1(\omega_i) + \underline{\beta}_2(\omega_i)x_i + e_i.$$

The considered null hypothesis is that both $\underline{\beta}_1(\omega_i)$ and $\underline{\beta}_2(\omega_i)$ are constant. Regarding the kernel function, the Gaussian kernel, $K(\cdot/h) = (2\pi)^{-1/2} \exp(-0.5(\cdot/h)^2)$, is used where $h = 1.06\hat{s}_u N^{-1/5}$ and \hat{s}_u is the standard error estimate available from the factor observations. For a detailed discussion of the nonparametric toolkit the reader is referred to Härdle, Müller, Sperlich and Werwatz (2004).

7.4.2 Size features

Table 7.1 shows the empirical size of the goodness-of-fit-tests at the 5% nominal level. Results from parametric and semiparametric modeling are shown in Panel A and Panel B, respectively. For parametric modeling, not only bootstrap (RB, WB, PB and FB) schemes for T_N and T_N^{hc} are considered, but also their corresponding F - and χ^2 -approximations.

First consider the results for parametric modeling (Panel A). The RB variant of T_N and the conventional F -test, $((N - K)/N)T_N \sim F(1, N - K)$, over-reject the null hypothesis when error terms are heteroskedastic, since both tests build upon an iid assumption. Under homoskedastic model disturbances, the empirical size is between 0.045 and 0.057 (RB) and 0.046 and 0.051 (F -test). In contrast, under heteroskedasticity of model disturbances the rejection probabilities of the WB, PB and FB tests increase only slightly. The empirical size obtained via FB is almost uniformly closer to the nominal level in comparison with WB and PB. Under homoskedastic model innovations WB and PB inference yield empirical size estimates close to the

Table 7.1: Size estimates

| N | $\sigma_i^2 = 1$ | | | | | $\sigma_i^2 = 0.8 x_i $ | | | | |
|---------------------------------------|------------------|-------------|-------------|------|-------------|-------------------------|-------------|-------------|-------------|-------------|
| | RB | WB | PB | FB | F/χ^2 | RB | WB | PB | FB | F/χ^2 |
| A - Parametric alternative: T_N | | | | | | | | | | |
| 50 | .049 | .078 | .065 | .051 | .048 | .144 | .092 | .076 | .058 | .142 |
| 100 | .045 | .064 | .059 | .047 | .046 | .161 | .077 | .072 | .054 | .159 |
| 200 | .057 | .062 | .060 | .058 | .055 | .151 | .067 | .064 | .060 | .152 |
| 400 | .046 | .052 | .051 | .045 | .051 | .147 | .054 | .056 | .050 | .144 |
| T_N^{hc} | | | | | | | | | | |
| 50 | .050 | .061 | .057 | .049 | .088 | .070 | .069 | .060 | .054 | .109 |
| 100 | .051 | .052 | .054 | .049 | .068 | .062 | .055 | .056 | .050 | .091 |
| 200 | .058 | .059 | .056 | .057 | .063 | .067 | .057 | .057 | .057 | .075 |
| 400 | .047 | .050 | .050 | .047 | .053 | .055 | .049 | .054 | .049 | .059 |
| B - Semiparametric alternative: T_N | | | | | | | | | | |
| 50 | .043 | .182 | .035 | .057 | - | .160 | .197 | .061 | .061 | - |
| 100 | .051 | .175 | .030 | .057 | - | .183 | .189 | .060 | .056 | - |
| 200 | .045 | .167 | .034 | .052 | - | .198 | .158 | .042 | .055 | - |
| 400 | .044 | .156 | .031 | .052 | - | .254 | .177 | .047 | .070 | - |

This table documents empirical sizes for the T_N and T_N^{hc} statistic employed under the null hypothesis of a constant coefficient model against a parametric (panel A) and semiparametric (panel B) alternative. MC exercises include RB, WB, PB and FB resampling. In the parametric case, inferential results for the common F - and χ^2 -statistic are also provided. The nominal significance level is 5%. Bold entries indicate violations of the nominal level diagnosed with 5% significance.

nominal size for $N = 400$ but tend to over-reject H_0 in smaller samples ($N = 50, 100, 200$).

With respect to the heteroskedasticity consistent test statistic T_N^{hc} , the first order asymptotic approximation yields empirical size estimates in excess of the nominal level for small to medium sample sizes. Similar as documented for T_N , WB inference is still oversized in small samples ($T = 50$). RB inference under heteroskedastic disturbances remains oversized even for larger sample sizes ($T = 100, 200$). Both FB and PB resampling deliver an empirical size close to 0.05.

Now turn to the results for semiparametric modeling (Panel B). Under heteroskedastic error terms only PB and FB based inference show valid em-

pirical significance levels. While size estimates from PB and FB are mostly close to 0.05, corresponding quantities from RB and WB are between 0.160 and 0.254 and 0.158 and 0.197, respectively. The slight increase of FB based rejection frequencies under H_0 to 0.07 is an artefact of the set of simulated samples. Under homoskedastic disturbances PB based inference constantly under-rejects the null hypothesis with empirical estimates between 0.03 and 0.035. At the same time WB based inference strongly over-rejects under H_0 with empirical size estimates between 0.156 and 0.182. In contrast, RB and FB resampling offer accurate empirical size features.

7.4.3 Power features

Size adjusted power estimates in Table 7.2 are documented for two alternative scenarios, $\eta = 0.1, 0.2$. The nominal significance level of alternative test procedures is adjusted such that the empirical size is 5%. Since size adjustment is not sensible in case of marked size distortions, power estimates are not reported for RB based inference of the T_N statistic under heteroskedasticity and WB resampling in the semiparametric modeling framework. It can be seen that power estimates under parametric modeling uniformly exceed the corresponding estimates obtained from semiparametric modeling. Overall, adjusted power estimates hint at consistency of all competing test procedures. Furthermore, for parametric modeling FB yields similar power properties compared with corresponding first order asymptotic approximations. For semiparametric modeling RB and PB perform slightly better than FB under homoskedastic errors. Under heteroskedastic disturbances PB inference is slightly more powerful than FB inference for $N = 50$.

Summarizing the scope of FB inference in applied functional modelling we note that among all competing bootstrap schemes FB has most accurate size features. FB inference is robust against heteroskedasticity and does not

go along with power loss under an actually functional model (H_1).

Table 7.2: Size adjusted power estimates

| N | $\sigma_i^2 = 1$ | | | | | $\sigma_i^2 = 0.8 x_i $ | | | | |
|---|------------------|------|------|------|------------|-------------------------|------|------|------|------------|
| | RB | WB | PB | FB | F/χ^2 | RB | WB | PB | FB | F/χ^2 |
| A - Parametric alternative: $T_N, \eta = 0.1$ | | | | | | | | | | |
| 50 | .066 | .061 | .060 | .064 | .065 | - | .065 | .066 | .058 | - |
| 100 | .097 | .089 | .084 | .086 | .091 | - | .078 | .081 | .084 | - |
| 200 | .125 | .130 | .129 | .123 | .132 | - | .101 | .111 | .110 | - |
| 400 | .233 | .228 | .231 | .244 | .232 | - | .203 | .204 | .206 | - |
| $T_N^{hc}, \eta = 0.1$ | | | | | | | | | | |
| 50 | .065 | .060 | .060 | .061 | .057 | .063 | .061 | .057 | .065 | .063 |
| 100 | .087 | .086 | .082 | .086 | .087 | .086 | .080 | .079 | .086 | .083 |
| 200 | .127 | .125 | .125 | .123 | .126 | .105 | .104 | .111 | .109 | .109 |
| 400 | .242 | .229 | .231 | .234 | .237 | .195 | .202 | .194 | .204 | .204 |
| B - Semiparametric alternative: $T_N, \eta = 0.1$ | | | | | | | | | | |
| 50 | .063 | - | .058 | .057 | - | - | - | .056 | .053 | - |
| 100 | .056 | - | .067 | .058 | - | - | - | .059 | .064 | - |
| 200 | .077 | - | .078 | .069 | - | - | - | .082 | .071 | - |
| 400 | .097 | - | .099 | .088 | - | - | - | .103 | .089 | - |
| A - Parametric alternative: $T_N, \eta = 0.2$ | | | | | | | | | | |
| 50 | .131 | .116 | .119 | .118 | .134 | - | .109 | .113 | .115 | - |
| 100 | .206 | .189 | .190 | .190 | .199 | - | .172 | .172 | .180 | - |
| 200 | .368 | .371 | .373 | .366 | .388 | - | .322 | .325 | .329 | - |
| 400 | .691 | .679 | .678 | .698 | .689 | - | .630 | .619 | .627 | - |
| $T_N^{hc}, \eta = 0.2$ | | | | | | | | | | |
| 50 | .115 | .108 | .110 | .123 | .117 | .113 | .101 | .098 | .110 | .108 |
| 100 | .184 | .179 | .175 | .182 | .181 | .178 | .176 | .168 | .180 | .173 |
| 200 | .365 | .373 | .364 | .360 | .367 | .325 | .326 | .318 | .332 | .333 |
| 400 | .689 | .677 | .672 | .686 | .688 | .616 | .625 | .599 | .623 | .633 |
| B - Semiparametric alternative: $T_N, \eta = 0.2$ | | | | | | | | | | |
| 50 | .085 | - | .077 | .079 | - | - | - | .088 | .077 | - |
| 100 | .096 | - | .111 | .085 | - | - | - | .102 | .115 | - |
| 200 | .169 | - | .174 | .155 | - | - | - | .194 | .172 | - |
| 400 | .297 | - | .313 | .284 | - | - | - | .333 | .289 | - |

This table documents size adjusted power estimates given $\eta = 0.1, 0.2$ for the T_N and T_N^{hc} statistic via parametric (A) and semiparametric modeling (B). Size adjustment is addressed by tuning the nominal level of a test such that the empirical size is 5%. '-' indicates that the provision of adjusted power is not sensible owing to massive size distortions. For further notes see Table 7.1.

7.5 An empirical example

This section provides a simple application of the functional coefficient model to between regressions of domestic investment on domestic saving. FH interpret the high SI relation in the between regression for OECD countries as a signal for low capital mobility. To find other factors than capital mobility governing the SI relation, they employed the following augmented between regressions,

$$\bar{I}_i = \beta_0 + (\beta_1 + \beta_2 \bar{w}_i) \bar{S}_i + e_i, \quad i = 1, \dots, N. \quad (7.15)$$

Variables in (7.15) are defined as those in (2.1). The time average $\bar{w}_i = 1/T \sum_t \tilde{w}_{it}$ is a measure of some factor characterizing the i -th member of the cross section. As particular factor variables \tilde{w}_{it} entering (7.15), FH use the rate of population growth, the degree of openness, measured as the sum of exports and imports in relation to GDP, and log GDP. To diagnose factor dependence, FH refer to the significance of coefficient estimates $\hat{\beta}_2$. However, they could not find that any of the considered factors significantly influences the link between domestic saving and investment.

A potential drawback of the specification (7.15) may lie in the presumed linear factor dependence. For this reason, the following functional coefficient model to the SI relation is implemented,

$$\tilde{\bar{I}}_i = \underline{\beta}(\tilde{\bar{w}}_i) \tilde{\bar{S}}_i + e_i. \quad (7.16)$$

The measure for domestic investment and saving is mean adjusted in (7.16) so that only the slope coefficient is factor dependent, i.e. $\tilde{\bar{I}}_i = \bar{I}_i - 1/N \sum_i \bar{I}_i$ and $\tilde{\bar{S}}_i = \bar{S}_i - 1/N \sum_i \bar{S}_i$. Thus, regression (7.16) is a semiparametric counterpart to regression (7.15). To provide comparable results over alternative measures \tilde{w}_{it} the factor variable in (7.16) is standardized, i.e. $\tilde{\bar{w}}_i = (\bar{w}_i - 1/N \sum_i \bar{w}_i) / \sigma(\bar{w}_i)$, with $\sigma(\bar{w}_i)$ being the cross sectional standard deviation of \bar{w}_i . Moreover, it is tempting to allow both the constant and slope

coefficient in the between regression to be factor dependent. Therefore, the functional coefficient model is also considered,

$$\bar{I}_i = \underline{\beta}_1(\tilde{w}_i) + \underline{\beta}_2(\tilde{w}_i)\bar{S}_i + e_i, \quad (7.17)$$

where \bar{I}_i , \bar{S}_i and \tilde{w}_i are defined as before. Models (7.16) and (7.17) are estimated via kernel smoothing to obtain the corresponding T_N statistics. The null hypotheses for (7.16) and (7.17) are $H_0 : \underline{\beta}(\tilde{w}_i) = \beta$ and $H_0 : \underline{\beta}_1(\tilde{w}_i) = \beta_1$, $\underline{\beta}_2(\tilde{w}_i) = \beta_2$, respectively, with β , β_1 and β_2 being constant parameters.

To investigate the factor dependent SI relation, W97, L68, and O26 with annual data from 1971 to 2002 are considered. Applying a t -test for $\beta_2 = 0$ in regression (7.15), only in L68 the population growth rate and the openness ratio have significant impacts on the SI relation at the 10% level, with estimated coefficients about -0.055 and 0.053 respectively.² Thus, similar to the results in FH, applying common t -statistics the evidence in favor of factor dependence is rather weak.

Bootstrap p -values obtained from $R = 499$ RB, WB, PB and FB samples are documented in Table 7.3. Results from models (7.16) and (7.17) are shown in Panel A and B, respectively. First of all, RB and WB based p -values are almost always much smaller than the FB counterparts. These results are consistent with MC analyses under heteroskedastic error terms. Since WB turned out to be vastly oversized in semiparametric modeling, only the results from RB, PB and FB are discussed henceforth.

