Legal Information

Publisher	The Bonn Journal of Economics Adenauerallee 24-42 53113 Bonn http://bje.uni-bonn.de bje@uni-bonn.de
V. i. S. d. P.	Jprof. Dr. T. Berg University of Bonn
Place of Publication	Bonn, Germany
ISSN	2195-3449

The Bonn Journal of Economics

Founded in 2012

VOLUME 4

ISSUE 2

Founded by

Johannes Hermle Justus Inhoffen Tobias Ruof

Department of Economics
University of Bonn

The Bonn
Journal of Economics

Recovering under a Fixed Exchange Rate - the Case of Estonia

Sonja Dobkowitz *

Introduction

Estonia stands out in the way it experienced the global financial crisis of 2008 in several ways. First, the downturn was extraordinarily severe as the financial crisis coincided with the bust of an overheating economy (Purfield and Rosenberg, 2010, p. 7). Second, the recovery took place relatively fast: in 2011 the economy was the fastest growing in Europe (Eurostat)[1]. Third, the policies adopted were rather unconventional: instead of devaluating its currency or stimulating the economy by expansionary fiscal policies, the country kept its exchange rate fixed to the euro and consolidated its budget.

Fiscal consolidation was primarily intended to stimulate foreign investment, restore confidence, and to fulfill EU accession criteria (Purfield and Rosenberg, 2010, p. 12). However, in theory, it also impacts the real economy and relative prices as it reduces aggregate demand. Fiscal contraction is therefore part of internal devaluation. The latter sums up policies directed at increasing compet-

*Sonja Dobkowitz received her degree (B.Sc.) from the University of Bonn in 2014. The present article refers to her bachelor thesis under supervision of Prof. Dr. Keith Kuester, which was submitted in August 2014.

[1] A detailed overview of statistics used is to be found in appendix A.

itiveness under a fixed exchange rate (Krugman, 2012). This summary is dedicated to analyze the mechanism of a general reduction in government spending. Did it, although seeming strange at first glance, contribute to the fast recovery in Estonia?

Being part of aggregate demand, fiscal spending has an effect on output and prices and, thus, the real interest rate. This is the reason why other components of aggregate demand are influenced. However, from a conventional Mundell-Fleming point of view, fiscal contraction appears to be a rather improper measure to boost recovery. Instead, it is presumed to crowd out other parts of aggregate demand, thereby worsening the downturn. On the contrary, when looking at fiscal contraction in a New Keynesian open economy environment the effect is different. Depending on the economy's characteristics, the multiplier is smaller than one. This is supported by an empirical study of OECD countries[2] conducted by Corsetti *et al.* (2012). The authors find a fall in short-term real rates and a real depreciation leading to a crowding in of consumption, investment, and net exports (pp. 21f).

Opinions differ concerning the impact of fiscal consolidation on recovery. Blanchard *et al.* (2013)[3] mention the coincidence of fiscal consolidation and recovery and merely point to fiscal consolidation having an effect on interest rates via the risk premium channel (p. 347). This assumption finds theoretical support as stated by Müller (2014) (p. 254); yet, this view neglects fiscal contraction's impact on prices as part of aggregate demand. Besides, Staehr (2013) underlines that especially the high degree of openness in Estonia makes it difficult to apply empirical findings of fiscal multipliers on the country (p. 30). Regarding openness, the European Commission Directorate-General for Economic and Financial

[2] Note that during the crisis Estonia was no member of the OECD but became one not later than June 2011. Therefore, the study seems suitable.

[3] The article focuses on fiscal contraction in Latvia. The country, however, conducted similar policies as Estonia.

Affairs (2010) estimates a multiplier for a model Baltic country and find a positive but low value (p. 67). However, this analysis disregards price flexibility as a determinant of the multiplier and the dynamic responses of the components of aggregate demand. This paper adds to the discussion by looking at the reasons for the size of the multiplier in the light of flexible prices and a high degree of openness, which is characteristic for Estonia.

To answer the research question a New Keynesian model of a small open economy will be applied to Estonia in 2009 when fiscal contraction began. The model developed by Corsetti *et al.* (2011) suits the country during a crisis as it provides for a currency peg and part of the households are excluded from the credit market.

The paper is structured as follows: the first section discusses a negative demand shock, as Estonia experienced it in 2008, in an open economy model. Afterwards, fiscal contraction under a fixed exchange rate will be analyzed in a New Keynesian model. Section 3 applies the model to Estonia in 2009. The last section concludes.

Recovering under Different Exchange Rate Regimes

The boom had been driven by private demand financed with foreign debt (Purfield and Rosenberg, 2010, p. 8). The following credit crunch and worsening of expectations (Staehr, 2013, p. 296) make it reasonable to model the crisis of 2008 in Estonia as a demand shock. The analysis in this section draws on a Mundell-Fleming and an AS-AD model (Blanchard and Illing, 2009, Chapters 7, 20, 21); both extended for an open economy.[4]

A flexible exchange rate depreciates when prices are too sticky to absorb the shock. The reason is that a decline in aggregate demand, ceteris paribus, translates into an excess supply of a country's currency. It is followed by an equal

[4] Appendix C1 deals with the models' equations.

reduction, a depreciation in the nominal exchange rate as per uncovered interest parity (UIP). This adds to the real depreciation which was caused by the initial fall in prices. A nominal depreciation makes the domestic currency cheaper stimulating aggregate demand for domestic goods.

Under a fixed exchange rate with prices somewhat sticky in the short run, there is no automatic stimulation of demand via the nominal exchange rate when a shock hits. The real exchange rate depreciates according to the initial change in prices only. Not until the driving factors of the AS curve, price expectations, wages, and prices, adjust to the new demand, output recovers. This shows that under a currency peg, wage and price flexibility gain in importance.

The Estonian authorities were reluctant to leave the peg. Nonetheless, the country recovered outstandingly fast (Eurostat).[5] Thus Estonia's recovery appears to be inconsistent with theory. Was the government right in adhering to 'the non-Keynesian effects of fiscal consolidation and its [...] expansionary effects'? (Raudla and Kattel, 2011, p. 15). To answer this question the following section will examine the fiscal consolidation methods adopted in Estonia, in particular cuts on the expenditure side, and their impact on recovery.

Fiscal Contraction under a Currency Peg

Estonia pursued an outstanding fiscal consolidation on both the expenditure and revenue side adding up to 8% of GDP in total (ECFIN, 2010, p. 67). Cuts in government expenditure made up the major part of fiscal consolidation (ECFIN, 2010, p. 67). Government expenditure fell by 1.58% of GDP in 2009.[6]

[5]Indeed, one can doubt the success of recovery. GDP growth rate and unemployment rate have not yet found back to their peak levels (August 2014) (Ibid.). Nonetheless, pre-crisis data needs to be examined against the background of an overheated economy with production over the natural level (Staehr, 2013, pp. 293 f) so that boom data should not be the benchmark.

[6] See appendix C4.

A New Keynesian Open Economy Model with a Fixed Exchange Rate

In general, the model takes the standard form of a New-Keynesian open economy model, for further explanation see Corsetti *et al.* (2011), henceforth CKM. The following equations give a first-order approximation of the equilibrium around a steady state (p. 8).[7] In the steady state, trade is balanced, debt is zero as is inflation, and purchasing power parity holds. The baseline model considers perfect financial markets and risk sharing; asset-market participation is unrestricted.

The dynamic IS (DIS) equation reads as follows

$$y_t = E_t y_{t+1} - \frac{(1-\chi)\bar{\omega}}{\gamma}(i_t - E_t \pi_{H,t+1}) - E_t \underbrace{(\hat{g}_{t+1} - \hat{g}_t)}_{\Delta \hat{g}_t + 1}, \tag{1}$$

and the New Keynesian Phillips curve (NKPC) is given by

$$\pi_{H,t} = \beta E_t \pi_{H,t+1} + \kappa(\varphi + \frac{\gamma}{(1-\chi)\bar{\omega}})y_t - \kappa \frac{\gamma}{(1-\chi)\bar{\omega}} \hat{g}_t \tag{2}$$

(CKM, p. 9). Incredibility of the peg is disregarded (CKM, p. 1), and monetary policy obeys

$$i_t = -\phi_\varepsilon e_t \tag{3}$$

(p. 9).[8]

[7] It holds that lower-case letters refer to a variable's percentage deviation from its steady-state value. The hat indicates that the respective deviations are expressed in percent of steady-state output. The meaning of the variables is to be found in appendix B.

[8] This version relies on the rule derived by Benigno *et al.* (2007). For reasons of consistency i denotes the deviations of the nominal interest rate from its steady state which differs from the model in CKM who use r instead.

Effects of Fiscal Contraction

This discussion is based on the model assuming complete financial markets. As shown by the DIS curve, a negative fiscal spending shock diminishes aggregate demand and reduces prices via the NKPC. The reason is that producers do not adjust production completely to the new amount demanded; they also lower prices in order to increase their sales. In contrast to that, a fall in government spending causes an increase in inflation. Assuming government spending to fall more than output, the former acts upon inflation as it leads to a depreciation of the terms of trade. Thus, foreign intermediate goods become relatively more expensive for home producers; their real marginal costs increase.[9] Overall, the negative impact from falling demand on inflation outreaches the contrary effect via marginal costs as can be seen by the coefficients of aggregate demand and government spending in the NKPC.

Keeping in mind an initial fall in inflation due to fiscal contraction, the short-term real interest rate rises one-to-one in accordance with the Fisher relation $i_t = \pi^e_{t+1} + r_t$ as nominal interest remains unchanged. However, essential to private demand is the long-run real rate as can be seen when solving the Euler equation forward (p. 14). Furthermore, the dynamic of the long-run real rate impacts the real exchange rate via UIP. According to the expectation hypothesis, the long-run interest rate is the sum of actual and future short-run ones; therefore, future inflation plays a crucial role in determining the fiscal multiplier.

CKM argue that, in the medium term, inflation increases thereby prompting future short-run real rates to fall. This is caused by expected inflation which is, under a peg, negatively related to its initial change. Regarding PPP as the driving factor of the price level under a fixed nominal exchange rate in the long run, the home price level needs to resemble its foreign equivalent. Per assumption, foreign

[9] See equation 9 in appendix C2.

variables are fixed, thus, the sum of home inflation needs to be zero. This starting point and the fact that nominal interest does not deviate result in the relation that the long-run real interest rate in the period of the shock equals inflation of the same period (CKM, p. 15).

The assumption of a constant nominal interest rate is essential to the finding that the initial change in the long-run real rate equals that of inflation. It needs to be considered against the background that there is no pressure exerted on the nominal exchange rate which needs to be corrected by monetary policy changing the policy rate. As the perfect credibility of the peg holds at every time (CKM, p. 1), the nominal interest rate does not move (Corsetti *et al.*, 2012, pp. 21f). Relaxing the assumption of the credibility of the peg, deviations of the nominal interest rate become indeed apparent. In their empiric study on fiscal multipliers, Corsetti *et al.* (2012) find a rise in the nominal interest rate as a reaction to government expansion (p. 22).

Having revealed the underlying mechanisms the effects of a reduction in government spending can be analyzed. The driving factor are price movements which determine the nominal exchange rate in the long run. When prices fall as a reaction to fiscal contraction, the value of the domestic currency rises; there is pressure on the nominal exchange rate to increase according to PPP. However, the monetary authority lowers the nominal interest rate which holds the exchange rate constant as per UIP.

Refocusing on the equilibrium model, an expected rise in inflation impacts aggregate demand positively: private consumption increases as the long-run real rate falls. Output rises according to the DIS relation. Eventually, domestic products become more competitive, yet, via the price channel and not for nominal devaluation.[10] Thus, the above analysis proves Estonian politicians right

[10] The movement in net exports will be analyzed in section 'Dynamics of the Baseline Simulation'.

in expecting fiscal contraction to have an expansionary effect and a multiplier smaller than one. However, hitherto, only a currency peg has been considered; the financial crisis has not yet been respected by the model.