As can be seen in Panel A of Table 7.3, the SI relation is found to be significantly affected by the population growth and the openness measure. Both factors have p -values in L68 and W97 smaller than 10%. Employing the semiparametric model (7.17) (Panel B), both factor variables impact on the constant and slope coefficient jointly with 10% significance. Although log GDP is found to have an insignificant impact on the SI relation (Panel

²The corresponding t -statistics are -2.59 and 2.14 .

Table 7.3: Tests for the factor independent SI relation in between regressions

| Factor | RB | | | WB | | | PB | | | FB | | |
|---------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| | O26 | L68 | W97 | O26 | L68 | W97 | O26 | L68 | W97 | O26 | L68 | W97 |
| A: Model without constant | | | | | | | | | | | | |
| LGDP | .194 | .345 | .639 | .072 | .693 | .774 | .188 | .667 | .832 | .353 | .559 | .820 |
| POPG | .150 | .002 | .010 | .070 | .000 | .002 | .257 | .120 | .096 | .333 | .036 | .040 |
| OPN | .731 | .060 | .000 | .824 | .158 | .020 | .874 | .329 | .074 | .804 | .220 | .018 |
| B: Model with constant | | | | | | | | | | | | |
| LGDP | .343 | .034 | .078 | .092 | .168 | .112 | .255 | .200 | .339 | .355 | .132 | .168 |
| POPG | .505 | .016 | .014 | .323 | .024 | .034 | .509 | .202 | .146 | .577 | .078 | .034 |
| OPN | .752 | .006 | .000 | .475 | .006 | .000 | .577 | .042 | .030 | .810 | .010 | .006 |

This table reports RB, WB, PB and FB implied p -values for goodness-of-fit test statistics (T_N). Three factors are considered: the logarithm of GDP (LGDP), the rate of population growth (POPG) and the openness ratio (OPN).

A), it affects both the constant and slope coefficient in model (7.17) jointly with 10% significance according to RB inference (Panel B). This may reflect partially the influence of GDP growth on the investment rate directly. In addition, it can be seen that conditional on OECD economies all p -values via RB, PB and FB exceed 10% for both models (7.16) and (7.17). On the one hand, this might be due to the reduced power of the test given only 26 observations. On the other hand, this result might reflect the smaller variation of factor variables in OECD countries, i.e. their relative homogeneity according to the population growth and the openness of goods markets.

To illustrate the variation of the SI relation, the functional estimates of SI relation over the support of $-2 \leq \tilde{w}_i \leq 2$ are displayed for the considered factors. Since the variations of $\hat{\beta}_2(\tilde{w}_i)$ from model (7.17) are very similar to those of $\hat{\beta}(\tilde{w}_i)$ from (7.16), we show only the latter in Figure 7.1. Furthermore, since the SI relation is found unaffected by any considered factors in O26 and also not affected by log GDP for all samples according RB, PB and FB based tests (Table 7.3, Panel A), plots in the first row and first column of Figure 7.1 are not discussed further. Focusing on the local estimates of the SI

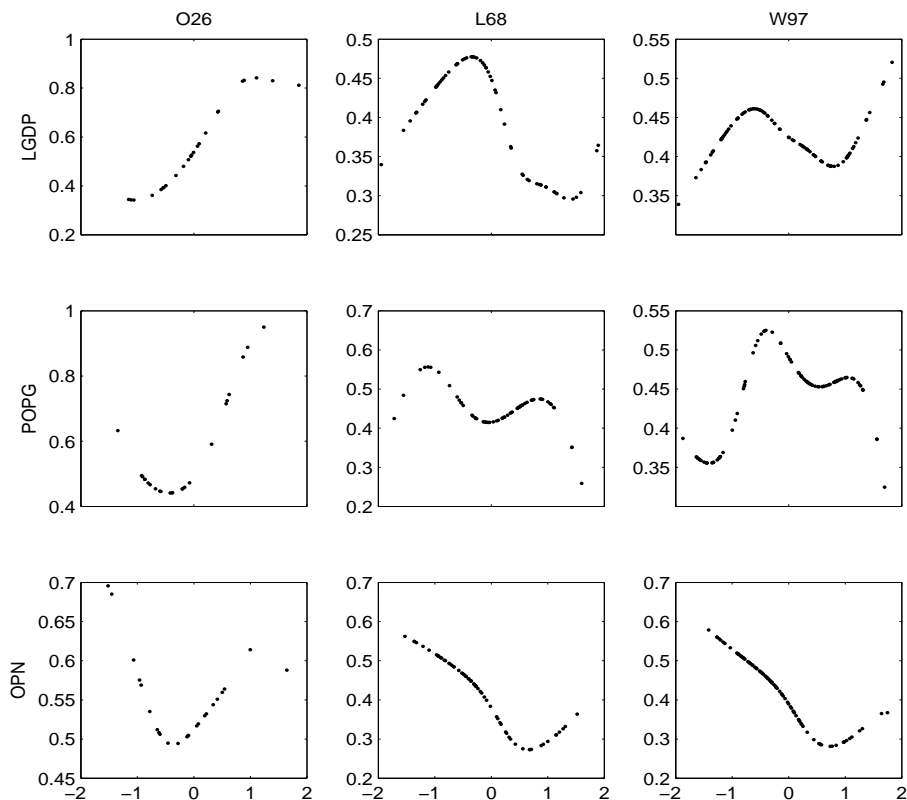


Figure 7.1: Functional estimates of the SI relation conditional on the logarithm of GDP (LGDP), the rate of population growth (POPG) and the openness ratio (OPN).

relation conditional on the population growth in L68 and W97 (the last two plots in the second row of Figure 7.1), a clear trending pattern of the estimated SI relation is not observed. However, conditioning the SI relation on the openness ratio obtains a decreasing functional pattern for L68 and W97 (the last two plots in the bottom row of Figure 7.1). This supports the view that economies with more integrated goods markets tend to have lower SI relations.

7.6 Summary

In this chapter, a new factor based bootstrap approach for functional coefficient models is introduced to test parameter invariance against a functional alternative. Modeling the functional coefficient parametrically, the bootstrap approximation of the distribution of the goodness-of-fit test statistic is shown to be valid asymptotically under the null hypothesis. The new bootstrap scheme is shown to be robust against heteroskedastic error distributions. Furthermore, for small samples factor based inference outperforms the wild bootstrap and the pairs bootstrap counterpart that are also known to cope with heteroskedasticity. When the functional coefficient is modeled nonparametrically, only tests via FB provide correct empirical size estimates. Applying functional coefficient model to between regressions of domestic investment on domestic saving, variations of the saving retention coefficient given the value of the population growth rate and the openness ratio is confirmed and illustrated. In the next chapter, the bivariate state dependent FCM is developed to investigate the determinants of the SI relation in detail.

Chapter 8

Application of FCM to SI relations

The preceding analyses have shown that the link between domestic saving and investment exhibits some downward trending behavior and country specific features. Moreover, profile regressions reveal that the correlation between domestic saving and investment may be explained conditional on some economic factor variables. Given the likelihood of parameter variation over two data dimensions, all empirical approaches followed so far carry the risk of providing biased results since at most one dimension of potential parameter dependence has been taken into account. From these observations, one may refrain from modeling the SI relation by means of econometric specifications presuming some restrictive form of (cross-sectional or time) homogeneity or state invariance. As a consequence one may alternatively opt for local models where the parameters of interest are given conditionally on some economic state variable measured over both dimensions of the panel. For these reasons semiparametric models that can be seen as a bivariate generalization of functional coefficient models presented in the last chapter are adopted. A further merit of this approach and its local implementation is that it might give valuable information on the accuracy of the restrictive nature of parametric models.

This chapter¹ is structured as follows: The next two sections introduce, for convenience, a one dimensional factor models fitting into the framework introduced by Cai, Fan and Yao (2000). While Section 8.1 provides the time dependent regression for the SI relation, Section 8.2 introduces the cross-section dependent regression. In Section 8.3 these two perspectives are combined to formulate the bivariate state dependent FCM. State dependent and invariant model representations are contrasted by means of CV criteria in Section 8.4. The implementation and inferential issues are discussed in Section 8.5 and 8.6 accordingly. The results for functional estimation is provided in Section 8.7. Section 8.8 summarizes the main results.

8.1 Time specific regression

To identify cross-sectional evidence for the factor dependency of the SI relation, the following time specific regression can be considered,

$$\begin{aligned} I_{it} &= \underline{\beta}_{1t}(w_{it}^{(t)}) + \underline{\beta}_{2t}(w_{it}^{(t)})S_{it} + e_{it}, \quad i = 1, \dots, N, \\ \equiv y_{it} &= \mathbf{z}'_{it}\underline{\boldsymbol{\beta}}_t(w_{it}^{(t)}) + e_{it}, \end{aligned} \quad (8.1)$$

where $\underline{\boldsymbol{\beta}}_t(\cdot) = (\underline{\beta}_{1t}(\cdot), \underline{\beta}_{2t}(\cdot))'$ and $\mathbf{z}'_{it} = (1, S_{it})$. The time specific factor measure $w_{it}^{(t)}$ is defined in a standardized fashion to enable a comparison of the functional estimates from different factors or time periods. Following this line, $w_{it}^{(t)} = (\tilde{w}_{it} - \bar{w}_t)/\sigma_t(\tilde{w})$, where \bar{w}_t and $\sigma_t(\tilde{w})$ denote the empirical (time dependent) cross-sectional mean, $\bar{w}_t = 1/N \sum_{i=1}^N \tilde{w}_{it}$, and time specific standard error of \tilde{w}_{it} , respectively. Note that \bar{w}_t , $t = 1, \dots, T$, might be interpreted as a factor's long run time path measured over the cross-section. In case the sample is as large as W97, \bar{w}_t approximates a factor's global evo-

¹For a different version of this chapter see "A functional coefficient model view of the Feldstein-Horioka puzzle", with Helmut Herwartz, University Kiel, Economics Working Paper 2007-14.

lution over time. For instance, with regard to the openness variable, \bar{w}_t is suitable to reflect the world wide trend towards global specialization and an intensified international exchange of goods.

For estimating the factor dependent parameter vector $\underline{\beta}_t(w_{it}^{(t)})$, a semi-parametric estimator similar to the Nadaraya Watson estimator (Nadaraya 1964, Watson 1964) can be obtained as

$$\begin{aligned}\hat{\underline{\beta}}_t(w^{(t)}) &= \mathcal{Z}_t^{-1}(w^{(t)})\mathcal{Y}_t(w^{(t)}), \\ \text{where } \mathcal{Z}_t(w^{(t)}) &= \sum_{i=1}^N z_{it}z'_{it}K_{t,h}(w_{it}^{(t)} - w^{(t)}), \\ \mathcal{Y}_t(w^{(t)}) &= \sum_{i=1}^N z_{it}y_{it}K_{t,h}(w_{it}^{(t)} - w^{(t)}).\end{aligned}$$

In this ‘weighted’ least square estimation $w^{(t)}$ denotes a local point in $\{w_{it}^{(t)}\}_{i=1}^N$ and $K_{\bullet,h}(u) = K_{\bullet}(u/h)/h$, where $K(\cdot)$ is a kernel function and h is the bandwidth parameter.

8.2 Cross-section specific regression

Analogously to the FCM for time specific regression in the last section, the following cross-sectional specific regression can be applied to obtain time-series evidence for the factor dependent SI relation, i.e.

$$\begin{aligned}I_{it} &= \underline{\beta}_{1i}(w_{it}^{(i)}) + \underline{\beta}_{2i}(w_{it}^{(i)})t + \underline{\beta}_{3i}(w_{it}^{(i)})S_{it} + e_{it}, \quad t = 1, \dots, T, \quad (8.2) \\ \equiv y_{it} &= \mathbf{z}'_{it}\underline{\beta}_i(w_{it}^{(i)}) + e_{it},\end{aligned}$$

where $\underline{\beta}_i(\cdot) = (\underline{\beta}_{1i}(\cdot), \underline{\beta}_{2i}(\cdot), \underline{\beta}_{3i}(\cdot))'$ and $\mathbf{z}'_{it} = (1, t, S_{it})$. This model formalizes the view that the SI relation responds to (changes of) some underlying factor, $w_{it}^{(i)}$, characterizing the state of economy i . The inclusion of a (local) trend parameter within the functional model allows to distinguish deterministic and time varying measurable economic conditions which are supposed to impact on the SI relation.

The inclusion of a deterministic trend term in (8.2) is thought to disentangle deterministic features of the SI relation from factor dependence. To measure economic states it is natural to represent the factor in some standardized form so that cross-sectional comparisons are facilitated. Owing to potential nonstationarity of the time path of a particular factor variable measured for a specific cross-section member, standardized factors are considered, i.e.

$$w_{it}^{(i)} = (\tilde{w}_{it} - \hat{w}_{it}^{(hp)}) / \sigma_i(\text{gap}(w)). \quad (8.3)$$

In (8.3) $\hat{w}_{it}^{(hp)}$ is the long run time path of a particular factor variable as obtained from applying the Hodrick-Prescott (HP) filter (Hodrick and Prescott 1997) to \tilde{w}_{it} , $t = 1, \dots, T$. Accordingly, the process $\tilde{w}_{it} - \hat{w}_{it}^{(hp)}$ describes the ‘factor gap’ for economy i having unconditional (cross-section specific) variance $\sigma_i^2(\text{gap}(w))$. To implement (8.3) with yearly factor observations the HP smoothing parameter is set to 6.25 as recommended by Ravn and Uhlig (2002). Note that the standardized ‘factor gap’ as defined in (8.3) has a unit unconditional variance. As an alternative for $\hat{w}_{it}^{(hp)}$ to measure a factor’s long run time path, one may a-priori also consider a cross-sectional mean, i.e. $\bar{w}_i = 1/T \sum_{t=1} \tilde{w}_{it}$. In case a particular factor variable is nonstationary, however, it is not clear what \bar{w}_i actually measures and, as such, it will not be representative for the factor over the entire sample period. In the opposite case of a stationary factor variable, \bar{w}_i is an efficient approximation of the factor’s ‘steady state’ but the efficiency loss implied by applying the HP filter might be moderate. Since controlling the time series features of diverse factor variables over a cross-section as large as W97 is not at the core of our analysis, the HP filter is preferred as an approximation of a factor’s long run time path. Along this line one may evaluate local SI relations conditional on scenarios where a particular factor variable for the i -th cross-section member is above, close to or below its long run time path.

Regarding, for instance, the ratio of exports plus imports over GDP as a factor, states of lower vs. higher ‘openness’ observed for a given economy over time could be distinguished to evaluate the SI relation locally.

The semiparametric estimator of the factor dependent parameter vector $\underline{\beta}_i(w_{it}^{(i)})$ then can be obtained as

$$\begin{aligned}\hat{\underline{\beta}}_i(w^{(i)}) &= \mathcal{Z}_i^{-1}(w^{(i)})\mathcal{Y}_i(w^{(i)}), \\ \text{where } \mathcal{Z}_i(w^{(i)}) &= \sum_{t=1}^T z_{it}z'_{it}K_{i,h}(w_{it}^{(i)} - w^{(i)}), \\ \mathcal{Y}_i(w^{(i)}) &= \sum_{t=1}^T z_{it}y_{it}K_{i,h}(w_{it}^{(i)} - w^{(i)}),\end{aligned}$$

where $w^{(i)}$ denotes a local point in $\{w_{it}^{(i)}\}_{t=1}^T$.

8.3 Bivariate state dependent model

Combining both cross-sectional and time dimensional evidence, both dimensions of a particular factor variable could be used to formalize a local view at the pooled regression model as

$$\begin{aligned}I_{it} &= \underline{\beta}_1(w^{(i)}, w^{(t)}) \\ &+ \underline{\beta}_2(w^{(i)}, w^{(t)})t + \underline{\beta}_3(w^{(i)}, w^{(t)})S_{it} + e_{it} \\ \equiv y_{it} &= z'_{it}\underline{\beta}(w^{(i)}, w^{(t)}) + e_{it} \\ &= z'_{it}\underline{\beta}(\omega) + e_{it},\end{aligned}\tag{8.4}$$

with $\omega = (w^{(i)}, w^{(t)})$ being a pair of local points in $\{\{w_{it}^{(i)}\}_{i=1}^N\}_{t=1}^T$ and $\{\{w_{it}^{(t)}\}_{i=1}^N\}_{t=1}^T$ accordingly.

To estimate the factor dependent parameter vector $\underline{\beta}(\omega)$ in (8.4) we proceed similar to a trivariate version of the Nadaraya Watson estimator. This

builds upon the following weighted sums of cross products of observations:

$$\mathcal{Z}(w^{(i)}, w^{(t)}) = \sum_{i=1}^N \sum_{t=1}^T \mathbf{z}_{it} \mathbf{z}'_{it} K_{i,h}(w_{it}^{(i)} - w^{(i)}) K_{t,h}(w_{it}^{(t)} - w^{(t)}), \quad (8.5)$$

$$\mathcal{Y}(w^{(i)}, w^{(t)}) = \sum_{i=1}^N \sum_{t=1}^T \mathbf{z}_{it} y_{it} K_{i,h}(w_{it}^{(i)} - w^{(i)}) K_{t,h}(w_{it}^{(t)} - w^{(t)}), \quad (8.6)$$

where the components of the bivariate factor variable $\omega_{it} = (w_{it}^{(i)}, w_{it}^{(t)})$ have been defined previously as

$$w_{it}^{(i)} = (\tilde{w}_{it} - \hat{w}_{it}^{(hp)}) / \sigma_i(\text{gap}(w)), \quad w_{it}^{(t)} = (\tilde{w}_{it} - \bar{w}_t) / \sigma_t(\tilde{w}).$$

In (8.5) and (8.6) $K_{\bullet,h}(u) = K_{\bullet}(u/h)/h$, where $K(\cdot)$ is a kernel function and h is the bandwidth parameter. From the moments given in (8.5) and (8.6), the semiparametric estimator is obtained as

$$\hat{\underline{\beta}}(\omega) = \hat{\underline{\beta}}(w^{(i)}, w^{(t)}) = \underline{\mathcal{Z}}^{-1}(\omega) \underline{\mathcal{Y}}(\omega). \quad (8.7)$$

As it is typical for kernel based estimation, the choice of the bandwidth parameter is of crucial importance for the factor dependent estimates given in (8.7) (Härdle, Hall and Marron 1988). For bandwidth selection, Scott's rule of thumb (Scott 1992) is used. Since the unconditional standard deviation of the factor variables over both data dimensions is (close to) unity by construction, the rule of thumb bandwidth is $h = (NT)^{-1/6}$. With regard to the kernel function, the Gaussian kernel is used, i.e. $K(u/h) = (2\pi)^{-1/2} \exp(-0.5(u/h)^2)$. Generally, NT is the number of observations available for the factor variable. For the practical implementation of the bivariate kernel estimator in the present case, it has to be pointed out that owing to missing observations the actual panel used for estimation is unbalanced for numerous factor variables. For convenience, the latter feature of the panel is suppressed by the employed notation.

8.4 Factor dependent model comparison

State dependent and invariant model representations could be contrasted by means of CV criteria (Allen 1974, Stone 1974, Geisser 1975), see Section 5.2. Since semiparametric estimates could become quite wiggly in the boundaries of the factor support, CV measures are only provided for those observations which correspond to ‘regular’ factor realizations such that

$$cv_{(mod)} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T |y_{it} - \hat{y}_{it}(mod)| \mathbb{I}(-2 \leq w_{it}^{(t)} \leq 2) \mathbb{I}(-2 \leq w_{it}^{(i)} \leq 2), \quad (8.8)$$

where $\mathbb{I}(\cdot)$ is an indicator function and *mod* refers either to state dependent or independent models.

For the three groups of factors considered in Section 3.2, CV criteria comparing the merits of local estimates (8.7) against pooled regressions are reported in Table 8.1. Since the results from model comparison based on squared and absolute forecast errors are very similar, only CV criteria for the latter are provided. CV estimates for the semiparametric model are given in Table 8.1 as a fraction of the pooled regression (state invariant version of (8.4)) CV statistics.

As can be seen from Table 8.1, the relative performance of the functional coefficient model (8.4) against the pooled regression differs over the alternative cross-sections as well as over the selected factor variables. For instance, used as a measure for capital market segmentation, an economies’ real interest rate differential measured against some world index (INTD, for details see Section 3.2) does not help to improve the pooled model since the relative CV measures are close to unity throughout. With only a very few exceptions, all relative CV estimates are less than unity and thereby indicate some gain in jackknife forecasting offered by the local model. To assess the significance of the relative measure, one should take into account that, depending on the cross-section, CV criteria are determined on the basis of

Table 8.1: Factor dependent model comparison

| Factor | W97 | L68 | D29 | O26 | O15 | E11 | F16 | E14 |
|---------|------|------|------|------|------|------|------|------|
| Group 1 | | | | | | | | |
| AGE | 0.97 | 0.98 | 0.93 | 0.92 | 0.93 | 0.94 | 0.94 | 0.96 |
| POPG | 0.99 | 0.99 | 0.98 | 0.98 | 0.91 | 0.93 | 0.95 | 0.91 |
| GDPC | 0.95 | 0.94 | 0.85 | 0.79 | 0.89 | 0.74 | 0.86 | 0.75 |
| GVBB | 0.99 | 0.98 | 0.99 | 0.96 | 0.97 | 1.02 | 0.92 | 0.96 |
| GVTT | 0.96 | 0.95 | 0.94 | 0.91 | 0.91 | 0.86 | 0.87 | 0.87 |
| GVIVM | 0.96 | 0.96 | 0.99 | 0.97 | 0.97 | 0.92 | 0.95 | 0.92 |
| GVCE | 0.97 | 0.98 | 0.92 | 0.90 | 0.92 | 0.85 | 0.87 | 0.86 |
| GVCON | 0.96 | 0.97 | 0.90 | 0.87 | 0.89 | 0.91 | 0.82 | 0.84 |
| Group 2 | | | | | | | | |
| OPN | 0.93 | 0.93 | 0.93 | 0.94 | 0.86 | 0.91 | 0.83 | 0.98 |
| EXPT | 0.94 | 0.94 | 0.90 | 0.90 | 0.81 | 0.96 | 0.81 | 0.97 |
| IMPT | 0.90 | 0.88 | 0.90 | 0.89 | 0.82 | 0.87 | 0.84 | 0.93 |
| INTD | 1.01 | 1.01 | 0.99 | 1.00 | 1.04 | 1.02 | 0.98 | 1.00 |
| Group 3 | | | | | | | | |
| LGDP | 0.95 | 0.98 | 0.90 | 0.89 | 0.92 | 0.78 | 0.74 | 0.75 |

This table reports CV criteria (absolute forecast errors) comparing of local estimates (8.7) against the factor invariant trivariate regression. Semiparametric CV estimates are given as a fraction of the pooled regression CV statistics.

a very large number of observations (up to 3100 for W97). Thus, moderate relative measures, varying between 0.90 and 0.97 say, may already signal a significant improvement of the invariant regression achieved by the local model. In some cases, the relative CV measures are clearly in favor of the local model. Conditioning, for instance, the SI relation on (the natural logarithm of) GDP when modeling F16 or E14 relative CV estimates are 0.74 and 0.75, respectively. And the CV estimates are also about 0.74 and 0.75 when SI relations are estimated conditional on GDP per capita for E11 and E14 accordingly. For the larger cross-sections W97 and L68, it is in particular the ratio of imports over GDP that provides the strongest improvement of the

pooled model. For this factor variable the relative CV measures are 0.90 and 0.88 for W97 and L68, respectively.

With regard to model evaluation by means of CV criteria, it is worthwhile mentioning that these statistics indicate overall model performance. Even if relative CV estimates are smaller than but close to unity it is still possible that over particular areas of the factor space local estimates differ significantly from corresponding quantities computed under an assumption of global homogeneity.

8.5 Implementation

In the last section it has been demonstrated that the factor dependent model outperforms the factor invariant model. This section discusses further the implementation of the factor dependent model.

The model formalized in (8.4) offers a local view at the SI relation conditional on a particular economic variable describing the state of an economy in two directions. As a consequence estimation results could be provided in terms of three dimensional graphs. However, since our interest here is focussed on some overall impact of a particular factor on the SI relation, estimation results from the model in (8.4) along particular paths of the state variables are displayed. The latter perspective has the advantage that estimation results can be provided in the familiar form of two dimensional functional estimates. To be explicit, estimates of the following local SI relations will be shown:

- (i) $\hat{\beta}_3(w^{(i)} = v, w^{(t)} = -1, 0, 1),$
- (ii) $\hat{\beta}_3(w^{(i)} = 0, w^{(t)} = v), v = -2 + 0.1k, k = 0, 1, 2, \dots, 40.$

Conditioning the evaluation of local estimates on states with either $w^{(i)} = 0$ or $w^{(t)} = 0$ provides different insights into the determinants of the SI rela-

tion that allow a classification into short and long run impacts. To get an intuition for the latter interpretations, the kernel based weighting schemes in (8.5) and (8.6) is discussed now in some more detail.

- Short run determinants

Conditional on $w^{(t)} = 0$ local SI relations are evaluated with putting higher weights on those members of a particular cross-section that follow closely the cross-sectional time trend (\bar{w}_t), as, for instance, the globally trending behavior towards an intensified exchange of goods. Similarly, conditional on positive (+1, say) or negative (-1) values of $w^{(t)}$, local SI relations are evaluated with those economies getting the highest weight which are above or below the factor specific trend. As a particular merit of the semiparametric approach it is noteworthy that the composition of the latter 'artificial' cross-sections is time dependent, i.e. the weighting scheme picks up effects of a country falling behind or keeping up with the global perspective. Apart from the time varying kernel weight, $K_{t,h}(\bullet)$, it is the 'inner factor variation' around its country specific trend that enters the local weighting scheme for the given country ($K_{i,h}(\bullet)$). In this sense, conditional estimates of the SI relation along (i) exploit short run factor variation. Since short run factor dependence might differ according to a countries' position relative to the cross-sectional average, it is tempting to compare various local estimates, conditioned upon $w^{(t)} = -1, 0, 1$ say.

- Long run determinants

Conditional on $w^{(i)} = 0$, country specific weights $K_{i,h}(\bullet)$ are the highest for those observations where a particular factor realization in country i is close to the long run time path characterizing this particular economy. Varying in the same time the location of $w^{(t)} = -2, \dots, 2$ allows to exploit 'outer factor variation' for quantifying local states

of the SI relation. In this case, the chosen support of $w^{(t)}$ will subsequently put high kernel weight, $K_{t,h}(\bullet)$, on those economies which are below, close to or above a factor's overall time path. Since changes of the latter relative positions are likely to reflect long term economic conditions or policy strategies, local SI relations conditional on (ii) are interpreted here as long run characteristics of the SI relation.