Limited Asset-Market Participation

During the crisis, credit granting was restricted (Staehr, 2013, p. 296). This was essential to the Estonian economy as the demand-driven boom had been based on credit and foreign direct investment; the debt to GDP ratio of the private sector almost reached 100% of GDP (Purfield and Rosenberg, p. 6, Figure 2). Therefore, the global credit crunch hit Estonian private demand especially hard.

CKM amend their model to capture the effects of restricted participation in the credit market (p. 18). Therefore, a fraction λ of households is excluded; they do not own firms and consume their disposable income each period being called 'hand-to-mouth consumers'. Their fraction is set to $\lambda = \frac{1}{3}$ (p. 22). The remaining households own firms and hold assets (p. 18). Hence, the households differ in their budget constraint and in that non-asset holders only maximize period utility.

Another extension is the introduction of imperfect international risk sharing (p. 20). As this would hedge national consumption perfectly against deviations in that country's GDP. To examine effects of a shock more realistically, this assumption is given up. There are two channels through which risk sharing is possible: through diversified asset holdings and through commodity trade.[11] Consequently, international financial markets are incomplete: traded bonds are restricted to non-contingent bonds (CKM, p. 18), and the trade price elasticity is set to $\sigma = \frac{2}{3} < 1$, so that demand does not adjust perfectly to price changes.

[11] For further information see Deutsche Bank (2009) and Cole and Obstfeld (1991), respectively.

Results and Analysis

Dynamics of the Baseline Simulation

While fiscal contraction under a peg leads to a crowding in of private demand and a multiplier slightly smaller than one, the result is now reversed. When asset markets are incomplete, the trade price elasticity of demand is smaller one, and asset-market participation is restricted private consumption is crowded out and the multiplier is bigger than one (p. 22).[12] Owing to restricted households being unable to smooth consumption by borrowing when income falls, part of private consumption depends on period income and therefore government demand (p. 22). The long-term real rate does not influence the consumption decision of constrained households. The discussion below relates to figure 2, the 'baseline' simulation, which reflects the responses in the model conducted using CKM data. It shows that the path of consumption is negative until output has recovered equaling output of the unrestricted scenario. Then, consumption begins to deviate positively from its steady-state level. The reason is that when income has almost recovered the stimulating effect of the long-run real rate on unrestricted households' consumption exceeds the income effect on hand-to-mouth consumers and consumption is crowded in.

Comparing the real exchange rate and the long-run real interest rate for the complete and restricted environment, it becomes apparent that the initial fall in both variables is stronger in the restricted scenario. The dynamics are as follows: the decline in inflation is the result of producers adjusting their prices to the fall in aggregate demand. Demand falls more in the restricted scenario due to hand-to-mouth consumers. Remember that the initial change in long-run real rates equals the initial fall in inflation (p. 15). Consequently, the long-run real interest rate declines more than in the complete market surrounding. The real

[12] In the baseline simulation calibration follows Corsetti *et al.* (2011).

exchange rate depreciates as per UIP.

Finally, under restricted participation, an initial crowding in of net exports is predicted. Finding the driving factors requires to examine each component of net exports separately: $NX \equiv Exports - \frac{Imports}{q}$. Furthermore, exports and imports are defined as

$$Imports = [1 - (1 - \omega)\Big(\frac{P_{H,t}}{P_t}\Big)^{-\sigma}]C_t{}^{13}$$

$$Exports = \omega\Big(\frac{P^*_{H,t}}{P^*_t}\Big)^{-\sigma}C^*_t$$

(p. 8). They are fractions of foreign and domestic demand, respectively. There are four factors impacting net exports. Obviously, the quantity imported increases with first, domestic consumption and second, domestic producer prices so that both lead to a deterioration of net exports. Third, exports fall with an increase in producer prices thereby deteriorating net exports. Fourth, net exports diminish with a real depreciation as a depreciation makes imports more expensive; this will be labeled price effect below.

The elasticity of demand to the terms of trade is essential to the deviation of net exports. With it equaling one, net exports do not deviate from their steady state in an unrestricted environment; see figure 1. Net exports do not react to fiscal contraction and real depreciation. Apparently, perfect international risk sharing ensures that net exports are in equilibrium. The terms of trade need to balance the deviations in output, consumption and government spending as per $\hat{nx} = (1 - \chi)\omega s_t + (y_t - \hat{g}_t - \hat{c}_t)$ (p. 29).

Introducing $\sigma < 1$, quantities exported and imported do not react sufficiently to outweigh the price effect; overall, net exports worsen.[14] Falling producer

[13]Derivation see appendix C.3.
[14] See dashed line in figure 2.

prices change the shares imported and exported which leads to a crowding in of net exports. However, as the quantities do not react one-to-one with the change in prices the crowding in is not big enough to balance the price effect.[15]

With a share of households being excluded from asset-market participation, the value of the trade price elasticity appears to have lost importance. The initial response of net exports is the same for $\sigma = 1$ and $\sigma = \frac{2}{3}$[16]; net exports better. A fall in domestic private consumption also reduces the base of imports. In addition, assuming a higher fall in prices in the restricted scenario because of a sharper downturn in domestic demand, exports rise. In summary, net exports rise. Even more, the fall in domestic consumption and the rise in exports also dominate the stronger rise in prices of those goods which are still imported, although the real exchange rate depreciates more than under unrestricted markets. Consequently, there is an amelioration of net exports; however, it is not clear whether the fall in imports or the rise in exports is stronger. Note that the fall in imports due to falling domestic demand does not direct more demand to domestic goods. Conversely, the sharper fall in prices does.

Calibration Estonian Style

As Estonia had already experienced a decline in demand and output when fiscal contraction was conducted, the analysis presented here considers the economy at the beginning of 2009 to be the steady state. Thus, changes in the variables need to be understood as further deviations from the recessionary starting point of 2009.

Average government spending as a share of GDP is set to 20% in CKM. 2008 stands out in having a relatively high share of government spending as the budget

[15] Setting $\sigma > 1$, net exports better as a reaction to fiscal contraction since the price effect is outreached by the rise in demand for domestic products.

[16] Compare figure 1 and figure 2.

was planned for an expected GDP growth of 14% in that year (ECFIN, 2010, p. 66). Therefore, it is not suitable in presenting a value of government spending in normal times. Hence, average government expenditure will be set to the average from 2004 to 2007, which is 21.2%.

During the crisis, Estonia profited from a flexible wage setting environment. Dabušinskas and Rõõm (2011) find a significant higher wage flexibility in Estonia compared to other European countries when examining Wage Dynamics Network (WDN) data. To capture the effect of flexible prices the model is calibrated as follows: in the baseline model the slope of the Phillips curve κ is rather steep: a price duration of 9 months corresponds to $\xi = \frac{2}{3}$. This is calculated using the average price duration $\frac{1}{1-\xi}$, the expectation value of the geometric distribution. However, as mentioned in CKM, the characteristics of the model lead to a steeper curve for any given value of price stickiness (CKM, p. 10) as real rigidities are neglected.[17] Thus, this analysis is conducted setting $\xi = 0.7$, which shall account for both the outstanding flexible prices and the disregard of real rigidities.[18]

With prices being more flexible one finds a faster recovery under a peg; see figure 3. This finding is in line with the prediction of the AS-AD model of the section 'Recovering under Different Exchange Rate Regimes'. The real exchange rate depreciates more than twice as much compared to sticky prices. Same is true for the long-run real rate. Interestingly, net exports do not double. Conversely, their initial crowding in is even smaller than in the sticky price environment. Striking at first glance, this provides support for the analysis of net exports in the baseline model. As the real exchange rate falls more strongly on impact, imports become more expensive depressing the crowding-in effect of net exports. In addition, private consumption falls less so that the decline in imports is smaller. Private

[17] For more information on real rigidities see Romer (2008).

[18] Galí *et al.* (2001) set ξ to 0.9 given constant returns to scale as in CKM (table 1). However, this would not account for Estonia's flexible prices.

consumption and output are interlinked, and the starting point is determined by the long-run real interest rate. Consumption of asset-holding households rises more explicitly due to the greater fall in the long-run real rate. This leads to a smaller fall in output, thus, income. The restricted households' consumption weakens less. All in all, the multiplier under flexible prices is smaller than under more rigid prices, but still slightly bigger than one.

The assessment of flexible prices is ambiguous. On the one hand, domestic products become more competitive which can be seen looking at the prices of consumption: the long-run real rate and the exchange rate. Internal devaluation in Estonia was thus amplified and accelerated by flexible prices. On the other hand, in the model, the trade balance does not better as much owing to the intensified price effect of imports.

The Estonian economy was marked by a high degree of openness. So that this paper sets the parameter for openness ω, which represents the fraction of imports per consumption bundle (p. 5), to 0.6 (Worldbank), which is the value of the import-to-GDP ratio of Estonia in 2009.[19] It is significantly lower than the average of import-to-GDP ratio of 0.8% from 2004 to 2008, hence, during EU membership (Ibid.). The reason is that the crisis was a global one dampening worldwide production and trade. In the baseline model ω equals 0.3.

With a higher degree of openness, the trade channel works as a shock absorber; hence, the multiplier diminishes.[20] The reason is that a high share of demand is directed to foreign products. Thus, a fall in income of hand-to-mouth consumers impacts foreign production relatively more than home production compared to a less open economy. Therefore, the initial fall in private consumption downsizes although the long-run real rate falls less than in the baseline scenario. This

[19] Galí (2008) sets this value to 0.4 which corresponds to the import-to-GDP ratio in Canada (p. 174).

[20] This is in line with ECFIN (2010) (pp. 65ff) who find a smaller multiplier for the Baltic states.

fact underlines the initial domination of hand-to-mouth consumers on the overall effect. Not until income has recovered sufficiently, the smaller reduction in the long-run real rate becomes effective: private consumption takes longer to recover and its positive deviation afterwards is smaller as well. Since initial movements of the long-run real rate equal the initial drop in the consumer price inflation (CKM, p. 15), this inflation must fall less. Consumer price inflation is the sum of changes in domestic and foreign goods' prices weighted using the share of imports in GDP, in fact, openness: $\pi_t = (1-\omega)\pi_{H,t} + \omega p_{F,t}$.[21] In the model holds: the more open an economy the higher the share of imports in the consumption bundle (p. 5), the smaller the impact of domestic producer inflation on consumer price inflation, and thus on the long-run real interest rate.

The real exchange rate does not react as much as in an autarkic economy. First of all, this reduces the price effect on net exports. Furthermore, since in the steady state the share of foreign goods in the home consumption bundle is high, the reduction in domestic consumption influences net exports to rise more. As a consequence, the positive deviation in net exports is bigger than in the less open environment.

Summing up, internal devaluation is less effective in the more open economy as the real exchange rate depreciates less. Nevertheless, only a smaller depreciation is needed in order to lead to an improvement in net exports because of the stronger trade channel. Though openness diversifies the shock, the diversification is not perfect because of incomplete markets and a trade price elasticity of demand smaller than one.

Amended for both features, openness and price rigidity, it seems as if the degree of openness predetermined the initial responses of the variables; see figure 5. Flexible prices, on the contrary, determine the shape of the recovery process. This

[21] Derived from equation (A.3) in CKM (p. 26).

is because the degree of openness does not change during the recovery process due to the model's setting. Conversely, inflation is dynamic and moves qualitatively in the same manner for a given degree of openness. Eventually, both amendments bring the restricted economy closer to the unrestricted one.