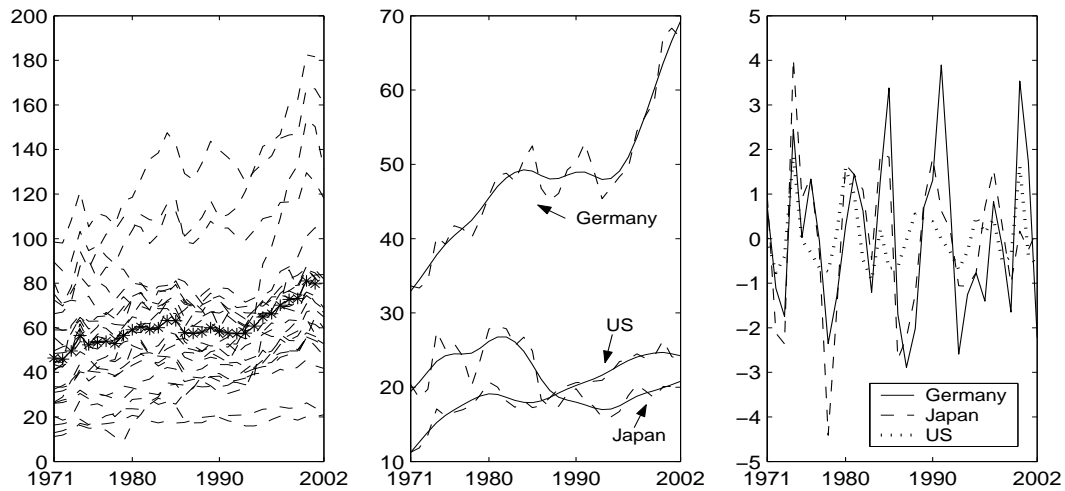


Figure 8.1: Dynamics of the openness ratio. The left hand side panel shows the observations for the countries in O26 (dashed line) and the corresponding cross-sectional averages (solid line with stars). The medium panel displays the openness degree of Germany, the US and Japan (dashed line), and the corresponding long run trend (solid line). The right hand side panel illustrates the deviations from the long run trend for the given three countries.

- An illustration

These perspectives of factor variation are illustrated for the case of the openness ratio (measured in %) in Figure 8.1. The left hand side panel shows time-series of the openness ratio (dashed line) for all countries in O26 jointly with the time path of the average openness degree (stars). The latter corresponds to $w^{(t)} = 0$. For three particular economies, Germany, the US and Japan, the openness ratio and it's corresponding country specific trend ($w^{(i)} = 0$) are shown in the

medium panel as dashed and solid curves, respectively. When evaluating long run dependence of the SI relation on openness, factor realizations close to the country specific trend enter the kernel regression with the highest weight. To estimate short run impacts of openness on the SI relation, factor variation around the long run trend (shown in right hand side panel of Figure 8.1) contributes to kernel based weighting while in the same time the relative location of a particular economy within the cross-section is fixed. Given the openness measure as displayed in the medium panel of Figure 8.1, it is likely that inner German variations get a higher/lower kernel based weight than factor variations measured for Japan or the US conditional on a relatively high/low degree of openness ($w^{(t)} = 1/w^{(t)} = -1$).

8.6 Inference

In this section, the factor based bootstrap approach introduced in the last chapter is adopted to infer on factor dependence against a structurally invariant model. Instead of the goodness-of-fit test statistic T_n , the confidence interval of the functional estimates for the factor invariant hypothesis are obtained by means of FB. The latter inference is of immediate interest since particular local pathes of the factor are chosen for the functional estimates in this study. Thus, T_n statistic requiring functional estimates conditional on every observations of the factors is not suitable. A further merit of using the confidence interval is that the (non)linear shape of the functional estimates conditional on the local path can be instantaneously visualized. The adopted approach is implemented along the following lines:

- 1) The local estimate in (8.7) can be seen as a function of the data and the

chosen bandwidth parameter, i.e.

$$\hat{\underline{\beta}}(\omega) = f\{y_{it}, z'_{it}, \omega_{it} = (w_{it}^{(i)}, w_{it}^{(t)}), h, i = 1, \dots, N, t = 1, \dots, T\}. \quad (8.9)$$

- 2) To distinguish the cases of factor dependence and factor invariance of the SI relation local estimates as given in (8.9) are compared with bootstrap counterparts

$$\hat{\underline{\beta}}^*(\omega) = f\{y_{it}, z'_{it}, \omega_{it}^* = (w_{it}^{(i*)}, w_{it}^{(t*)}), h, i = 1, \dots, N, t = 1, \dots, T\}, \quad (8.10)$$

where bivariate tuples $\omega_{it}^* = (w_{it}^{(i*)}, w_{it}^{(t*)})$ are drawn with replacement from the set of bivariate variables $w_{it} = (w_{it}^{(i)}, w_{it}^{(t)})$. Since sample information on the y_{it} and z'_{it} is not affected by the bootstrap the adopted scheme will disconnect any potential link between the selected factor variable on the one hand and the SI relation on the other hand. If the true underlying SI relation is state invariant estimates $\hat{\underline{\beta}}(\omega)$ and $\hat{\underline{\beta}}^*(\omega)$ are likely to differ only marginally over the support of the state variable.

- 3) Drawing a large number (for instance $R = 1000$) of bootstrap estimates $\hat{\underline{\beta}}^*(\omega)$ allows to decide if the null hypothesis of a state invariant SI relation can be rejected at some state $\omega = (w^{(i)}, w^{(t)})$. For this purpose, estimates $\hat{\underline{\beta}}_3(\omega)$ are contrasted with a confidence interval constructed from its bootstrap distribution $\hat{\underline{\beta}}_3^*(\omega)$. For this study, the 2.5% and 97.5% quantiles of $\hat{\underline{\beta}}_3^*(\omega)$ are used as a 95% confidence interval to hold for the parameter $\underline{\beta}_3$ under the null hypothesis of state invariance. Accordingly, we regard the actual estimate to differ locally from the unconditional relation with 5% significance if $\hat{\underline{\beta}}_3(\omega)$ is not covered by the bootstrap confidence interval.

8.7 Results for functional estimation

In this section, results obtained from the state dependent model (8.4) are reported. The discussion will not cover local estimates of the intercept ($\underline{\beta}_1(\omega)$) and trend parameter ($\underline{\beta}_2(\omega)$) of the model. Rather we will concentrate on the empirical features of the SI relation, i.e. on local estimates $\hat{\underline{\beta}}_3(\omega)$. As mentioned, the inclusion of a deterministic trend term in (8.4) was meant to allow an evaluation of factor impacts on the SI relation conditional on deterministic time features. The local model excluding the deterministic trend term is also estimated. Explicit results obtained from these exercises are not provided. It is confirmed that functional relationships turn out to be invariant in shape under inclusion or exclusion of a deterministic trend variable. For most factors, however, slopes of functional forms were more pronounced for the model without deterministic trend. In addition, evaluating estimation uncertainty by means of resampling schemes obtains confidence intervals for the SI relation which are throughout wider for the functional regression model including the deterministic trend term.

8.7.1 Factors impacting on saving or investment

Age dependency ratio

As displayed in the left hand side panels of Figure 8.2, the age dependency ratio affects significantly the SI relation in the long run for all displayed cross-sections of developed countries (E14, F16, E11, O15, O26, and D29). Conditional on country specific long run trends ($w^{(i)} = 0, w^{(t)} = [-2, \dots, 2]$), the empirical SI relation is decreasing in the time specific age dependency ($w_{it}^{(t)} = (\tilde{w}_{it} - \bar{w}_t)/\sigma_t(\tilde{w})$).

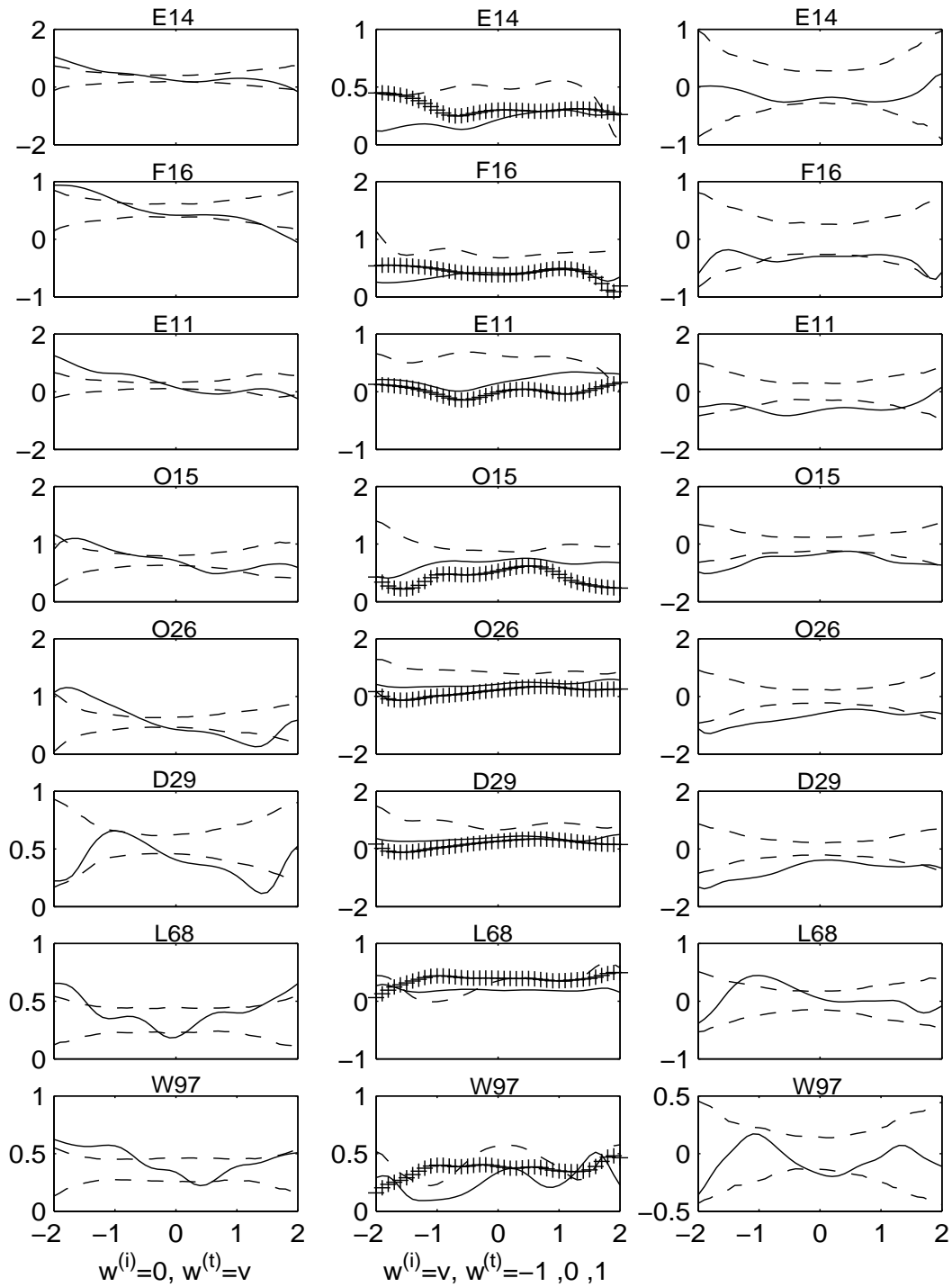


Figure 8.2: Functional estimates of the SI relation conditional on the age dependency ratio for five selected cross sections. The left hand side panels show the estimated long run effects $\hat{\beta}_3(w^{(i)} = 0, w^{(t)} = v)$ (solid). Dashed lines are the corresponding 95% confidence intervals. The medium panels display the short run effects for three local paths, i.e. $\hat{\beta}_3(w^{(i)} = v, w^{(t)} = 0)$ (solid line), $\hat{\beta}_3(w^{(i)} = v, w^{(t)} = -1)$ (dashed or minus line) and $\hat{\beta}_3(w^{(i)} = v, w^{(t)} = 1)$ (plus line). The right hand side panels show the difference between two estimated short run effects (solid), i.e. $\hat{\beta}_3(w^{(i)} = v, w^{(t)} = 1) - \hat{\beta}_3(w^{(i)} = v, w^{(t)} = -1)$, and the corresponding 95% confidence intervals (dashed).

The observed negative impact of the age dependency ratio on the SI relation is consistent with the “Life Cycle Hypothesis” (LCH) suggested by Modigliani and Brumberg (1954). According to this theory, consumption or saving is affected by the age distribution of the population. Most households do not have a constant flow of income over their lifetimes. In order to smooth their consumption path, young agents should borrow and retired agents shall finance themselves from their past savings. Therefore, if the age dependency ratio, the ratio of the dependent population to the working-age population, is high, the aggregate saving rate shall be low. The latter might disconnect the links between domestic saving and investment. In the empirical literature (Modigliani 1970, Masson et al. 1998) the influence of the age dependency ratio on the saving ratio has been mainly confirmed by means of studies with cross-country or pooled data.

Regarding the level of the functional SI relations, it is worthwhile to point out that the between estimates given in Table 6.1 are likely not representative for the entire cross-sections. For instance, the estimated between coefficient for E11, $\hat{\beta}_3 = -0.16$, is far below the SI relation measured over states of a relatively low age dependency ratio. As such, homogeneous models like the between regression (2.1) run the risk of providing biased approximations of the link between domestic saving and investment. Note that this caveat of a homogeneous model formalization may also be illustrated with other potential factor variables.