The initial responses of the other variables are less devastating than in the baseline model: the multiplier equals one, net exports are crowded in more strongly, and private consumption deviates less negatively. Nevertheless, in the long run, the recovery is dampened. On closer inspection, real depreciation is bigger than in the baseline model owing to the flexibility of prices. However, the price effect on imports is intensified so that the crowding-in effect of the higher degree of openness is diminished. Net exports deviate by 20% from their steady state in the Estonian-like calibration. When consumption has recovered after five quarters, the deviation becomes negative. It is provoked by an intensified price effect in addition to the strengthened private consumption. The path of the long-run real rate is somewhat different to the scenarios discussed so far: it becomes positive after approximately four years. This is striking, as every factor taken on its own does not lead to a positive deviation. The finding accounts for openness absorbing the deflationary pressure on the long-run real rate in the beginning and flexible prices leading to a faster adjustment. Therefore, in the long term, consumption does not deviate positively; instead, it remains on its steady-state level after five quarters of recovery. Finally, output does not recover as fast as in the baseline scenario. The Estonian features account for an extended recovery period of slightly more than 25 quarters compared to 20 quarters in the baseline calibration[22].

[22] The simulation has also been conducted amending the size of the shock to 1.58%. This aims at exploring the effect of the real size of fiscal contraction in Estonia; see figure 6.

Possible Reasons for Estonian Recovery

The responses in foreign and domestic private demand were not big enough to establish a negative multiplier. A contraction in government spending as part of aggregate demand extended recovery. What was then the reason for the fast recovery in Estonia?

Rzońca and Cizkowicz (2005) point to a growth-enhancing effect of fiscal consolidation because demand rises due to an expected increase in cumulated disposable income (pp. 9f). The perspective of joining the euro area added to this effect. Furthermore, in the Estonian case, the rise in interest rate spreads at the beginning of the crisis (Purfield and Rosenberg, 2010, p. 9) caused a reduction in the multiplier (Müller, 2014, p. 252). Effectively, the recovery was export-driven, but not only because of fiscal contraction: the harsh collapse in exports might have led to a rebound effect (Staehr, 2013, pp. 300f). Moreover, the usage of EU structural funds during the crisis set countercyclical impetus (ECFIN, 2010, p. 58). Besides, the fiscal shock was not exogenous, as modeled by CKM, but triggered by the crisis. Yet, fiscal contraction might probably not have been expected, compared to fiscal expansion, which makes the model more suitable again. Rzońca and Cizkowicz (2005) underline that the effectiveness of fiscal consolidation depends on the spontaneity and the decisiveness of the policy (p. 5). Those facts support the idea of a non-Keynesian multiplier in Estonia.

Conclusion

Estonia recovered outstandingly fast after the global financial crisis of 2008 although it kept its exchange rate fixed to the euro. This paper examined the impact of fiscal contraction as an element of aggregate demand on that fast recovery.

Proponents of fiscal consolidation pointed to its expansionary effects with a multiplier smaller than one. It seems as if they did ignore the restricted credit granting during the crisis. Assuming that only a third of Estonians did not have access to credit, the multiplier increases above one. However, thanks to the flexible prices and the high degree of openness existing in Estonia, it converges closer to one again. Both features taken together reduced the multiplier and increased the crowding in of net exports and mitigated the crowding out of private consumption on impact. Simultaneously, those characteristics led to a prolonged recovery process. In the long run, there was even a crowding out of net exports, no crowding in of private demand, and a positive deviation in the long-run real interest rate. Thus, when looking at its transition as part of aggregate demand solely, fiscal contraction did not contribute to the country's fast recovery it even intensified the harsh downturn.

Fiscal devaluation might have been an alternative with less deteriorating effects. Applying such a policy to Estonia during the crisis might be an interesting subject for further research against the background of this paper's findings.

References

Benigno, G., Benigno, P. and Ghironi, F. (2007). Interest rate rules for fixed exchange rate regimes. *Journal of Economic Dynamics and Control*, **31** (7), 2196 – 2211.

Blanchard, O., Griffiths, M. and Gruss, B. (2013). Boom, bust, recovery: forensics of the latvia crisis. *Brookings Papers on Economic Activity*, pp. 325–388.

— and Illing, G. (2009). *Makroökonomie*, vol. 5. Pearson.

Cole, H. L. and Obstfeld, M. (1991). Commodity trade and international risk sharing. *Journal of Monetary Economics*, **28**, 3–24.

Corsetti, G., Kuester, K. and Müller, G. J. (2011). *Floats, pegs and the transmission of fiscal policy*. Working Papers 11-9, Federal Reserve Bank of Philadelphia.

—, Meier, A. and Müller, G. J. (2012). What determines government spending multipliers? *Economic Policy*, **27** (72), 521–565.

Dabušinskas, A. and Rõõm, T. (2011). *Survey evidence on wage and price setting in Estonia*. Working Paper 6/2011, Eesti Pank.

Deutsche Bank (2009). *Monthly Report*. Report, Deutsche Bank.

European Commission Directorate-General for Economic and Financial Affairs (2010). *Cross-country study economic policy challenges in the Baltics*. Occasional Papers 58, European Commission.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle*. Princton University Press.

Galí, J., Gertler, M. and Lopez-Salido, J. D. (2001). *European inflation dynamics*. Working Paper 8218, National Bureau of Economic Research.

Krugman, P. (2012). The unbearable slowness of internal devaluation. *The New York Times*.

Müller, G. J. (2014). Fiscal austerity and the multiplier in times of crisis. *German Economic Review*, **15** (2), 243 – 258.

Purfield, C. and Rosenberg, C. B. (2010). *Adjustment under a currency peg: Estonia, Latvia and Lithuania during the global financial crisis 2008-09*. Working Paper 10/213, IMF.

Raudla, R. and Kattel, R. (2011). Why did estonia choose fiscal retrenchment after the 2008 crisis? *Journal of Public Policy*, **31**, 163–186.

Romer, D. (2008). real rigidities. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.

Rzońca, A. and Cizkowicz, P. (2005). *Non-Keynesian effects of fiscal contraction in new member states.* Working Paper 519, European Central Bank.

Staehr, K. (2013). Austerity in the baltic states during the global financial crisis. *Intereconomics*, **48** (5), 293–302.

Appendix

A - Statistics Used

Eurostat

GDP and main components - volumes, *namq_gdp_k*;

Available from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=namq_gdp_k&lang=en.

Government revenue, expenditure and main aggregates - million euro, *gov_a_main*;

Available from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=gov_a_main&lang=en.

Gross domestic product at market prices - current prices in million euro, *tec00001*;

Retrieved from

http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do;jsessionid=9ea7d07d30e8fcb6835bcbdf4896a712f17b779c9a49.e34OaN8PchaTby0Lc3aNchuNa3uMe0?tab=table&plugin=1&pcode=tec00001&language=en.

Real GDP growth rate - volume, *tec00115*;

Retrieved from http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tec00115.

Unemployment rate by sex and age groups - monthly average, %, *une_rt_m*;

Retrieved from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=une_rt_m&lang=en.

Statistics Estonia

Average hourly gross wages (salaries) by year, economic activity (EMTAK 2008) and indicator, *WS5212*;
Retrieved from `http://pub.stat.ee/px-web.2001/Dialog/Saveshow.asp`.

Worldbank

Imports of goods and services (% of GDP);
Retrieved from `http://data.worldbank.org/indicator/NE.IMP.GNFS.ZS`.

B - Variables

It holds that lower-case letters refer to a variable's percentage deviation from its steady-state value. The hat indicates that the respective deviations are expressed in percent of steady-state output. An asterix refers to foreign variables, and the index to the period.

y	Output
χ	Average share of government spending
ω	Degree of openness
γ	Coefficient of relative risk aversion
i	Nominal interest rate
E	Expectation operator
ϕ_H	Producer price inflation
$\hat{g}$	Government spending
β	Discount factor
ξ	Degree of nominal price rigidity
φ	Inverse of the Frisch elasticity of labour supply

$\bar{\omega}$	$= 1 + \omega(2 - \omega)(\sigma\gamma - 1)$
e	Nominal exchange rate
ϕ_ε	Parameter of the interest rate rule under a peg
$\varepsilon_t = \frac{P_t^*}{P_{F,t}}$	Nominal exchange rate; not in deviations from its steady state;
p	Consumption price index
$p_{F,t}$	Price index of intermediate goods produced abroad
$p_{H,t}$	Price index of domestically produced intermediate goods
$s_t = p_{H,t} + e_t$	Terms of trade
q	Real exchange rate
$Q_t = \frac{P_t \varepsilon_t}{P_t^*}$	Real exchange rate; not in deviatons from its steady state
λ	Fraction of households excluded from the asset market
C_t	Final consumption good
σ	Trade price elasticity of demand
mc^r	Real marginal cost
w^r	Real wage

C - Derivations and Calculations

C.1 - Model of Section 2

The AS curve is given by

$$P = (1 + \mu)\underbrace{P^e F\big(1 - \frac{Y}{L}, z\big)}_{W} \tag{4}$$

(Blanchard and Illing, 2009, p. 622). The AD curve reads as

$$Y = Y\big(\underbrace{\frac{EP}{P^*}}_{-}, \underbrace{G}_{+}, \underbrace{T}_{-}\big) \tag{5}$$

(Ibid.), and the open economy IS-LM model is based on the following equations (Ibid., p. 598):

$$LM : \frac{M}{P} = L(\underbrace{Y}_{+}, \underbrace{i}_{-}), \tag{6}$$

$$IS : Y = C\underbrace{(Y-T)}_{+} + I(\underbrace{Y}_{+}, \underbrace{i}_{-}) + G + NX(\underbrace{Y}_{-}, \underbrace{Y^*}_{+}, \underbrace{\frac{1+i}{1+i^*}\bar{E}^e}_{\text{as per UIP; -}}). \tag{7}$$

The derivation and the meaning of the variables is to be found in chapter 21 (Ibid.). Note, however, that a decline in the nominal exchange rate E constitutes a depreciation, and that the expected nominal exchange rate is fixed per assumption.

C.2 - Government Spending and Real Marginal Cost

Deviation of real marginal cost from steady state is given by

$$mc_t^r = w_t^r - \omega s_t \tag{8}$$

(CKM, p. 27). Substituting the equilibrium equation $s_t = -\frac{\gamma}{(1-\chi)\bar{\omega}}(y_t - \hat{g}_t)$ (p. 31), one finds

$$\uparrow mc_t^r = w_t^r + \omega\frac{\gamma}{(1-\chi)\bar{\omega}}(y_t - \downarrow \hat{g}_t). \tag{9}$$

C.3 - Derivation of Imports

Aggregate demand for home products is given by

$$Y_t = (1-\omega)\left(\frac{P_{H,t}}{P_t}\right)^{-\sigma} C_t + \underbrace{\omega\left(\frac{P^*_{H,t}}{P^*_t}\right)^{-\sigma} C^*_t}_{\text{Exports}} + G_t \tag{10}$$

(CKM, p. 8). Consequently, the fraction of domestic consumption which is directed to domestic products is given by $(1-\omega)\left(\frac{P_{H,t}}{P_t}\right)^{-\sigma}$; the other part of domestic consumption is directed to foreign goods. Therefore, $[1-(1-\omega)\left(\frac{P_{H,t}}{P_t}\right)^{-\sigma}]C_t$ represents imports.

C.4 - General Government Spending as a Share of 2008 GDP

	2008	**2009**
GDP	16235.1	
Gross Capital Formation	892.7	727.4
Share of 2008 GDP	5.5%	4.48%
Final Consumption Expenditure	3128.2	3037.6
Share of 2008 GDP	19.3%	18.71%
Sum	4020.9	3765
Percentage Difference to 2008		-6.36%
Share of 2008 GDP	24.77%	23.19%
Change relative to GDP	$\frac{3765-4020.9}{16235.1} = -0.01576$	= **-1.58%**

All numbers other than percentages are provided in million euro.

Source: Eurostat

D - Figures

Results from Simulations

The solid lines depict the restricted scenario with incomplete asset markets and the dashed ones the unrestricted scenario with complete asset markets. The horizontal axes represent the course of time measured in quarters, the vertical axes the variables' deviations from their steady state, with the exception of deviations of government spending and output which are measured against steady-state output. For the figure showing five responses holds from top to bottom: the first panel shows deviations of private consumption, the second the long-run real interest rate, the third output, the fourth net exports, and the fifth the real exchange rate.

Figure 1: **Baseline Simulation** $\sigma = 1$

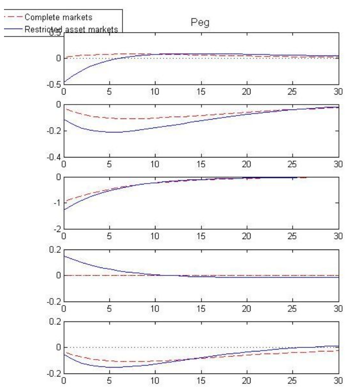


Figure 2: **Baseline Simulation** $\sigma = 2/3$

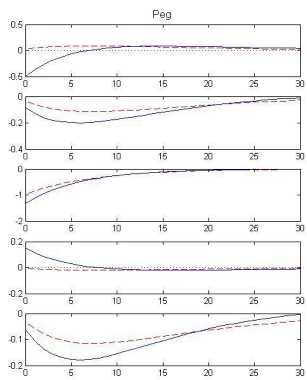


Figure 3: **Flexible Prices** $\xi = 0.7$

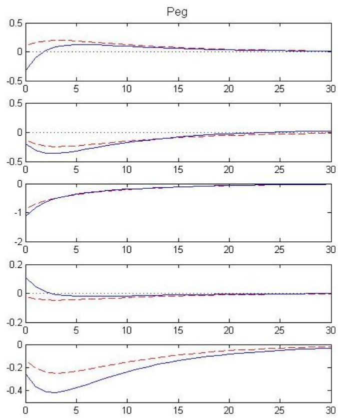


Figure 4: **Higher Degree of Openness** $\omega = 0.6$

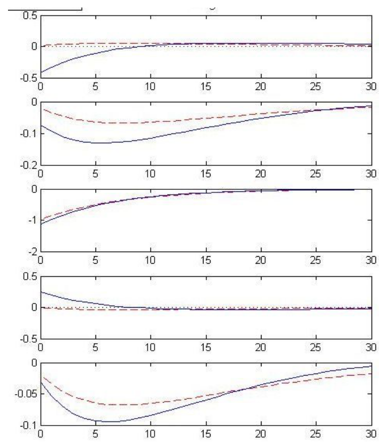


Figure 5: **Flexible Prices and Higher Degree of Openness**

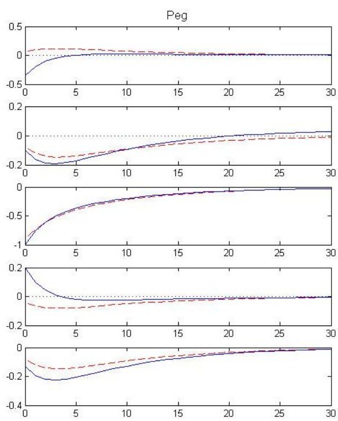
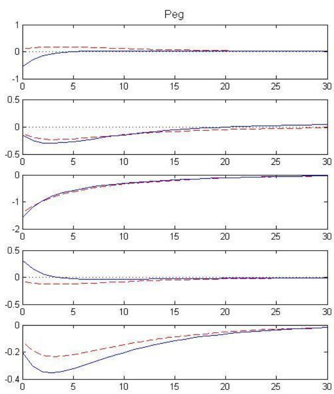


Figure 6: **Fiscal Contraction of 1.58%**



Causes of the Chilean Financial Crisis in the 1980's

Joshua Hruzik *

Introduction

The Chilean financial crisis of the 1980's is in many ways worthy of a concise scientific inquiry. Not only that there was socialist economic planning from 1970 to 1973, it was also replaced by unprecedented market reforms introduced by the Pinochet regime. It therefore raises the question what causes led to the meltdown of the Chilean financial system. Did market reforms fail to bring about everlasting change and prosperity into Chile?

The coup d'état by parts of the Chilean military, with the installation of the Pinochet regime in September 1973, acts as a historical structural break and it is therefore an appropriate point in time to start this inquiry. To date the beginning and the end of the financial crisis proves to be more challenging. With different definitions of a financial crisis come different time windows. It seems gainful to keep the temporal terminus of this paper in the blurry mid-1980's. To construct more robust time series it will be necessary to refer as early as to 1962.

The downfall of the Chilean financial system has already been subject of many economic and historic publications. A well-balanced presentation of the state of

*Joshua Hruzik received his degree (B.A.) from the Heinrich-Heine University in 2013. The present article refers to a term paper under supervision of PD Dr. Boris Gehlen which was submitted in February 2015.

research can be found in Peter Montiel's monograph *Ten Crises* (2014). Montiel emphasizes the impact of financial market liberalization, the opening up of the capital account, and an overvalued real exchange rate (Montiel, 2014, p. 42). This paper will not reject the current state of research. Rather, it will shift the focus of the discussion.

Methodically, this paper will use a cliometric and contra-factual model to explain the stifling level of foreign borrowing in the Chilean banking sector. A more theoretical examination of the effects which this foreign borrowing exercised will be made by using the Austrian business cycle theory.

Statistical modeling will be done based on the quarterly International Financial Statistics of the International Monetary Fund (IMF).[1] The World Bank's statistics on commerce will help to supplement the statistical structure by adding important price statistics.

The evolution of the Chilean financial system since 1973

Privatization of the Chilean banking system initiated in 1974 with the gradual denationalization of commercial banks, the dismantling of restrictions concerning the operational sphere of banks, and the relaxation of entry barriers for foreign banks (Parkin, 1983; Edwards, 1985; Magendzo and Titelman, 2008). In spite of the fact that this procedure was embedded in the general privatization strategy of the Pinochet regime, commercial banks were one of the very first institutions that were sold to the private sector (Edwards and Edwards, 1991, p. 96).

The denationalization of the banking system ended in a considerable concentration of banks in the hands of a few grupos.[2] In the year 1979, approx. 80% of the Chilean banking capital was concentrated in the balance sheets of these few

[1] Cliometric modeling on a monthly basis is not possible due to incomplete statistical records of the IMF.

[2] Grupos are internationalized conglomerates that operate in different sectors of the economy.

grupos (Edwards and Edwards, 1991, p. 99).

This process implied two fundamental problems: First, the amalgamation of creditors and debtors under the same grupo caused a fall in the diversification of debtors. Grupos financed their own businesses with loans granted by their own banks. Second, single banks assumed a vital role within the Chilean economy because of their dominant market share and their interlinkage with most businesses in Chile.

This process was not a significant problem for the recovery of the Chilean economy. Although being highly cartelized, the Chilean banking system was not able to establish a monopolistic price structure since there was enough (potential) international competition after opening the economy and relaxing the entry barriers for foreign banks (Edwards and Edwards, 1991, p. 100). Liberalizing the granting of loans, respectively the operative business of banks, was the second cornerstone of the market reforms inside the financial system. The act of liberalization was carried out by abolishing quantitative restrictions on the granting of loans, limiting the role of the central bank, freeing interest rates, progressively lowering the minimum reserve specification, and by abolishing the restrictions on the level of foreign borrowing (Magendzo and Titelman, 2008, p. 298).

The preceding clarifications concerning the structure of the Chilean banking system should have rendered evidently that the banking system was operating under a high degree of cartelization and a gradual liberalization of loan granting. This could have had significant implications for the vitality of the system itself. As long as these banks were operating within a framework of market competition and market prices, the personal liability for losses should have prevented the accumulation of toxic credits by giving incentives to proper risk assessment and debtor diversification. The freeing of interest rates on deposits and loans in the year 1976 introduced a financial system that was in principle based on free market prices.

However, in the year 1974, the Chilean government realized that the savings and loans institute SINAP (modeled after the American Savings-and-Loans) would fail due to an unhedged interest risk. In April 1974, the government reacted by announcing a government guarantee for the institute. Due to to the recession in 1974 and 1975, that was brought about by a windfall drop in the price of copper in 1974/Q3, the Chilean government lacked the financial resources to effectively save SINAP (Cuadra, Sergio de la and Valdés-Prieto, 1991, p. 75-76). After the government failed to put in practice its announced guarantees all further prevailing guarantees lost their credibility and therefore Chile entered a short-lived episode of free banking (Vittas, 1992, p. 13).

The end of the free banking experiment in Chile was brought about by January 6th, 1977. By that time the depositors and foreign creditors at Banco Osorno y la Union and numerous other smaller banks were bailed out while the concerned banks were recapitalized by the Chilean government (Brock, 2000; Cuadra, Sergio de la and Valdés-Prieto, 1990).

After Chile abandoned its short-lived free banking era, adequate banking regulations would have been necessary since government guarantees overrode the regulative mechanisms of the market economy. After Banco Osorno went bankrupt, the regulative framework was altered only marginally. Besides a strengthening of the financial and personnel resources of the banking supervision and the right to correct the book value of investments in the balance sheets of the banks, there have only been tougher restrictions concerning the capital base of financieras (Cuadra, Sergio de la and Valdés-Prieto, 1990, p. 43-44).

Chile's financial system was under a severe moral hazard that implied significant effects on the volume of loans granted and eliminated incentives to assess the risk taken by the banks. The regulative framework in act was appropriate for a banking system embedded in a market system that relied on self-regulation by personal liability while in reality the very basis for a system as such was nonexistent.

The prevailing moral hazard was further augmented by two policies: Shortly after saving Banco Osorno, the Chilean government relaxed regulations concerning the level of foreign indebtedness. This paved the legal way to excessive moral hazard lending by means of mediating foreign credit. The second policy, augmenting the prevailing moral hazard, was fixing the Peso to the dollar. The fixed exchange rate acted as a government guarantee when banks engaged in mediation of foreign credit (McKinnon, 1999, p. 25) .

International capital flows into Chile

Chile quit its short-lived free banking era by rescuing the failing Banco Osorno and other financial institutes while at the same time it missed the opportunity to adjust its regulative framework to the new realities. As a consequence, banks began excessive credit mediation whereby the foreign indebtedness of Chilean commercial banks reached a stifling level.[3] A question that remains, however, is why foreign creditors were willing to lend to Chilean banks. Government guarantees are only capable of explaining why Chilean banks engaged in moral hazard lending but not why other banks provided liquidity. The unique historical setting in Chile coincided with a high degree of international liquidity in the aftermath of the first oil price shock (Zahler, 1983, p. 512).

In order to test this hypothesis empirically, the deflated foreign liabilities of Chilean commercial banks up until the bailout of Banco Osorno will be investigated. Based on quarterly data since 1962, an Autoregressive-Integrated-Moving-Average (ARIMA) model will estimate a forecast for the subsequent time frame.

[3] See figure 1.

If the prevailing moral hazard and the opening up of the capital account did not exert any influence on the level of foreign liabilities, the forecast should not deviate noticeably away from the actual foreign liabilities which should also linger within the 95% confidence interval.

Since the data is evidently not stationary and contains an exponential trend, the data needs to be transformed by taking its logarithm and must be differentiated at least once.

Although the time series' first difference oscillates around a constant mean and shows constant variance, some outliers before and after the coup d'état[4], as well as the outcomes of the ADF-Test and the KPSS-Test, are worrisome. The fact that both tests are insignificant indicates the presence of a non-stationary process according to the ADF-Test (-2.8558) and the presence of a stationary process according to the KPSS-Test (0.19377). With recourse to the Lex Parsimoniae, a stationary process is assumed in order to avoid over-differencing the time series.