For less developed economies (L68), the estimated SI relation shows a U-shaped behavior when interpreted as a function of the age dependency ratio. To explain this evidence, one may conjecture that for less developed economies age dependency affects saving (consumption smoothing) and investment (growth prospect) in a more symmetric fashion than implied by the LCH for developed economies. As the most comprehensive cross-section, the results for the long run SI relation given for W97 can be

seen as an aggregate over the features of developed (D29) and less developed (L68) economies with the latter introducing some mild, i.e. insignificant, U-shaped pattern. In sum, the results for W97 underscore that the negative impact of age dependency on the SI relation dominates according to the significantly decreasing functional estimates over the factor support $-2 \leq w^{(t)} \leq 0.5$.

Effects of short run variations in the age dependency on the SI relation are not observed (medium panels of Figure 8.2). Conditional on $(w^{(i)} = [-2, \dots, 2], w^{(t)} = -1, 0, 1)$ the estimated functional forms are more or less constant. However, comparing conditional estimates for $w^{(t)} = 1$ and $w^{(t)} = -1$, it turns out that the former are almost uniformly below the latter for all developed economies. The right hand side panels show the difference between these two estimated short run effects, i.e. $\hat{\beta}_3(w^{(i)} = v, w^{(t)} = 1) - \hat{\beta}_3(w^{(i)} = v, w^{(t)} = -1)$, and the corresponding 95% confidence intervals. The significantly negative difference is confirmed for E11 and F16 over the supports $-1 \leq w^{(i)} \leq 1$, for O26 and D29 given $-2 \leq w^{(i)} \leq 1.6$, and for O15 over $-2 \leq w^{(i)} \leq 2$.

Similar to the latter results on the short run behavior of the SI relation conditional on age dependency, analyses conditional on other factors also reveal that the link between domestic saving and investment is mostly stable in response to inner country factor variation. For this reason, the following discussion will focus on the functional relations characterizing the SI relation in the long run.

Population growth

Following Obstfeld (1986), population growth might govern saving as well as investment and thereby explain a high positive correlation between the latter variables. Long run effects of population growth on the SI relation for W97 are shown in the medium panel of Figure 8.3. Apart from boundary

effects, the conditional estimates are well stabilized around the between estimates $\hat{\beta}_3 = 0.43$ documented in Table 6.1. A clear trending pattern of the functional estimates cannot be diagnosed.

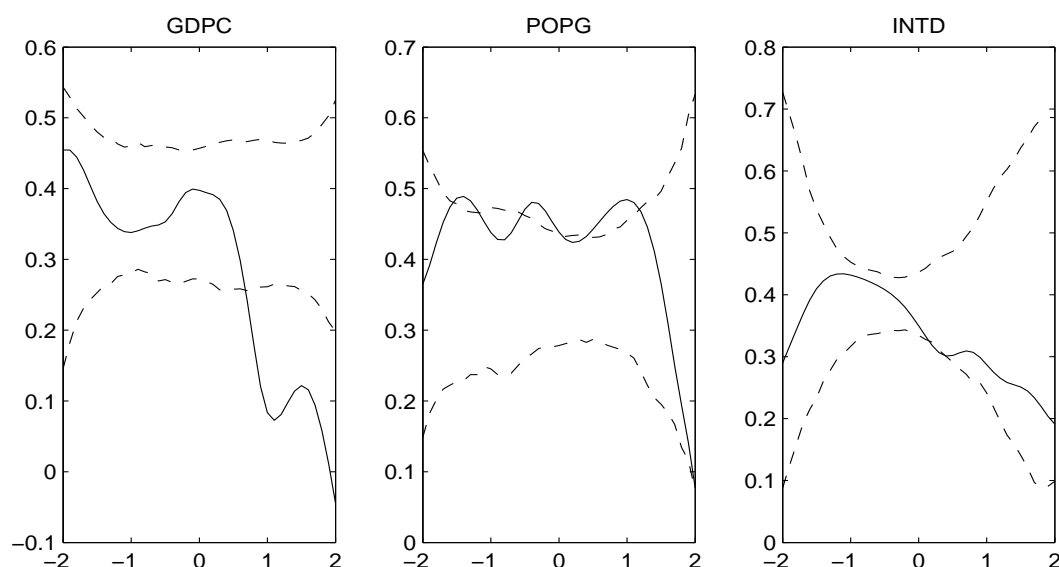


Figure 8.3: Functional estimates of the SI relation conditional on per capita income, the population growth rate and the absolute real interest rate differential. Estimated long run effects $\hat{\beta}_3(w^{(i)} = 0, w^{(t)} = v)$ are displayed for W97. The solid line shows the point estimates and the two dashed lines are the corresponding 95% confidence intervals.

Per capita income

As a potential measure of an economies' state of development, the impact of global variation (W97) of per capita income on the SI relation is shown in the left hand side panel of Figure 8.3. From a-priori reasoning one may expect that for a less developed country, the domestic investment ratio is high in response to high rates of return, and the domestic saving ratio is lower owing to a high growth prospect. In contrast, for rich industrial economies with high per capita income, the domestic investment ratio is low because of low rates of return, and the domestic saving ratio is high owing to a low growth prospect. Hence, a hump-shaped SI relation is expected conditional

on an increase of per capita income. Our empirical evidence on the impact of per capita income on the SI relation confirms the latter considerations merely to some extent. Conditional on economies having per capita income above the cross-sectional average (W97), a significantly decreasing trend is visible from Figure 8.3. Functional estimates, similar to W97 in shape as well as in level, are found for L68. For the remaining cross-sections a hump like pattern cannot be detected which might be addressed to a higher degree of factor homogeneity within these subsamples.

Fiscal variables

Firstly, the government budget balance is considered as a fiscal variable which might have an influence on the SI relation. A full offset of private saving to government deficits (Ricardian equivalence) is generally rejected in the empirical literature. Bernheim (1987) shows that a unit increase in the government deficit is related with a decrease in consumption of 0.5 to 0.6. This evidence supports the view that government deficits might be positively correlated with current account deficits, thereby describing so-called "Twin Deficits". Based on this argument, we shall expect a hump-shaped SI relation conditional on an increasing government budget balance since a high current account imbalance is consistent with a low SI relation. Although a significant left part of a hump shape is found for developed economies (D29), a significant influence of the government budget balance on the SI relation is not observed for W97, as can be seen in the left hand side panels of Figure 8.4.

In the second place, the influence of the composition of government expenditures on the SI relation is investigated. As can be seen from the upper right hand side panel of Figure 8.4, a significantly decreasing estimated SI relation is obtained for W97 conditional on increasing total government expenditure. For the remaining cross-sections, similar effects are found.

A high government deficit spending might higher interest rates, and thus crowd out private investment and induce an increase in private saving. Furthermore, a partial offset of private saving to government deficit spending due to the expectation of increasing future tax may provoke a further increase in private saving. Therefore, high government spending might be related with a low SI relation. When total government expenditures are decomposed to government capital, current and consumption expenditure for W97, significantly decreasing functional estimates are also obtained for the latter two components (the lower right hand side panels in Figure 8.4). Since government capital expenditure is generally viewed as productive, increasing future taxes might not be expected, which leaves private saving unaffected. Therefore, no significant influence of the government capital expenditure on the SI relation is observed.

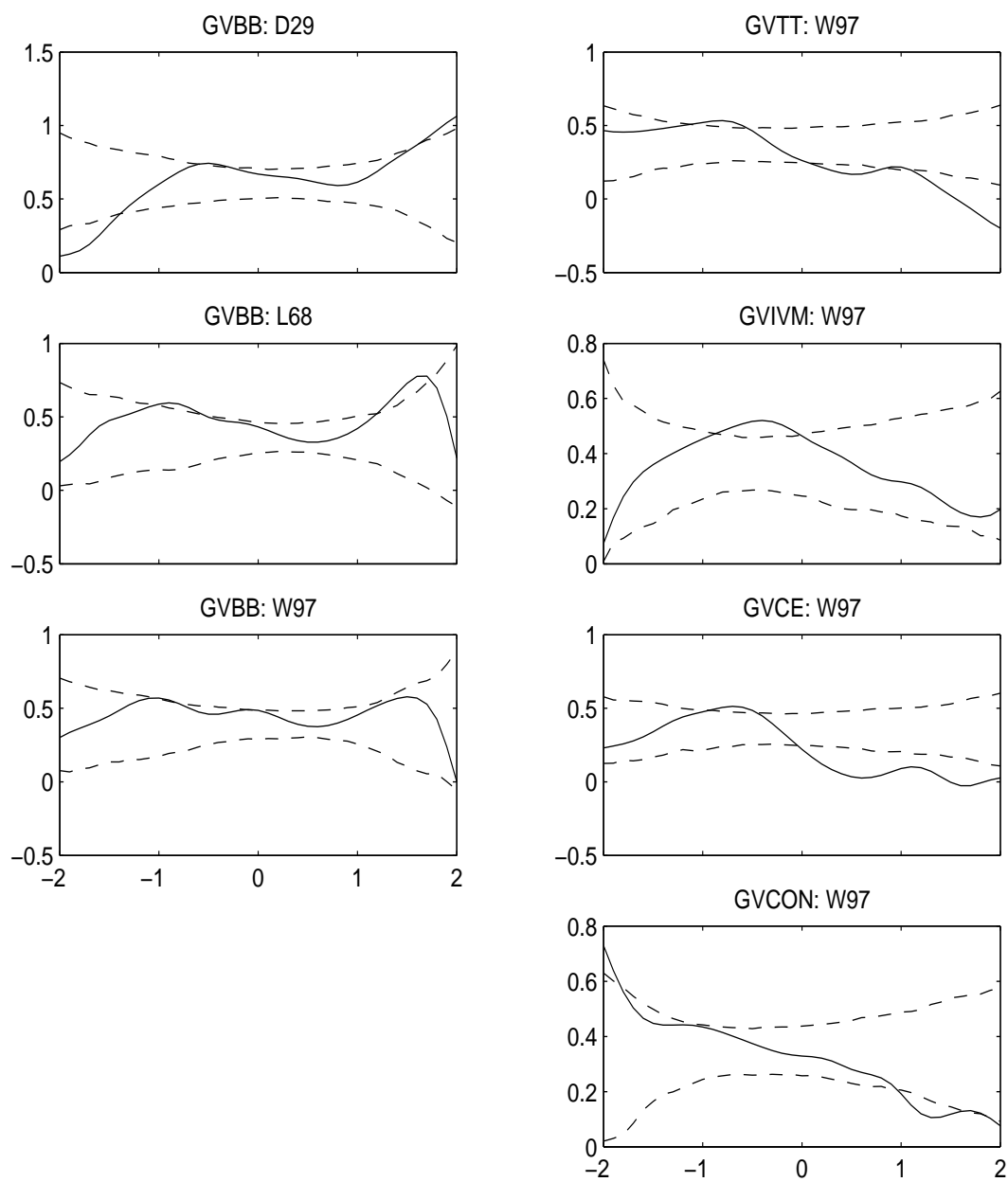


Figure 8.4: Functional estimates of the SI relation conditional on fiscal variables. The long run effects $\hat{\beta}_3(w^{(i)} = 0, w^{(t)} = v)$ are displayed. The solid line shows the point estimates and the dashed lines are the corresponding 95% confidence intervals.