Dependent variable			
Name		Deflated foreign liabilities of Chilean commercial banks in billion Pesos	
Transformation		log	
Coefficients		**Model evaluation**	
AR(1) (s.e.)	0.4562 (0.1150)	AICc	-6.73
Drift (s.e.)	0.0315 (0.0195)	σ^2	.04728

Table 1: ARIMA model specifications

The ARIMA model (table 1) with the order (1,1,0) was estimated in order to give a forecast. It gives some evidence for the presence of moral hazard credit

[4]See figure 2.

mediation. Shortly after the bailout of Banco Osorno and the liberalization of the capital account, there has been an exponential increase in deflated foreign liabilities of Chilean commercial banks. In the long run, the factual time series passes over the 95% confidence interval.[5]

The model's residual diagnostics show no significant lags. Unfortunately, the residuals are not normally distributed and they seem to contain a unit root in the last quarters.[6] The residual diagnostics of models with different orders (not reported) support the finding that a (1,1,0) order is still the best order to choose. Non-normal errors might have serious implications for the confidence interval. An ex-post evaluation of alternative models proves that a (1,1,0) model gives the best point estimate. The point estimates of higher order models point towards the wrong direction and the confidence intervals are much wider.[7] Also, the model suffers from the characteristic limitations of a univariate time series model. The model could easily be criticized by pointing out that only one intervention (e.g. the relaxation of rules concerning the level of foreign indebtedness shortly after introducing the moral hazard) influenced the growing level of foreign indebtedness. It is for this reason and the technical shortcomings that the model stands or falls with its theoretical premises.

Economic implications of the capital inflows

Although, the capital inflows delivered considerable benefits, for without them a recovery of the economic catastrophe that was brought about by the Allende government would have never been possible, its excessive volume, triggered by a moral hazard paired with an inadequate regulatory framework, brought along fundamental economic distortions. These distortions can be studied by using the

[5]See figure 3.
[6]See figure 4.
[7]See figure 5.

Austrian Business Cycle Theory, represented by their most dominant exponents Ludwig von Mises and Friedrich August von Hayek.

In a nutshell, the Austrian Business Cycle Theory argues that a business cycle commonly starts with an expansion of credit, whereby the gross market rate lags behind the originary interest rate plus the positive price premium. This lag leads to a heightened demand for loans that will be used to expand the production facilities of businesses. The drop in the market rate of interest induces lower costs in the calculation for future investments, raising the profit expectations for new investments. When credit expansion ends and market rates raise towards the equilibrium rate of interest, production plans that seemed profitable before will prove themselves as unremunerative (Mises, 2007; Hayek, Friedrich A. von, 2012). A subsequent crisis is the correction of production plans.

Year	**Lending interest rate (%) in Chile**
1977	163.15
1978	86.13
1979	62.11
1980	47.14
1981	52.01
1982	63.86

Table 2: Interest rates on loans in Chile. Source: World Bank

An examination of interest rates (table 2) in Chile shows a common trend with the volume of credit in Chile[8]:

Perceived wealth and expectations in the future profitability of the existing production structure increased with an ever advancing volume of credit. Con-

[8] Please note that a finite distributed lag model can not be fit because interest rates on loans in Chile are only available on a yearly basis.

sumption expenditure increased and foreign savings substituted domestic savings without raising investments in capital goods above its persistent low level (Edwards and Edwards, 1991, p. 60). The gainers of this consumption spending boom were the construction sector, the commercial sector, and the financial sector (Condon, Corbo, and Melo, 1985, p. 381). Increased capital inflows directly affected the commercial and financial sector since commercial banks mediated the increased foreign indebtedness of the Chilean economy while a relative augmentation of imports (financed by foreign liabilities rather than equalizing exports) increased the number of import traders. Booming constructions in Chile might be seen as a typical symptom for cheap money. Falling market rates on loans made financing real estate much easier.

Chile in Crisis

Crisis appeared to be an inevitable corrective action the longer the expansion of credit endured. How did the sudden end of the boom come about?

After fixing the Peso to the Dollar in 1979, Chile imported the monetary policy of the United States. The real appreciation of the Dollar to the currencies of Chile's major trading partners worsened the competitiveness of Chilean businesses on the world market (Valdés-Prieto, 1994; Magendzo and Titelman, 2008). The real appreciation of the Dollar was also accompanied by a consumption of American savings due to the loose fiscal policy of the Reagan-administration. Increased capital demand in the United States raised interest rates on the American capital market which led to a reversal in capital flows. Capital inflows into the United states increased while at the same time capital outflows dried up (Aldcroft and Oliver, 1983; Eichengreen, 2000). The following increase of interest rates in Chile revealed the misallocation of resources within the Chilean economy. Businesses that formally appeared to be profitable were not able to finance themselves on the credit market and went into bankruptcy. 1982 by itself registered more than 800 bankruptcies in the private sector (Collier and Sater, 2004, p. 370). A second wave of bankruptcies happened in November 1981, when eight financial institutions were rescued by the government (Montiel, 2014, p. 32). The peak of the financial crisis was in January 1983, when the Chilean government bailed out several major banks and thereafter owned more than 50% of the banking system (Brock, 2000, p. 76).

The events of the Chilean financial crisis are very much in line with what the Austrian Business Cycle Theory says. Interest rates were depressed by excessive credit mediation and when credit mediation ceased it also ended the damping of the interest rate. Rising interest rates ended the Chilean credit boom and production plans had to be rearranged.

How could the Chilean banking sector become so vulnerable? The concentration process, that accompanied the process of denationalization, led to a close conjunction between private sector businesses and credit mediators. Loans were mainly granted to companies owned by the parent grupo (Harberger, 1985, p. 543). Little diversification of risk and practically non-existing risk assessment linked the survival of financial institutions directly to the survival of a small group of businesses. As soon as the huge number of toxic credits was uncovered, risk premiums surged and therefore no new inflows of capital could be lured into Chile.

Conclusion

It has been shown that the roots of the Chilean financial crisis lay much deeper than in exogenous shocks. The foundation for the downfall of the Chilean financial system was laid much earlier. Government guarantees combined with a simultaneous liberalization of the financial system proved to be a toxic mixture.

It led to a massive expansion of credit by means of foreign indebtedness. Enterprises seemed lucrative due to the artificially low market rate of interest, that was below the equilibrium rate of interest, and were kept alive by toxic loans granted by banks belonging to the same parent grupo. The economic crisis was an inevitable correction towards market prices.

The question of whether market reforms failed to bring about everlasting change and prosperity into Chile has to be negated. Facts prove quite the contrary. The Chilean financial crisis was an adjustment of wrong price signals that were introduced by government intervention.

There is still historical research to be done. Studies concerning the corporate history of large Chilean banks could prove themselves as highly prolific. Such studies could expand our knowledge on the decision-making horizon of Chilean bankers and thereby help us to properly analyze the implications of the moral hazard. Also, more data is needed to construct sophisticated models. Multivariate models might be able to overcome the problems of this paper's ARIMA model.

References

ALDCROFT, D. H., AND M. J. OLIVER (1983): *Exchange Rate Regimes in the Twentieth Century*. Edward Elgar Publishing, Cheltenham and Northampton.

BROCK, P. L. (2000): "Financial Safety Nets: Lessons from Chile," *The World Bank Research Observer*, 15(1), 69–84.

COLLIER, S., AND W. F. SATER (2004): *A history of Chile, 1808 - 2002*, vol. 82 of *Cambridge Latin American studies*. Cambridge University Press, Cambridge [USA], 2 edn.

CONDON, T., V. CORBO, AND J. D. MELO (1985): "Productivity Growth, External Schocks, and Capital Inflows in Chile: A General Equilibrium Analysis," *Journal of Policy Modeling*, 7(3), 379–405.

CUADRA, SERGIO DE LA, AND S. VALDÉS-PRIETO (1990): "Myths and Facts about Financial Liberalization in Chile: 1973-1982," .

——— (1991): "Bank Structure in Chile," in *Banking Structures in Major Countries*, ed. by G. G. Kaufman, pp. 59–112. Springer, Norwell.

EDWARDS, S. (1985): "Stabilization with Liberalization: An Evaluation of Ten Years of Chile's Experiment with Free Market Policies, 1973-1983," *Economic Development and Cultural Change*, 33(2), 223–254.

EDWARDS, S., AND A. C. EDWARDS (1991): *Monetarism and liberalization: The Chilean experiment*. University of Chicago Press, Chicago.

EICHENGREEN, B. (2000): *Vom Goldstandard zum Euro: Die Geschichte des internationalen Währungssystems*. Klaus Wagenbach Verlag, Berlin.

HARBERGER, A. C. (1985): "Observations on the Chilean Economy, 1973-1983," *Economic Development and Cultural Change*, 33(3), 451–462.

HAYEK, FRIEDRICH A. VON (2012): *Monetary Theory and the Trade Cycle*. Martino Fine Books, New York.

MAGENDZO, I., AND D. TITELMAN (2008): "Chile," in *Macroeconomic volatility, institutions and financial architectures*, ed. by J. M. Fanelli, pp. 283–315. Palgrave Macmillan, Basingstoke and New York.

MCKINNON, R. I. (1999): "Exchange-Rate Regimes for Emerging Markets: Moral Hazard and International Overborrowing," *Oxford Review of Economic Policy*, 15(3), 19–38.

MISES, L. V. (2007): *Human Action: A Treatise on Economics*. Liberty Fund, Indianapolis, 4 edn.

Montiel, P. (2014): *Ten crises*. Routledge Chapman & Hall, London and New York.

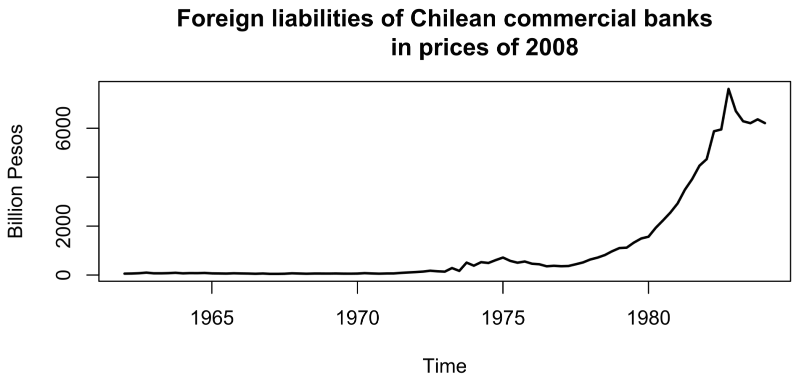
Parkin, V. (1983): "Economic liberalism in Chile, 1973-82: a model for growth and development or a recipe for stagnation and impoverishment?," *Cambridge Journal of Economics*, 7, 101–124.

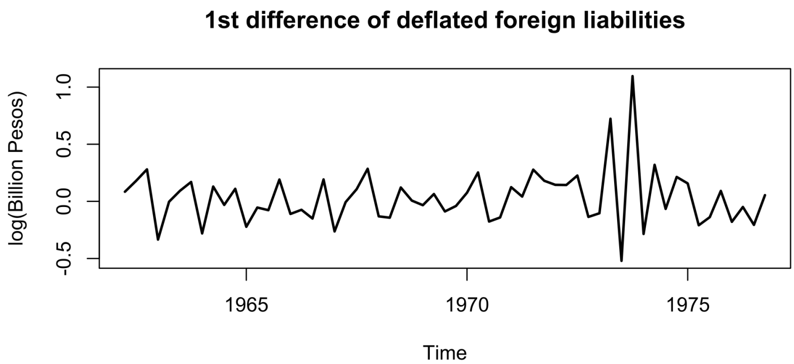
Valdés-Prieto, S. (1994): "Financial liberalization and the capital account: Chile, 1974-84," in *Financial reform*, ed. by G. Caprio, I. Atiyas, and J. A. Hanson, pp. 357–410. Cambridge University Press, Cambridge [UK] and New York.

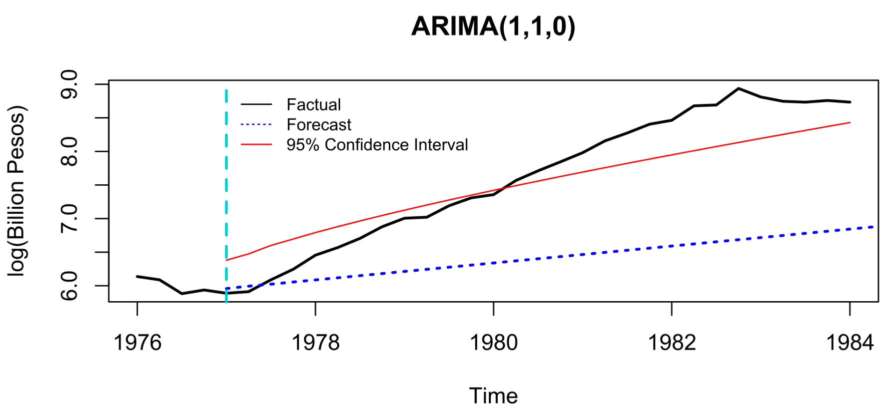
Vittas, D. (1992): "Policy Issues in Financial Regulation," .

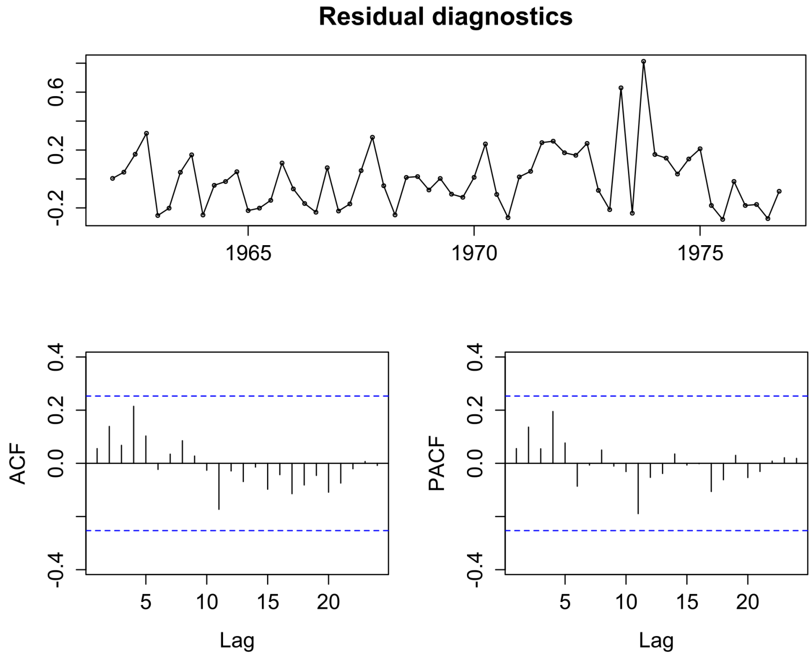
Zahler, R. (1983): "Recent Southern Cone Liberalization Reforms and Stabilization Policies: The Chilean Case, 1974-1982," *Journal of Interamerican Studies and World Affairs*, 25(4), 509–562.

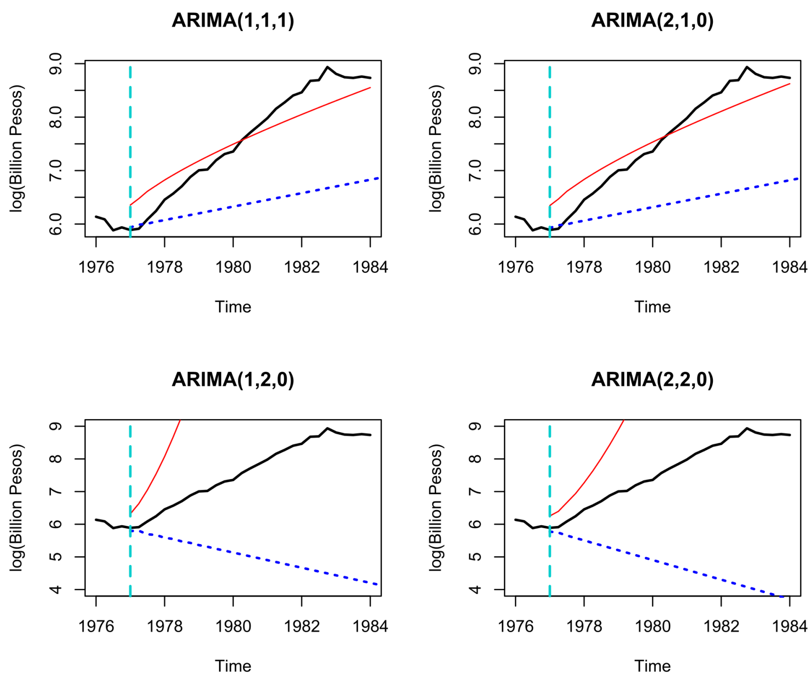
Appendix


Foreign liabilities of Chilean commercial banks
in prices of 2008
Billion Pesos
0
2000
6000
1965
1970
1975
1980
Time


1st difference of deflated foreign liabilities
log(Billion Pesos)
-0.5
0.0
0.5
1.0
1965
1970
1975
Time


ARIMA(1,1,0)
log(Billion Pesos)
6.0
7.0
8.0
9.0
Factual
Forecast
95% Confidence Interval
1976
1978
1980
1982
1984
Time


Residual diagnostics
-0.2
0.2
0.6
1965
1970
1975
ACF
-0.4
0.0
0.2
0.4
5
10
15
20
Lag
PACF
-0.4
0.0
0.2
0.4
5
10
15
20
Lag


ARIMA(1,1,1)
ARIMA(2,1,0)
log(Billion Pesos)
6.0
7.0
8.0
9.0
1976
1978
1980
1982
1984
Time
ARIMA(1,2,0)
ARIMA(2,2,0)
log(Billion Pesos)
4
5
6
7
8
9
1976
1978
1980
1982
1984
Time

The Bologna Reform in the Job Market Signaling Model

Jan Knöpfle *

Introduction

With the objective of increasing the competitiveness of European Universities and increasing the mobility as well as employability of their graduates, the members of the European Higher Education Area (EHEA)[1] agreed to participate in a voluntary harmonisation process, culminating in the Bologna Declaration signed in 1999.

As a result of this declaration, many reforms were instated in order to create an internationally comparable and straightforward system of degrees, fostering the European dimension in higher education (e.g. through the Erasmus programme) and to establish a structure of mainly two degrees.

The latter objective is what the reform is mostly identified with in Germany, and which led to the most profound changes of the country's higher educational system.

From 2005 to 2010, the agreement establishes, the structure with a first job-qualifying degree after an overall duration of mostly five years, the Diploma

*Jan Knöpfle received his degree (B.Sc.) from the University of Bonn in 2014. The present article refers to his bachelor thesis under supervision of Prof. Dr. Dezsö Szalay, which was submitted in August 2014.

[1]The EHEA encompasses all members of the European Union and further countries. (See www.ehea.info)

should be transformed into one with a first qualifying degree (bachelor) after three years and a second (master) after two additional years. The doctoral degree or PhD remained to be the subsequent level of qualification offered and its structure was altered only punctually.

The change from the diploma to the bachelor's and master's degrees will be at the heart of this study. It does not aim to evaluate the performance with respect to its qualitative ambitions mentioned above, but rather to analyse the impact on students' utility, determined by effort for education and resulting wages.

The next section gives a brief overview of recent studies concerning the Job Market Signaling Model and the competing model of human capital accumulation; section three and four introduce the model and equilibrium concept, respectively which are then used for the welfare analysis. Concluding remarks follow in the last section.

Literature overview

Two theories prevail the analysis of the effect, education has on employees' wages: There is the human capital theory according to Becker (1962) which focuses on the increased productivity induced by education and the Job Market Signaling Model (Spence, 1973), which will be the groundwork for this thesis. It argues that education works as a signal in a job market of asymmetric information. Its sole purpose in this model is to convey a higher, yet unobservable, productivity to potential employers since this lowers the cost of obtaining a degree.

Empirical studies on which of the effects dominates vary in their methodology as well as findings:

Some derive their conclusions by evaluating reforms that have a direct effect only on certain groups of students, and analysing if there are indirect effects on other groups. Observing a positive correlation between the access to higher education and high school drop-outs in US data (Bedard, 2001) can be seen as argument for the Signaling Theory with the following rationale: Increasing the access to university lowers the average productivity and therefore the prospective wages of students with "only" a high school degree. This makes more pupils drop out. However, similar evaluation on the minimum years of schooling in the UK and the absence of effects on the educational choices of those in higher tiers let Chevalier, Harmon, Walker, and Zhu (2004) conclude that the human capital effect dominates.

Diploma effects, a high wage increase attributed to the last educational year before obtaining a degree and very low gains for the years before, observed in US data are explained best within the signalling framework (Frazis, 2002).

Observations for Germany include Rinne and Zhao (2010), using the reunification as a natural experiment and observing that the lower wage effects of East-German degrees in comparison to their West-German equivalents are an indicator for the human capital theory. Whereas Dill and Hammen (2011) use the Bologna Reform and the transition time, during which both the Bachelor- and Master as well as the Diploma paths could be chosen to observe clear signalling effects in the different wages both groups of graduates received.

Model

The effects of the reform towards Bachelor and Master degrees are analysed in the signalling game of the following structure:

1. Nature determines the productivity (type) of an individual θ, uniformly distributed on the interval $[0, 1]$.

2. The individual chooses her level of effort $e \in \mathbb{R}$, incurring cost $c(e, \theta) = \frac{e}{2\theta}$ where

before the reform: $e \in \{0, d, p\} = \{NoSignal, Diploma, Doctorate\}$

after the reform: $e \in \{0, b, m, p\} = \{NoSignal, Bachelor, Master, Doctorate\}$.

3. Potential employers observe the level of education, not the productivity, and offer a wage $w(e) \in \mathbb{R}_+$.

The model assumes that the productivity level of an individual is innate and cannot be altered through education. In a job market with asymmetric information regarding the workers' productivity levels, effort however, can be used as a signal to convey information about the type if the cost of education is negatively correlated with the productivity level.

Costs should be thought of as including all monetary aspects as tuition fees and materials, but specially non-monetary dimensions such as foregone leisure and the strain involved in obtaining a certain degree. The latter gives rise to the negative correlation of cost and productivity.

The wage can be interpreted as value of a discounted profile of future payments. Assuming perfect competition among risk-neutral employers, they simultaneously offer wages to the applicant, who then chooses to accept the highest; if any. This implies that firms in equilibrium will offer a wage at the expected productivity level, given the observed choice of effort: $w(e) = \mathbb{E}[\theta|e]$.

This set-up differs from the canonical Job Market Signaling model due to Spence (1973) in the continuous, instead of binary type space and the discrete choice set of effort. The cost function was altered to $c(e, \theta) = \frac{e}{2\theta}$ instead of $\frac{e}{\theta}$ in the original model for ease of exposition. Results also hold without this change.

In what follows the productivity levels which lead to indifference between two levels of effort will play a central role in the analysis. We denote by $\theta_{b,m}$ the type which induces indifference between obtaining a bachelor's or a master's degree and correspondingly for all pairs of successive choices of e.

Equilibrium Analysis

The Equilibrium concept used in the following will be that of Perfect Bayesian Equilibrium (PBE) which requires the following conditions in the present model of Osborne and Rubinstein (1994):

- **Sequential Rationality**: Firms choose to offer $w(e) = \mathbb{E}[\theta|e]$ and individuals choose an effort level $e^* \in \arg\max_e w(e) - c(e, \theta)$.
- **Correct initial beliefs** about the distribution of θ. Implies $\mathbb{E}[\theta] = \frac{1}{2}$.
- **Action-determined beliefs**: The firm's belief about the type is a function of the effort only.
- **Baeysian updating**: Firm builds beliefs with all available information through Bayes Rule.