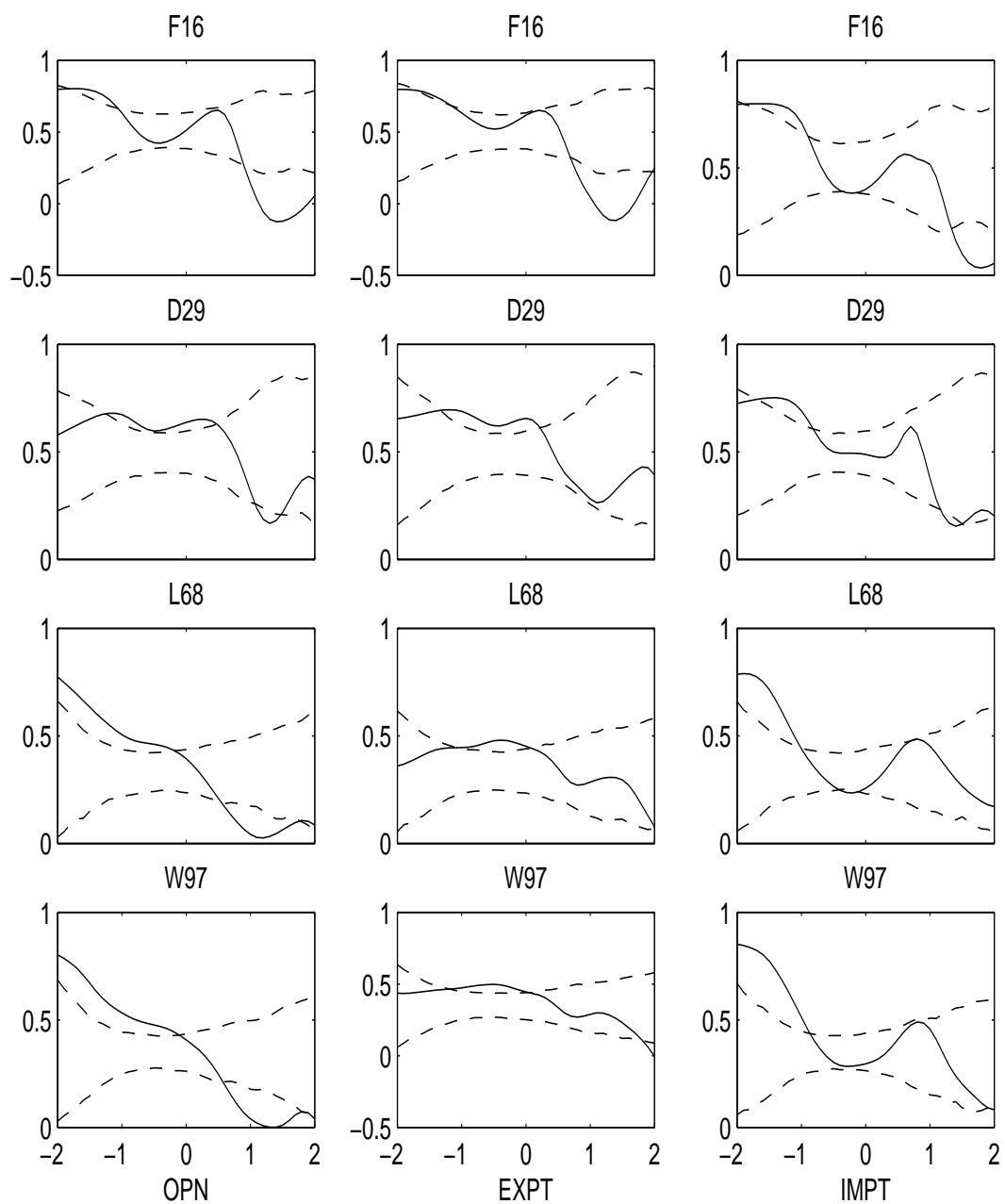


Figure 8.5: Functional estimates of the SI relation conditional on the openness ratio, the ratio of exports and imports to GDP for four selected cross sections. Estimated long run effects $\hat{\beta}_3(w^{(i)} = 0, w^{(t)} = v)$ are displayed. The solid line shows the point estimates and the two dashed lines are the corresponding 95% confidence intervals.

8.7.2 Factors measuring integration of markets

Openness

Conditioning the SI relation on the long run path of an economies' openness measured as the sum of imports and exports over GDP obtains significantly decreasing functional estimates for W97 (Figure 8.5). The latter reflects that domestic investment is naturally bounded by domestic saving for a closed economy. Separating W97 in its two divisions D29 and L68, it is found that the overall trend is most obvious for the group of less developed economies. The latter impression might mirror that L68 is likely more heterogenous with regard to country specific degrees of openness. Considering the 16 OECD members in F16, nevertheless, the decreasing trend is evident as well. When alternatively decomposing the openness measure in its two components, exports over GDP and imports over GDP, we obtain that the common factor results documented for more closed economies ($-2 < w^{(t)} < -1$) are most obvious for the import over GDP measure. At the opposite, for more open economies ($1 < w^{(t)} < 2$), it appears to be the export over GDP component having the strongest impact on the SI relation evaluated conditional on openness. By construction, the 'openness' variable is a measure reflecting good markets integration. As such, our results for the conditional SI relation motivate the view that the SI relation is perhaps not only reflecting capital market separation as stated by FH (1980) but also barriers of international trade (Obstfeld and Rogoff 2000).

Interest rate parity

Having discussed the impact of openness as a measure of good markets integration on the SI relation it is also tempting to relate this relation to absolute real interest differential approximating capital market integration. As documented in the right hand side panel of Figure 8.3, a significant

impact of the absolute real interest rate differential on the SI relation for W97 is not found in our analysis for the global perspective which is also representative for all remaining cross-sections. By using the real interest rate differential, we are aware that this measure might not only correspond to capital mobility, as argued by Frankel (1992). As another potential measure of capital mobility, the nominal interest rate differential is considered. However, significant impacts of this measure on the SI relation are also not obtained.

8.7.3 Large country effect

As can be seen in Figure 8.6, significantly positive long run impacts of the log GDP on the SI relation can be diagnosed for developed economies (as D29 and F16) and W97, thereby supporting a large country effect. A large country might have a higher SI relation than a small country owing to an endogenous domestic interest rate. For the cross-section of less developed economies the large country effect is not confirmed. This might be expected given that L68 collects small economies by definition.

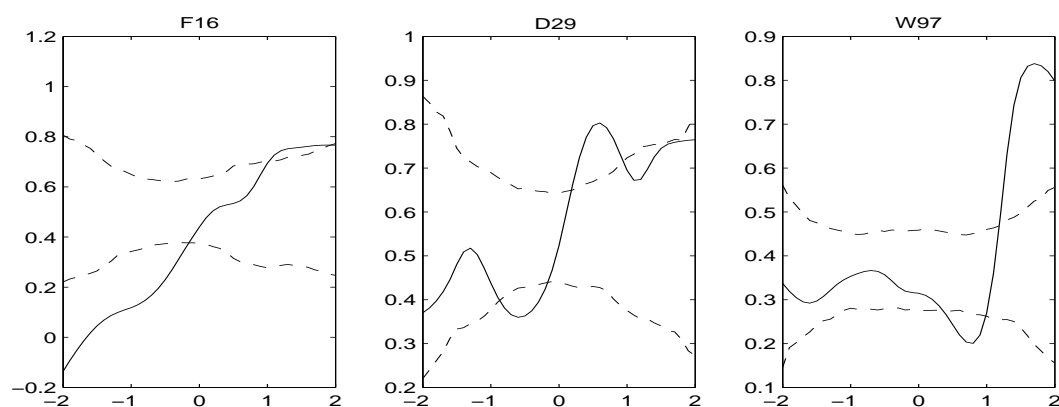


Figure 8.6: Functional estimates of the SI relation conditional on the natural logarithm of GDP. Estimated long run effects $\hat{\beta}_3(w^i = 0, w^t = v)$ are displayed. The solid line shows the point estimates and the two dashed lines are the corresponding 95% confidence intervals.

To summarize, according to functional coefficient estimation the SI relation is found to be rather stable in the short run, but factor dependent in the long run. A low SI relation might be due to a high degree of the trade openness, high age dependency ratio, or high government current and consumption expenditures. In addition, small countries tend to have a lower SI relation in comparison with larger economies.

8.7.4 Implications

In light of the diagnosed factor dependent nature of empirical SI relations it is natural to address its' potential implications. Two issues are discussed in turn.

Firstly, the increasing integration of capital markets alone does not automatically decrease the correlation between domestic saving and investment. The integration of global capital markets is necessary but not sufficient for a high net capital in(out)-flow and thus a low SI relation. The extent to which the domestic saving and investment are disconnected is depending on variables as the openness ratio, the age dependency ratio and government current and consumptions expenditures according to our analysis. Economies with high age dependency ratios or high exports have more money to lend and thus seek internationally the highest return. Comparably, economies with high government consumption and current expenditures or high imports might have more incentive to borrow internationally at the lowest costs. For these economies the SI relation may be low.

Secondly, since a low SI relation tend to correspond to a high current account imbalance, determinants of the SI relation can also be regarded as determinants of the current account. Based on our results, high government current and consumption expenditure may induce a high current account imbalance for most economies. For OECD countries, a high current account

imbalance might also mirror a high age dependency ratio. Furthermore, the increasing degree of openness in good markets might provide countries with the possibility to sustain long run current account imbalances.

8.8 Summary

In this chapter a new framework of bivariate functional coefficient models is applied to estimate conditional SI relations. The factor based resampling scheme is applied to address inferential issues. This bivariate functional approach allows to separate factor dependence of the SI relation in the short and long run. In the short run, the factor dependent SI relations are found to be rather stable. In the long run, however, a set of economic factors is found to impact the SI relation. The latter are an economies' openness ratio, the age dependency ratio and government expenditures. Supporting evidence for the large country effect on the SI relation is also found.

Chapter 9

Concluding remarks

The influence on the link between domestic saving and investment measured as a fraction of GDP has drawn extensive attentions in both theoretical and empirical literature since Feldstein and Horioka (1980). This thesis develops a new semi-parametric approach along with factor based bootstrap inference to analyse the determinants of the saving-investment relation. Annual data from 1971 to 2002 for eight cross-sections comprising the world economy, developing countries, the OECD, the EU and the Euro area are applied.

In Chapter 4 we contrast tests on nonstationarity and bounded nonstationarity of current account imbalances for 26 OECD countries. The standard approach of unit root testing by means of the Phillips-Perron statistic ignores the limited support of current account imbalances expressed as percentages of the gross domestic product. Taking account of the bounded nature of actual current account processes weakens the evidence against the prevalence of stochastic trends, in every respect. At the aggregate level the null hypothesis of bounded integration cannot be rejected for the current account imbalances. Therefore, policy controls or economic crises might cause mean reverting patterns of CA imbalances on the one hand. On the other hand, however, they could hide important stochastic features point-

ing to unsustainability of the CA for numerous developed economies. The direct implication of these results is that saving and investment are likely not cointegrated. The high association between saving and investment may be partially due to the existence of the bounds on the current account imbalances. Given this evidence, we refrain from viewing error correction models for saving and investment as derived from a system of cointegrating variables. Rather we focus on its empirical performance in comparison with static panel based formalizations of the saving-investment relation.

Performances of various specifications for the saving-investment relation are compared by means of cross-validation criteria in Chapter 5. From static model performance it is found that the best performing parametric description of the saving-investment relation is cross-section specific. Thus, saving-investment relations might be subject to country specific economic conditions and policies. Contrasting static and weakly dynamic model formalizations we find no hint at the necessity of a weakly dynamic model specification. This evidence might be due to individual effects governing investment ratios. Moreover, adding an error correction term in dynamic models does not improve model performance.

Chapter 6 provides a preliminary view on the empirical features of the saving-investment relation given our data sets. By means of between, pooled, cross-sectional and time-series regressions, time decreasing and country heterogenous estimates of the saving-investment relations are found. Regressing the estimated cross-section specific saving-investment relation on potential economic factors, the age dependency ratio and the openness ratio are diagnosed as influencing the saving-investment relation significantly. As such, the panel-heterogenous saving-investment relation might be due to economic state variables, rather than a purely deterministic feature or some random processes.

Before applying functional coefficient models to analyze the factor de-

pendency of the saving-investment relation, we introduce the factor based bootstrap approach to test parameter invariance against a functional alternative in Chapter 7. Modeling the functional coefficient parametrically, the bootstrap approximation of the distribution of the goodness-of-fit test statistic is shown to be valid asymptotically under the null hypothesis. The new bootstrap scheme is shown to be robust against heteroskedastic error distributions. Furthermore, for small samples factor based inference outperforms the wild bootstrap and the pairs bootstrap counterpart that are also known to cope with heteroskedasticity. When the functional coefficient is modeled nonparametrically, only tests via the factor based bootstrap provide correct empirical size estimates.

In Chapter 8, the framework of bivariate functional coefficient models is introduced and applied to estimate conditional saving-investment relations. This approach enables separating the factor dependence of the saving-investment relation in the short and long run. In the short run, the factor dependent saving-investment relations are found to be rather stable. However, in the long run a set of economic factors is found to impact the saving-investment relation. The latter are an economies' openness ratio, the age dependency ratio and government expenditures. Supporting evidence for the large country effect on the saving-investment relation is also found. According to these results, the interpretation of a high saving-investment relation as a signal for low capital mobility has to be treated with care. Empirically high saving-investment relations could reflect goods market friction, demographic development or fiscal consolidation rather than being puzzling.

As an issue for future research, it is tempting to develop theoretic models to explain the influence of variables such as the openness ratio on the saving-investment relation.

Bibliography

- Abbott, A. and Vita, G. D. (2003). Another piece in the Feldstein-Horioka puzzle, *Scottish Journal of Political Economy* **50**: 69–89.
- Afxentiou, P. C. and Serletis, A. (1993). International capital mobility and the long run investment and saving in canada, *Economia-Internationale* **46**: 147–167.
- Alexakis, P. and Apergis, N. (1994). The Feldstein-Horioka puzzle and exchange rate regimes: evidence from cointegration test, *Journal of Policy Modeling* **16**: 459–472.
- Allen, D. M. (1974). The relationship between variable selection and data augmentation and a method for prediction, *Technometrics* **16**: 125–127.
- Amirkhalkhali, S. and Dar, A. A. (1993). Testing for capital mobility: A random coefficients approach, *Empirical Economics* **18**: 523–541.
- AmirKhalkhali, S., Dar, A. A. and AmirKhalkhali, S. (2003). Saving-investment correlations, capital mobility and crowding out: some further results, *Economic Modelling* **20**: 1137–1149.
- Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**: 817–858.

- Artis, M. and Bayoumi, T. (1992). Global capital market integration and the current account, in M. Taylor (ed.), *Money and Financial Markets*, Cambridge, MS and Oxford: Blackwell, pp. 297–307.
- Backus, D., Kehoe, P. and Kydland, F. (1992). International real business cycles, *Journal of Political Economy* **100**: 745–775.
- Bajo-Rubio, O. (1998). The saving-investment correlation revisited: The case of Spain, 1964-1994, *Applied Economics Letters* **5**: 769–772.
- Balke, N. S. and Fomby, T. B. (1997). Threshold cointegration, *International Economic Review* **38**: 627–645.
- Banerjee, A. and Zanghieri, P. (2003). A new look at the Feldstein-Horioka puzzle using an integrated panel, *CEPII Working Paper 2003-22*.
- Baxter, M. and Crucini, M. J. (1993). Explaining saving-investment correlations, *The American Economic Review* **83**: 416–436.
- Bayoumi, T. (1989). Saving-Investment correlations: Immobile capital, government policy or endogenous behavior?, *IMF Working paper WP/89/66*.
- Beran, R. (1988). Prepivoting test statistics: A bootstrap view of asymptotic refinements, *Journal of the American Statistical Association* **83**: 687–697.
- Bernheim, B. D. (1987). Ricardian equivalence: An evaluation of theory and evidence, *NBER Macroeconomics Annual 1987*, Cambridge, Mass.: MIT Press.
- Blanchard, O. and Giavazzi, F. (2002). Current account deficits in the euro area. the end of the feldstein horioka puzzle?, *MIT Department of Economics Working Paper 03-05*.