Note that the above characterization does not restrict the off-equilibrium beliefs in any sense. The concept of PBE doesn't lead to a unique prediction: A pooling equilibrium where all agents choose the same level of education and get a wage of $\frac{1}{2}$ as well as some "hybrids" where only certain degrees aren't chosen could be supported by the off equilibriums belief(s) $\mathbb{E}[\theta|e'] = 0$ for all e' not chosen in equilibrium, since this belief renders deviations to such an effort level unprofitable.

To further restrict the set of equilibria we will make use of the intuitive criterion presented by Cho and Kreps (1987). This refinement discards equilibrium-dominated beliefs off the equilibrium path of a game. In the above example, deviating from the pooling equilibrium with zero effort towards a higher signal is equilibrium dominated for the lowest types, in the sense of higher wages not compensating for the high cost that this type faces. The intuitive criterion requires that the firms adapt their beliefs accordingly; therefore - when observing a deviation towards a higher effort - assume it comes from a more productive individual. Hence, only the following separating equilibrium in which all effort levels are chosen will survive the refinement concept:

After observing a signal e firms pay wages in the middle of the interval from the two adjacent indifference levels. For example a diploma graduate will have a wage of

$$w(d) = \mathbf{E}[\theta|d] = \frac{1}{2}(\theta_{0,d} + \theta_{d,p})$$

Potential employees anticipate this belief formation and choose among the discrete effort set, the level which maximizes their utility. This implies that for all types $\theta \leq \theta_{0,d}$ it is better not to choose a costly signal in the sense that

$$w(0) - c(0,\theta) = \frac{0 + \theta_{0,d}}{2} - \frac{0}{2\theta} \quad \geq \quad \frac{\theta_{0,d} + \theta_{d,p}}{2} - \frac{d}{2\theta} = w(d) - c(d,\theta) \quad (1)$$

with equality holding at $\theta = \theta_{0,d}$. Similarly we have the following inequality for all $\theta \in [\theta_{0,p}, \theta_{d,p}]$ with equality at $\theta = \theta_{d,p}$

$$w(d) - c(d,\theta) = \frac{\theta_{0,d} + \theta_{d,p}}{2} - \frac{d}{2\theta} \quad \geq \quad \frac{\theta_{d,p} + 1}{2} - \frac{p}{2\theta} = w(p) - c(p,\theta) \quad (2)$$

Combining 1 and 2 and applying the same procedure to the effort levels to be chosen from after the reform, leads to the following indifference levels and wages:

Before Reform	**After Reform**
Indifference levels:	
	$\theta_{0,b} = \frac{bp}{m}$
$\theta_{0,d} = \frac{d}{p}$	$\theta_{b,m} = \frac{m}{p}$
$\theta_{d,p} = p$	$\theta_{m,p} = p$
Wages:	
$w(0) = \frac{d}{2p}$	$w(0) = \frac{bp}{2m}$
	$w(b) = \frac{1}{2}(\frac{bp}{m} + \frac{m}{p})$
$w(d) = \frac{1}{2}(\frac{d}{p} + p)$	$w(m) = \frac{1}{2}(\frac{m}{p} + p)$
$w(p) = \frac{1+p}{2}$	$w(p) = \frac{1+p}{2}$

For the above values to be economically relevant we require that $0 < \theta_{0,d} < \theta_{d,p} < 1$ and $0 < \theta_{0,b} < \theta_{b,m} < \theta_{m,p} < 1$ hold. This is satisfied if $d < p^2$ as well as $b > (\frac{m}{p})^2$ and $m < p^2$.

Since the duration from the beginning of higher education and finishing a master's degree is equal to the time that was necessary to finish a diploma under the old regime, and the number of graduates in both categories has been approximately the same, it is assumed for the rest of the analysis that $m = d$. The results presented below hold without this simplification in a broad range of plausible parameter values for m and d.

Welfare evaluation

From the perspective of a utilitarian social planner who puts the same weight on every individual, evaluating the welfare implications of the Bologna Reform within the above model comes down to comparing the total sums of costs and wages resulting from the equilibrium effort levels.

Note that due to the fact that each groups' individuals get exactly the respective average productivity level paid as wage, the sum will always be $\frac{1}{2}$; independent of the distribution of groups. Therefore, total welfare increases after the reform iff the sum of costs before the reform exceeds that afterwards:

$$\int_0^{\theta_{0,d}} c(0,z)dz + \int_{\theta_{0,d}}^{\theta_{d,p}} c(d,z)dz + \int_{\theta_{d,p}}^{1} c(p,z)dz >$$

$$\int_0^{\theta_{0,b}} c(0,z)dz + \int_{\theta_{0,b}}^{\theta_{b,m}} c(b,z)dz + \int_{\theta_{b,m}}^{\theta_{m,p}} c(m,z)dz + \int_{\theta_{m,p}}^{1} c(p,z)dz \qquad (3)$$

Inserting the cost function $c(e,\theta) = \frac{e}{2\theta}$ and the indifference levels from table 1 this leads to:

$$\frac{d}{2}\left[ln(p) - ln\left(\frac{d}{p}\right)\right] > \frac{b}{2}\left[ln\left(\frac{m}{p}\right) - ln\left(\frac{bp}{m}\right)\right] + \frac{m}{2}\left[ln(p) - ln\left(\frac{m}{p}\right)\right]$$

which under the above condition $m = d$ simplifies to:

$$1 > p^{-2b}m^{2b}b^{-b}$$

This is equivalent to $b > (\frac{m}{p})^2$; precisely contradicting the condition derived above to ensure an increasing sequence of indifference levels. It is established that the total cost of education rises after the introduction of a new signal; what happens to the individuals?

If we have $d = m$, the upper groups, master/ diploma and PhD graduates are as well off after the reform as they were before. Their wages and costs remain unchanged. For the Groups without degree the wages shrink after the reform given that we have $b > (\frac{m}{p})^2$. This is intuitive since the more productive top of this interval separates with the new signal, lowering the expected productivity of those choosing no signal.

The more interesting group to consider is the bachelor graduates who now gain the possibility of separating from the least productive types they were pooled with in the equilibrium before the reform. If for any, this separation is clearly profitable for the higher types within the new group of bachelor graduates. Denoting by θ^* the lowest type for which the new regime presents an improvement, this type can be derived by the lowest type for which

$$\frac{d}{2p} < \frac{1}{2}\left(\frac{dp}{m} + \frac{m}{p}\right) - \frac{2}{2\theta^*}$$

holds. This simplifies to $\theta^* > \frac{m}{p} = \theta_{b,m}$ which is precisely the first type to choose the master's, rather than bachelor's degree.

This result - that under the model with continuous types and discrete signals, the cost of separation exceeds the gains from a higher wage - generalizes as follows:

Suppose that types are uniformly distributed on the interval $[\underline{\theta}, \overline{\theta}] \subset \mathbf{R}_{++}$, and that initially there is no possibility of separation, so that the wage (and therefore utility) of each individual is $\frac{1}{2}(\underline{\theta} + \overline{\theta})$. Introducing an arbitrary signal s at cost $\frac{s}{2\theta}$ leads to an indifferent type

$$\theta_{0,s} = \frac{s}{\overline{\theta} - \underline{\theta}}$$

For this indifferent type to be within the interval $[\underline{\theta}, \overline{\theta}]$ (such that the signal is chosen) s needs to satisfy

$$\underline{\theta}(\overline{\theta} - \underline{\theta}) \leq s \leq \overline{\theta}(\overline{\theta} - \underline{\theta}) \tag{4}$$

Type θ^* is better off without the new signal if the wage increase induced by it is lower than the cost of obtaining the signal:

$$\frac{1}{2}\left(\frac{s}{\overline{\theta}-\underline{\theta}}+\overline{\theta}\right)-\frac{1}{2}(\underline{\theta}+\overline{\theta})<\frac{s}{2\theta^*} \tag{5}$$

If any types will be better off after a new signal, clearly the highest possible type would be among them. Therefore inserting $\theta^* = \overline{\theta}$ we get, $s\underline{\theta} < \underline{\theta}\overline{\theta}(\overline{\theta}-\underline{\theta})$ which simplifies precisely to the condition on s to be under the upper bar shown in 4. Hence 5 necessarily holds for all admissible θ, proving that the new signal worsens all types' situations.

Within the presented framework, introducing a new signal as a possibility to further separate according to productivity levels will not leave any individual better off.

Conclusion

Are students after the Bologna Reform worse off than preceding generations, studying in the old system? From a utilitarian perspective this theoretical framework leads to the conclusion that the former system was preferable.

While former diploma and present master graduates' utility is left unchanged, the less skilled individuals loose. For the least productive within this group, those who still don't obtain a higher education degree, this is due to a lower wage. For those who now obtain a bachelor's, the reason is that the higher effort involved in obtaining the newly introduced degree is not compensated by the wage increase induced by this signal.

This does not mean that the Bologna Reform should be seen as a failure. It was successful in achieving a higher mobility and decreased drop-out rates among students to mention only a few of the explicit goals. Furthermore, this analysis does not take certain factors on society as a whole into account, such as a more efficient allocation of workers to tasks.

It should further be mentioned, that the model's perspective on education is rather pessimistic. Since there is no human capital formation whatsoever the productivity isn't increased by education. The opportunity of signalling their own productivity level through effort always decreases overall utility.

References

BECKER, G. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70.

BEDARD, K. (2001): "Human Capital versus Signaling Models: University Access and High School Dropouts," *Journal of Political Economy*, 109(4), 749–775.

CHEVALIER, A., C. HARMON, I. WALKER, AND Y. ZHU (2004): "Does Education Raise Productivity, or Just Reflect it?," *The Economic Journal*, 114(499), F499–F517.

CHO, I.-K., AND D. M. KREPS (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102(2), 179–221.

DILL, V., AND A. HAMMEN (2011): "Einkommensdifferenziale zwischen Bachelor- und Diplomabsolventen - Humankapital oder Signal?, Working paper," *Research papers in Economics (University of Trier)*, (4/11).

FRAZIS, H. (2002): "Human capital, signaling, and the pattern of returns to education," *Oxford Economic Papers*, 54(2), 298–320.

OSBORNE, M., AND A. RUBINSTEIN (1994): *A Course in Game Theory.* MIT Press.

RINNE, U., AND Z. ZHAO (2010): "Human Capital vs. Signaling - The German Reunification as Natural Experiment," *Conference paper (IZA Bonn).*

SPENCE, M. (1973): "Job Market Signaling," *The Quarterly Journal of Economics*, 87(3), 355–374.

The Relationship between Time Preferences and Schooling Choices: An Empirical Analysis

Felix Lukowski *

Introduction

Time preferences play an important role in an individual's decision-making process at different stages of life. How long do I stay in school? How much effort do I devote to my job search? How much money do I save for my retirement? These decisions involve a trade-off between the present and the future. Intertemporal decision-making has a long tradition in economics and the discounted utility model, which has been the standard for most economic applications, goes back to Samuelson (1937).

The goal of this article is to investigate, whether an individual's time preferences are relevant for schooling choices. The notion is that allocating more time to education, one accepts a longer period of foregone earnings to achieve higher returns in the future (see Schmillen and Stüber (2014) for current statistics from Germany). This requires a certain degree of future-orientation.[1] Following this line of thought, one would expect a more patient individual to invest more time

*Felix Lukowski received his degree (M.Sc.) from the University of Bonn in 2014. The present article refers to his master thesis under supervision of Prof. Dr. Thomas Dohmen, which was submitted in December 2014.

[1]The expressions patience, being future-oriented, low discount rate and low rate of time preferences are used synonymously in the course of this article.

in schooling and choose longer educational pathways.