- Cai, Z., Fan, J. and Yao, Q. (2000). Functional-coefficient regression models for nonlinear time series, *Journal of the American Statistical Association* **95**: 941–956.
- Cavaliere, G. (2005). Limited time series with a unit root, *Econometric Theory* **21**: 907–945.
- Cavaliere, G. and Xu, F. (2007). ADF tests for a unit root in bounded time series, *mimeo*, Christian-Albrechts-University of Kiel.
- Chen, R. and Tsay, R. (1993). Functional-coefficient autoregressive models, *Journal of the American Statistical Association* **88**: 298–308.
- Chinn, M. and Prasad, E. S. (2003). Medium-term determinants of current accounts in industrial and developing countries: An empirical exploration, *Journal of International Economics* **59**: 47–76.
- Cleveland, W. S., Grosse, E. and Shyu, W. M. (1992). Local regression models, in J. Chambers and T. Hastie (eds), *Statistical Models in S*, Wadsworth and Brooks, Pacific Grove, California, pp. 309–376.
- Coakley, J., Fuertes, A.-M. and Spagnolo, F. (2004). Is the Feldstein-Horioka puzzle history?, *The Manchester School* **72**: 569–590.
- Coakley, J. and Kulasi, F. (1997). Cointegration of long span saving and investment, *Economics Letters* **54**: 1–6.
- Coakley, J., Kulasi, F. and Smith, R. (1996). Current account solvency and the Feldstein-Horioka puzzle, *Economic Journal* **106**: 620–627.
- Coakley, J., Kulasi, F. and Smith, R. (1998). The Feldstein-Horioka puzzle and capital mobility: A review, *International Journal of Finance and Economics* **3**: 169–188.

- Coiteux, M. and Olivier, S. (2000). The saving retention coefficient in the long run and in the short run, *Journal of International Money and Finance* **19**: 535–548.
- Corbin, A. (2001). Country specific effect in the Feldstein-Horioka paradox: a panel data analysis, *Economics Letters* **72**: 297–302.
- Davidson, R. and Flachaire, E. (2001). The wild bootstrap, tamed at last, *Working Papers 1000*, Queen's University, Department of Economics.
- Debelle, G. and Faruquee, H. (1996). What determines the current account? a cross-sectional and panel approach, *Technical Report 96/58*, IMF Working Papers.
- Demetrescu, M., Hassler, U. and Tarcolea, A.-I. (2006). Combining significance of correlated statistics with application to panel data, *Oxford Bulletin of Economics and Statistics* **68**: 647–663.
- Dooley, M., Frankel, J. and Mathieson, D. J. (1987). International capital mobility: What do saving-investment correlations tell us?, *IMF Staff Papers* **34**: 503–529.
- Edwards, S. (1995). Why are saving rates so different across countries?: An international comparative analysis, *NBER Working Paper 5097*.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *The Annals of Statistics* **7**: 1–26.
- Feldstein, M. (1983). Domestic saving and international capital movements in the long run and the short run, *European Economic Review* **21**: 129–151.

- Feldstein, M. and Bachetta, P. (1991). National saving and international investment, in D. Bernheim and J. Shoven (eds), *National Saving and Economic Performance*, Chicago: University of Chicago Press, pp. 201–226.
- Feldstein, M. and Horioka, C. (1980). Domestic saving and international capital flows, *Economic Journal* **90**: 314–329.
- Frankel, J. A. (1985). International capital mobility and crowding out in the u.s. economy: imperfect integration of financial markets or of goods markets?, *NBER Working Paper* 1773.
- Frankel, J. A. (1992). Measuring international capital mobility: A review, *American Economic Review* **82**(2): 197–202.
- Freedman, D. A. (1981). Bootstrapping regression models, *The Annals of Statistics* **9**: 1218–1228.
- French, K. and Poterba, J. (1991). Investor diversification and international equity markets, *American Economic Review* **81**: 222–226.
- Geisser, S. (1975). The predictive sample reuse method with applications, *Journal of the American Statistical Association* **70**: 320–328.
- Ghosh, A. R. (1995). International capital mobility amongst the major industrialised countries: Too little or too much, *The Economic Journal* **105**.
- Glick, R. and Rogoff, K. (1995). Global versus country specific productivity shocks and the current account, *Journal of Monetary Economics* **35**: 159–192.
- Gulley, O. D. (1992). Are saving and investment cointegrated?, *Economics Letters* **39**: 55–58.
- Haan, J. D. and Siermann, C. L. (1994). Saving, investment, and capital mobility. a comment on Leachman, *Open economies review* **5**: 5–17.

- Hall, P. and Titterton, D. M. (1989). The effect of simulation order on level accuracy and power of Monte-Carlo tests, *Journal of the Royal Statistical Society, Series B*, **51**: 459–467.
- Harberger, A. C. (1980). Vignettes on the world capital market, *American Economic Association* **70**: 331–337.
- Härdle, W., Hall, P. and Marron, J. S. (1988). How far are automatically chosen regression smoothing parameters from their optimum?, *Journal of the American Statistical Association* **83**: 86–99.
- Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits, *The Annals of Statistics* **21**: 1926–1947.
- Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and semiparametric models*, Springer, Berlin.
- Harrison, J. M. (1985). *Brownian motion and stochastic flow systems*, Wiley.
- Hartung, J. (1999). A note on combining dependent tests of significance, *Biometrical Journal* **41**: 849–855.
- Hastie, T. and Tibshirani, R. (1993). Varing-coefficient models, *Journal of the Royal Statistical Society, Series B*, **55**: 757–796.
- Herwartz, H. and Xu, F. (2006a). Panel data model comparison for the investigation of the saving-investment relation, *Applied Economic Letters*, forthcoming.
- Herwartz, H. and Xu, F. (2006b). Reviewing the sustainability/stationarity of current account imbalances with tests for bounded integration, *The Manchester School*, forthcoming.

- Herwartz, H. and Xu, F. (2007a). A functional coefficient model view of the Feldstein-Horioka puzzle, *Economics Working Paper 2007-14*, Christian-Albrechts-University of Kiel.
- Herwartz, H. and Xu, F. (2007b). A new approach to bootstrap inference in functional coefficient models, *Economics Working Paper 2007-15*, Christian-Albrechts-University of Kiel.
- Ho, T.-W. (2002a). The Feldstein-Horioka puzzle revisited, *Journal of International Money and Finance* **21**: 555–564.
- Ho, T.-W. (2002b). A panel cointegration approach to the investment-saving correlation, *Empirical Economics* **27**: 91–100.
- Hodrick, R. J. and Prescott, E. C. (1997). Postwar u.s. business cycles: An empirical investigation, *Journal of Money, Credit and Banking* **29**: 1–16.
- Hoffmann, M. (2004). International capital mobility in the long run and the short run: can we still learn from saving-investment data, *Journal of International Money and Finance* **23**: 113–131.
- Horowitz, J. L. and Härdle, W. (1994). Testing a parametric model against a semiparametric alternative, *Econometric Theory* **10**: 821–848.
- Im, K. S., Pesaran, M. H. and Shin, Y. (2002). Testing for unit roots in heterogeneous panels, *Journal of Econometrics* **115**: 53–74.
- Jansen, W. J. (1996). Estimating saving-investment correlations: evidence for OECD countries based on an error correction model, *Journal of International Money and Finance* **15**: 749–781.
- Jansen, W. J. (1997). Can the intertemporal budget constraint explain the Feldstein-Horioka puzzle?, *Economics Letters* **56**: 77–83.

- Jansen, W. J. (1998). Interpreting saving-investment correlations, *Open Economies Review* **9**: 205–217.
- Jansen, W. J. (2000). International capital mobility: evidence from panel data, *Journal of International Money and Finance* **19**: 507–511.
- Jansen, W. J. and Schulze, G. G. (1996). Theory-based measurement of the saving-investment correlation with an application to Norway, *Economic Inquiry* **34**: 116–132.
- Johansen, S. (1992). Cointegration in partial systems and the efficiency of single-equation analysis, *Journal of Econometrics* **52**: 389–402.
- Kao, C. and Chiang, M. H. (2001). On the estimation and inference of a cointegrated regression in panel data, in B. Baltagi (ed.), *Advances in Economics*, pp. 179–222.
- Karfakis, C. and Kim, S.-J. (1995). Exchange rates, interest rates and current account news: Some evidence from Australia, *Journal of International Money and Finance* **14**: 575–595.
- Kasuga, H. (2004). Saving-investment correlations in developing countries, *Economics Letters* **83**: 371–376.
- Krol, R. (1996). International capital mobility: evidence from panel data, *Journal of International Money and Finance* **15**: 467–474.
- Leachman, L. L. (1991). Saving, investment, and capital mobility among OECD countries, *Open economies review* **2**: 137–163.
- Lemmen, J. J. and Eijffinger, S. C. (1995). The quantity approach to financial integration: The Feldstein-Horioka criterion revisited, *Open economies review* **6**: 145–165.

- Liu, R. (1988). Bootstrap procedures under some non i.i.d. models, *The Annals of Statistics* **16**: 1696–1708.
- Maddala, G. and Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test, *Oxford Bulletin of Economics and Statistics* **61**: 653–670.
- Mamingi, N. (1994). savings-investment correlations and capital mobility in developing countries, *The World Bank Policy Research Working Paper* 1211.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models, *The Annals of Statistics* **21**: 255–285.
- Masson, P. R., Bayoumi, T. and Samiei, H. (1998). International evidence on the determinants of private saving, *The World Bank Economic Review* **12**: 483–501.
- McKinnon, R. I. (1981). The exchange rate and macroeconomic policy: Changing postwar perceptions, *Journal of economic literature* **19**: 531–557.
- Mendoza, E. G. (1991). Real business cycles in a small open economy, *American Economic Review* **81**: 797–818.
- Milesi-Ferretti, G. M. and Razin, A. (1996). Sustainability of persistent current account deficits, *NBER Working Paper* 5467.
- Milesi-Ferretti, G. M. and Razin, A. (1997). Sharp reductions in current account deficits: An empirical analysis, *NBER Working Paper* 6310.
- Milesi-Ferretti, G. M. and Razin, A. (1998). Current account reversals and currency crises: Empirical regularities, *NBER Working Paper* 6620.

- Miller, S. M. (1988). Are saving and investment cointegrated?, *Economics Letters* **27**: 31–34.
- Modigliani, F. (1970). The life-cycle hypothesis of saving and intercountry differences in the saving ratio, in W. Eltis, M. Scorr and J. Wolfe (eds), *Induction, Trade, and Growth: Essays in Honour of Sir Roy Harrod*, Clarendon Press, Oxford.
- Modigliani, F. and Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross-section data, in K. K. Kurihara (ed.), *Post-Keynesian Economics*, Rutgers University Press.
- Moreno, R. (1997). Saving-investment dynamics and capital mobility in the Us and Japan, *Journal of International Money and Finance* **16**: 837–863.
- Murphy, R. G. (1984). Capital mobility and the relationship between saving and investment rates in oecd countries, *Journal of International Money and Finance* **3**: 327–342.
- Nadaraya, E. A. (1964). On estimating regression, *Theory of Probability and its Applications* **10**: 186–190.
- Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models, *Journal of the Royal Statistical Society, Series A*, **135**: 370–384.
- Obstfeld, M. (1986). Capital mobility in the world economy: Theory and measurement, *Carnegie-Rochester Conference Series on Public Policy* **24**: 55–104.
- Obstfeld, M. and Rogoff, K. (2000). The six major puzzles in international macroeconomics: Is there a common cause?, in B. S. Bernanke and K. Rogoff (eds), *NBER Macroeconomics Annual 2000*, The MIT Press, pp. 339–390.

- Özmen, E. and Parmaksiz, K. (2003a). Exchange rate regimes and the Feldstein-Horioka puzzle: The French evidence, *Applied Economics* **35**: 217–222.
- Özmen, E. and Parmaksiz, K. (2003b). Policy regime change and the Feldstein-Horioka puzzle: The UK evidence, *Journal of Policy Modeling* **25**: 137–149.
- Payne, J. E. and Kumazawa, R. (2006). Capital mobility and the Feldstein-Horioka puzzle: Re-examination of less developed countries, *The Manchester School* **74**: 610–616.
- Pelagidis, T. and Mastroiannis, T. (2003). The saving-investment correlation in Greece, 1960-1970: Implications for capital mobility, *Journal of Policy Modeling* **25**: 609–616.
- Pelgrin, F. and Schich, S. (2004). National saving-investment dynamics and international capital mobility, *Working Paper 2004-14*, Bank of Canada.
- Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics* **68**: 79–113.
- Phillips, P. C. B. (1987). Time series regression with a unit root, *Econometrica* **55**: 277–301.
- Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika* **75**: 335–346.
- Ravn, M. O. and Uhlig, H. (2002). On adjusting the HP-filter for the frequency of observations, *Review of Economics and Statistics* **84**: 371–380.
- Sachs, J. (1981). The current account and macroeconomic adjustment in the 1970s, *Brookings Papers on Economic Activity* **1**: 201–268.

- Sachsida, A. and Caetano, M. A.-R. (2000). The Feldstein-Horioka puzzle revisited, *Economics Letters* **68**: 85–88.
- Schmidt, M. B. (2001). Savings and investment: Some international perspectives, *Southern Economic Journal* **68**: 446–456.
- Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley & Sons, New York, Chichester.
- Sheffrin, S. M. and Woo, W. T. (1990). Present value tests of an intertemporal model of the current account, *Journal of International Economics* **29**: 237–253.
- Sinha, T. and Sinha, D. (2004). The mother of all puzzles would not go away, *Economics Letters* **82**: 259–267.
- Sinn, S. (1992). Saving-investment correlations and capital mobility: On the evidence from annual data, *Economic Journal* **102**: 1162–1170.
- Stone, M. (1974). Cross-validation choice and assessment of statistical predictions, *Journal of the Royal Statistical Society, Series B* **36**: 111–147.
- Summers, L. H. (1988). Tax policy and international competitiveness, in J. Frankel (ed.), *International Aspects of Fiscal Policies*, Chicago: University of Chicago Press, pp. 349–386.
- Swamy, P. (1970). Efficient inference in a random coefficient regression model, *Econometrica* **38**: 311–323.
- Swamy, P. and Mehta, J. (1975). Bayesian and non-bayesian analysis of switching regressions and of random coefficient regression models, *Journal of the American Statistical Association* **70**: 593–602.
- Taylor, A. M. (1996). International capital mobility in history: The saving-investment relationship, *NBER Working Paper* 5743.

- Taylor, A. M. (1998). Argentina and the world capital market: Saving, investment, and international capital mobility in the twentieth century, *Journal of Development Economics* **57**: 147–184.
- Taylor, A. M. (2002). A century of current account dynamics, *Journal of International Money and Finance* **21**: 725–748.
- Tesar, L. L. (1991). Savings, investment and international capital flows, *Journal of International Economics* **31**: 55–78.
- Tesar, L. L. (1993). International risk-sharing and non-traded goods, *Journal of International Economics* **35**: 69–89.
- Vita, G. D. and Abbott, A. (2002). Are saving and investment cointegrated? an ardl bounds testing approach, *Economics Letters* **77**: 293–299.
- Watson, G. (1964). Smooth regression analysis, *Sankhyā, Series A*, **26**: 359–372.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, **48**: 817–838.
- White, H. (2001). *Asymptotic theory for econometricians*, revised edn, Academic Press, San Diego.
- Wong, D. Y. (1990). What do saving-investment relationships tell us about capital mobility?, *Journal of International Money & Finance* **9**: 60–74.
- Wu, C. (1986). Jackknife, bootstrap, and other resampling methods in regression analysis (with discussion), *The Annals of Statistics* **14**: 1261–1343.

Index of Notation

Abbreviations related to models

| | |
|--|---|
| $B(\cdot)$ | standard Brownian motion |
| $B_{\underline{c}-c_0}^{\bar{c}-c_0}(\cdot)$ | regulated Brownian motion |
| $[\underline{b}, \bar{b}]$ | nominal bounds of a series |
| $[\underline{c} - c_0, \bar{c} - c_0]$ | standardized bounds of the regulated Brownian motion |
| $Cov(\cdot, \cdot)$ | covariance function |
| cv | absolute forecast errors |
| \tilde{cv} | normalized absolute forecast errors such that cv from the cross-section specific model is set to unity |
| cv^2 | squared forecast errors |
| \tilde{cv}^2 | normalized squared forecast errors such that cv^2 from the cross-section specific model is set to unity |
| $d_p(\cdot, \cdot)$ | mallows metric of order p |
| $E(\cdot)$ | expectation function |
| H_0 | null hypothesis |
| H_1 | alternative hypothesis |
| h | bandwidth parameter |
| I_{it} | ratio of gross domestic investment to GDP in period t and country i |
| \bar{I}_i | average of gross domestic investment to GDP for country i |
| $\mathbb{I}(\cdot)$ | indicator function |
| $K(\cdot)$ | kernel function |
| $\max(\cdot)$ | maximum function |

| | |
|----------------------|---|
| $\min(\cdot)$ | minimum function |
| $P(\cdot)$ | probability function |
| p -value | probability value associated with test statistics |
| S_{it} | ratio of gross domestic saving to GDP in period t and country i |
| \bar{S}_i | average of gross domestic saving to GDP for country i |
| $\sin(\cdot)$ | sine function |
| sup | supremum or least upper bound |
| Δ | first difference operator |
| \int | integral |
| $ \cdot $ | absolute value |
| \rightarrow | converge to |
| $\xrightarrow{a.s.}$ | converge almost surely to |
| \xrightarrow{p} | converge in probability to |
| \xrightarrow{d} | converge in distribution to |
| BI(1) | bounded integrated of order 1 |
| CF&Y | Cai, Fan and Yao |
| CV | cross-validation |
| DF | Dickey-Fuller |
| ECM | error correction model |
| FB | factor based bootstrap |
| FCM | functional coefficient model |
| FH | Feldstein-Horioka |
| HP | Hodrick-Prescott |
| I(0) | integrated of order 0 |
| I(1) | integrated of order 1 |
| iid | independently and identically distributed |
| MC | Monte Carlo |
| OLS | ordinary least square |

| | |
|---------|---|
| PB | pairs bootstrap |
| RB | residual based bootstrap |
| RSS_0 | residual sum of square under the null hypothesis |
| RSS_1 | residual sum of square under the alternative hypothesis |
| T_N | goodness-of-fit test statistic |
| VECM | vector error correction model |
| WB | wild bootstrap |

Abbreviations of economic variables

| | |
|-------|--|
| AGE | age dependency ratio (%) |
| EXPT | ratio of exports of goods and services to GDP (%) |
| CA | current account |
| GDP | gross domestic product |
| GDPC | natural logarithm of GDP per capita |
| GVBB | ratio of government overall budget balance (including grants) to GDP (%) |
| GVCE | ratio of government current expenditure to GDP (%) |
| GVCON | ratio of government consumption expenditure to GDP (%) |
| GVIVM | ratio of government capital expenditure to GDP (%) |
| GVTT | ratio of government total expenditure to GDP (%) |
| IMPT | ratio of imports of goods and services to GDP (%) |
| INTD | real interest rate differential(%) |
| LGDP | natural logarithm of GDP |
| OPN | ratio of export plus import to GDP (%) |
| POPG | growth rate of the population (%) |
| SI | saving-investment |

Abbreviations of cross-sections

| | |
|-----|--|
| E11 | 11 Euro area economies |
| E14 | 14 major countries of the European Union |
| F16 | 16 OECD countries considered by FH (1980) |
| D29 | 29 developed countries |
| L68 | 68 less developed economies |
| O15 | 15 OECD economies which are not in Euro area |
| O26 | 26 OECD countries |
| W97 | 97 countries from all over the world; |

Abbreviations of countries

| | |
|-----|--------------------------|
| DZA | Algeria |
| ARG | Argentina |
| AUS | Australia |
| AUT | Austria |
| BGD | Bangladesh |
| BRB | Barbados |
| BEL | Belgium |
| BEN | Benin |
| BWA | Botswana |
| BRA | Brazil |
| BFA | Burkina Faso |
| BDI | Burundi |
| CMR | Cameroon |
| CAN | Canada |
| CAF | Central African Republic |
| CHL | Chile |
| CHN | China |
| COL | Colombia |
| ZAR | Congo, Dem. Rep. |
| COG | Congo, Rep. |
| CRI | Costa Rica |
| CIV | Ivory Coast |
| DNK | Denmark |
| DOM | Dominican Republic |
| ECU | Ecuador |
| EGY | Egypt, Arab Rep. |

| | |
|-----|------------------|
| SLV | El Salvador |
| FJI | Fiji |
| FIN | Finland |
| FRA | France |
| GAB | Gabon |
| GMB | Gambia, The |
| GER | Germany |
| GHA | Ghana |
| GRC | Greece |
| GTM | Guatemala |
| GUY | Guyana |
| HTI | Haiti |
| HND | Honduras |
| HKG | Hong Kong, China |
| HUN | Hungary |
| ISL | Iceland |
| IND | India |
| IDN | Indonesia |
| IRL | Ireland |
| ISR | Israel |
| ITA | Italy |
| JAM | Jamaica |
| JPN | Japan |
| KEN | Kenya |
| KOR | Korea, Rep. |
| KWT | Kuwait |
| LUX | Luxembourg |
| MDG | Madagascar |
| MWI | Malawi |

| | |
|-----|--------------|
| MYS | Malaysia |
| MLI | Mali |
| MLT | Malta |
| MRT | Mauritania |
| MEX | Mexico |
| MAR | Morocco |
| MMR | Myanmar |
| NPL | Nepal |
| NLD | Netherlands |
| NZL | New Zealand |
| NER | Niger |
| NGA | Nigeria |
| NOR | Norway |
| PAK | Pakistan |
| PRY | Paraguay |
| PER | Peru |
| PHL | Philippines |
| PRT | Portugal |
| RWA | Rwanda |
| SAU | Saudi Arabia |
| SEN | Senegal |
| SGP | Singapore |
| ZAF | South Africa |
| SPA | Spain |
| LKA | Sri Lanka |
| SUR | Suriname |
| SWZ | Swaziland |
| SWE | Sweden |
| SWI | Switzerland |

SYR Syrian Arab Republic
THA Thailand
TGO Togo
TTO Trinidad and Tobago
TUN Tunisia
TUR Turkey
UGA Uganda
UK United Kingdom
US United States
URY Uruguay
VEN Venezuela, RB
ZMB Zambia
ZWE Zimbabwe

Curriculum Vitae

Personal Data

Name: Xu
Surname: Fang
Date of Birth: 10.11.1977
Place of Birth: Jiang Xi, China
email: fangxu[at]stat-econ.uni-kiel.de

Education

09/2003 - 04/2008 International Doctoral Programme in Quantitative Economics, Christian-Albrechts-University Kiel, Germany.
03/2007 - 12/2007 Advanced Studies Program in International Economic Policy Research, Kiel Institute for the World Economy, Germany.
09/2000 - 07/2003 Master of Arts in Economics and Management Science, Humboldt-University Berlin, Germany.
04/2000 - 07/2000 Department of Economics, Ludwig-Maximilians-University Munich, Germany.
09/1995 - 06/1999 Bachelor of Economics, Field "Money and Banking", Nankai University, China.

Academically Work Experience

09/2003 - 04/2008 Research and Teaching Assistant, Institute for Statistics and Econometrics, Christian-Albrechts-University Kiel, Germany.
07/2001 - 07/2003 Student Research Assistant, Institute for Statistics and Econometrics, Humboldt-University Berlin, Germany.

Teaching Experience

Econometrics I, Christian-Albrechts-University Kiel, Germany, M.A. Level, Winter 04/05 - 06/07.
Econometrics II, Christian-Albrechts-University Kiel, Germany, M.A. Level, Summer 2005 and 2007.
Introduction in Econometrics, Christian-Albrechts-University Kiel, Germany, M.A. Level, Summer 2004.
Econometric Software, Christian-Albrechts-University Kiel, Germany, M.A. Level, Summer 2004 - 2006 and Winter 04/05 - 06/07.

Publications in Refereed Journals

- 2006 "Reviewing the sustainability/stationarity of current account imbalances with tests for bounded integration", with H. Herwartz, forthcoming in *The Manchester School*.
- 2006 "Panel data model comparison for empirical saving-investment relations", with H. Herwartz, forthcoming in *Applied Economic Letters*.
- 2005 "Does consumption-wealth ratio signal stock returns? - VECM results for Germany", *Economics Bulletin* 3(30), 1-13.

Working Papers

- 2007 "A functional coefficient model view of the Feldstein-Horioka puzzle", with H. Herwartz, Kiel Economics Working Paper 2007-14.
- 2007 "A new approach to bootstrap inference in functional coefficient models", with H. Herwartz, Kiel Economics Working Paper 2007-15.

Professional Interests

Time series analysis, panel data analysis, and international macroeconomics.

Presentations

- 03/2008 Finance-Seminar, Limburg Institute of Financial Economics, Maastricht University, Maastricht.
- 09/2007 Statistical Week, Kiel.
- 09/2007 62nd European Meeting of the Econometric Society, Budapest.
- 08/2007 Cemapre Conference on Advances in Semiparametric Methods and Applications, Lisbon.
- 06/2006 International Conference on Policy Modeling, Hong Kong.

Refereeing

Referee for *Applied Economics*, *Applied Economic Letters*, and *Economics Bulletin*.

Selected Awards

- 2002 Dean's Honor List, Humboldt-University Berlin, Germany.
- 1997 Prize of Chinese friendly peace and development of the Panasonic Foundation, China.
- 1996 & 1997 Nankai University Scholarship, China.

Professional Affiliations

Econometric Society, European Economic Association,
Verein für Socialpolitik and Deutsche Statistische Gesellschaft.

Kiel, April 2008

Ich erkläre hiermit an Eides Statt, dass ich meine Doktorarbeit "A functional coefficient model view of the saving-investment relation" selbständig und ohne fremde Hilfe angefertigt habe und dass ich alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehnenden Ausführungen meiner Arbeit, besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert habe.

Kiel, April 2008