Using data from the German Socio-Economic Panel (SOEP), this article contributes to research in two ways. First, it investigates how time preferences can be measured. A direct measure provided by the SOEP as well as a self-created, indirect measure of time preferences are presented. In a second step, the two measures are used to examine the decision of German secondary school graduates who have completed their *Abitur*.[2] For this purpose, a multinomial logit model is applied. After obtaining their *Abitur*, young adults have basically three options. They can attend a university, pursue vocational training or neither (e.g., do temporary work). This article provides insights on the role of individual patience in this important life decision.

Related Literature

There exists a variety of articles examining the formation of time preferences, their empirical measurement as well as their role in predicting differences in economic or social outcomes. This overview begins with the latter.

Sutter *et al.* (2013) relate experimentally elicited time preferences to field behavior of adolescents. According to the authors' findings, impatience is a significant predictor of misconduct at school, lower savings, higher Body Mass Index (BMI) and negative health behavior, such as smoking and drinking alcohol. Using data from the National Longitudinal Survey of Youth (NLSY), Cadena and Keys (2012) find that being impatient has a negative effect on human capital investment and subsequently leads to a lower income when entering the labor force. Linking survey data on children's time preferences to Swedish administrative data, Golsteyn *et al.* (2014) show that, mediated by human capital investment during childhood, lower discount rates have a substantial, positive effect on dif-

[2] The *Abitur* is the highest secondary school qualification in Germany.

ferent lifetime outcomes, such as performance in school, health and labor market success. So several articles find a positive relationship between patience and different behaviors including human capital investment. This article builds upon this research and examines the role of time preferences in explaining German secondary school graduates' schooling choice.

Regarding the link between time preferences and education, there also exists literature postulating the converse relationship, i.e. education leads to lower discount rates. Becker and Mulligan (1997) provide a theoretical framework for the formation of time preferences and argue that schooling plays an important role in developing future-orientation. Exploiting the admission process of a public college in Mexico City, where admissions are determined through a lottery, Perez-Arce (2011) provides empirical evidence that education leads to a lower rate of time preferences. Taking these different results into account, this could indicate that causality works in both directions, so that patience and human capital investment mutually reinforce each other. Since this article only investigates one direction, choosing the appropriate timing is crucial. In order to avoid simultaneity bias, time preferences are measured before schooling choice takes place.

One important question for researchers is how to measure time preferences. While in laboratory experiments participants are asked to make lottery choices (Andersen *et al.*, 2008), these are usually not available using field data. For this reason some authors rely on proxies for impatience, e.g. smoking (Munasinghe and Sicherman, 2006). In addition to a SOEP survey item eliciting an individual's patience, this article employs an innovative approach by DellaVigna and Paserman (2005). In order to examine the effect of impatience on the length of unemployment spells, the authors create an indirect measure for time preferences. By conducting a factor analysis, they identify one common factor "impatience" among a multitude of variables. The authors find impatience to increase the length of unemployment spells.

Concerning the method to model schooling choices, an approach similar to Caner and Okten (2010) is applied in this article. The authors analyze the choice of students' college majors in Turkey. Since students face a limited set of alternatives, the choice is modeled using a multinomial logit. Establishing a risk and return framework of different college majors, the authors provide evidence that family income and parental risk attitudes influence young adults' selection.

Data and Measures

The SOEP

The German Socio-Economic Panel (SOEP) is a representative, longitudinal panel of private households in Germany. It has been conducted in western Germany since 1984 and was expanded to include eastern Germany in 1990. Every year about 30,000 individuals from nearly 11,000 households in Germany are interviewed about different domains of their life such as education, employment, income and health as well as attitudes and different satisfaction measures. For a detailed description, see Wagner *et al.* (2007).

Schooling Choice and Sample Selection

This article uses the SOEP waves from 2008 to 2012 for the analysis. The dependent variable, schooling choices by secondary school graduates, determines how the observation sample is selected. Time preferences are measured in 2008 and their effect on an individual's schooling choice in the subsequent years from 2008 to 2012 is examined. Only respondents who were between 17 and 21 years old in 2008 are included. Statistics from the Federal Statistical Office (2008) indicate that over 95% of the secondary school graduates in 2008 were under 22 years old, so taking age 21 as a cutoff captures the vast majority of the students. The minimum age of SOEP respondents and therefore the lower bound of the analysis is 17. This will not prove problematic, since cases, where a 16-year old already graduates from secondary school, can be considered very rare. The sample is restricted to students who have obtained their general maturity certificate (*Abitur*) or an equivalent degree at most one year prior to the observation period and have started to pursue tertiary education (category [2]), vocational training (category [1]) or neither (category [0]) in the observation period.[3] By its construction, the last category does not have a clear interpretation. There are different reasons as to why somebody has not made a choice yet. Individuals might carry out military or civil service, spend a voluntary social year, do temporary work, be unemployed or take a gap year and travel. In the following, this category is going to be referred to as *other*. In light of these considerations, this restricts the sample to a total of 349 students.

As can be seen in figure 01, the majority of the respondents (53.58%) enter a tertiary education program. Furthermore, 26.93% decide to enter vocational training and for the rest, the decision is postponed or their status is not clear. The gender-specific age distribution is illustrated in figure 02. In terms of age, one can see that a large proportion of the respondents is between 18 and 20 years old, when patience is measured. There are slightly more females (180) than males (169), so there is an almost equal gender distribution among age groups.

Patience Measures

This article applies two different approaches to measure time preferences, a direct and an indirect one. The strategy is to conduct an analysis with both measures

[3] It is important to note that only the initial choice is taken into account when forming the categories, because the intention of this article is to model the decision after graduation. So somebody switching from tertiary education to vocational training is considered to be in category [2].

separately as well as conjointly and then evaluate their predictive power in the context of schooling choices.

Direct Measure The SOEP provides one survey item designed to elicit an individual's time preferences. In 2008, respondents were asked to rate their patience on a Likert scale from "very impatient" (0) to "very patient" (10). This ultra-short measure of patience has been validated by Vischer *et al.* (2013) and shown to predict an individual's inter-temporal choice behavior in an incentivized experiment. One potential drawback of this patience measure could be the perception of the word patience in everyday language, since it might also be used to describe endurance (e.g., being patient with a child) instead of a trade-off between now and the future. Similarly to Jaeger *et al.* (2010), who have used an equally coded SOEP measure for an individual's risk attitude, a patience index using the responses on the scale as well as a binary variable are used. The latter is scaled as 1 if a value of 7 or higher is chosen and 0 otherwise. It is going to be referred to as patience indicator. Constructing the indicator helps to mitigate problems arising from respondents who interpret the scale differently. The main analysis of this article employs the patience indicator. Figure 03 displays the distribution of individuals' responses to the question eliciting their patience.

Indirect Measure For creating an indirect measure of patience, an approach similar to DellaVigna and Paserman (2005) is used. It is based on the assumption that individuals indirectly reveal their time preferences by engaging in certain activities. The idea is that relatively patient individuals are more likely to pursue activities having immediate costs and delayed rewards. Reversely, this means that impatient individuals are more likely to take on activities with immediate gratification and delayed costs.

Principal Component Analysis A Principal Component Analysis (PCA) is conducted to identify a component "patience" from a set of variables. PCA[4] is a data reduction method that can be used to reduce the dimensionality of a set of interrelated variables, while keeping the highest variation possible. A new set of uncorrelated variables, called the Principal Components (PCs), is created, where the first few PCs already contain most of the variation in the dataset. In this way, a multitude of variables can be summarized with relatively few components capturing as much information as possible from the original variables. It should be noted that these measures by themselves would be very noisy proxies for patience, so that the strength of PCA lies in reducing the information. In the following, the variables chosen to be included in the PCA are presented.

Smoking Smoking is probably the most frequently used proxy for time preferences (e.g., Munasinghe and Sicherman (2006)). Someone who smokes, accepts serious health problems in the long-run in order to have the immediate gratification of smoking a cigarette. SOEP respondents indicate, whether they smoke or not. A dummy is constructed, which is coded as 1 for smokers.

Regular Alcohol Consumption This variable follows the same notion as smoking cigarettes. Drinking alcohol on a regular basis, one accepts negative effects on health in the future. In the SOEP, there are four categories of alcohol, namely beer, (sparkling) wine, spirits and mixed drinks, and for each category the respondent indicates how often he or she consumes it. The answer options are "regularly" (1) , "once in a while" (2), "seldom" (3) and "never" (4). A dummy is constructed, which is 1, if the respondent consumes alcohol regularly according to one of the four categories.

[4]The explanations are based onJolliffe (2005).

BMI The SOEP also provides information on the respondent's BMI. Sutter *et al.* (2013) have shown empirically that a high rate of time preferences could lead to a higher BMI. A more present-oriented person might exercise less or eat more unhealthy food than a more future-oriented person.

Health-Conscious Diet The idea is similar as for BMI, healthy nutrition reveals patience. Respondents are asked, whether they follow a health-conscious diet. The answer options are "very much"(1), "much"(2), "not so much" (3) and "not at all" (4). The items are recoded so that "very much" corresponds to the highest value.

Leisure Activities Respondents indicate how often they take part in active sports and in artistic and musical activities. The options are "daily" (1), "weekly" (2), "monthly" (3), "less frequently" (4) and "never" (5). Both activities are included because they involve frequent practicing and, depending on the goal, it might take a long time to achieve it. The items are reversed so that a higher number means practicing more frequently.

Household's Saving Attitude The SOEP household questionnaire provides one item asking, whether a monthly amount of money is saved for emergencies or large purchases. This could indicate patience, because one keeps money aside for (unexpected) future events. Although the item does not capture, if the individual himself saves money, one could argue that this is something a child learns from the parents. Hence, if it is common in the household to save money, it is more likely that the individual has picked up this behavior. Since no amount of money is specified, this variable does not so much measure wealth or income (although this might play a role), but rather the household's attitude towards saving. A dummy is created, coded 1 if a monthly amount is saved and 0 if not.

Correlations To get a first overview of how the items are related, their correlations are examined. It is interesting to see if, for example, activities assumed to reveal impatience are negatively correlated with activities assumed to reveal patience. Table 1 displays the pairwise correlations of the items and their corresponding p-values. The correlations provide mixed evidence. Some measures are correlated in a way one would expect, some are not. For example smoking and monthly household savings have a significant, negative correlation of -0.1774. BMI on the other hand does show a significant correlation with any of the other measures. Not finding correlations between all measures does not necessarily imply that they do not capture a common factor "patience". Clear correlations would strengthen the argument, but it could also be the case that they measure different aspects of patience.

PCA Results The relevant components from the PCA can be chosen by two criteria indicating the components already explaining the largest share of the variance. The first is to look at the scree plot of the eigenvalues and maintain only components before the kink in the plot. The second is to consider only components with an eigenvalue ≥ 1. Applying these two criteria, the first four respectively the first three components have to be taken into account. The scree plot and the corresponding results for all components are displayed in figure 4 and table 2. Table 3 displays the items' loadings for the first four components. One can see that the first component's eigenvectors have a direction which would be in line with a common component "patience". Regular alcohol consumption, smoking and BMI have a negative sign, whereas the other four items have a positive sign. The rest of the components do not reveal loadings that could be interpreted as patience. However, it is not clear, whether the first component really extracts an underlying factor "patience" or whether it captures something else. Smoking, following a health-conscious diet and active sports receive the highest loadings in absolute value, so the underlying factor could also be a preference for sports or a healthy lifestyle. On the other hand, a healthy lifestyle might also indicate a strong future-orientation. These considerations show that one has to be very cautious in interpreting the results of the PCA. In the following, the identified component is going to be referred to as indirect patience measure. However, the reader should keep in mind that there might be some objections against this measure.

In a next step, the correlation between the indirect patience measure and the direct one elicited in the SOEP is examined. The two measures have a correlation of 0.0980 (p-value = 0.0874). Considering that both measures are supposed to capture the same trait, this might seem relatively low. Nevertheless, it is not uncommon to observe low correlations for different measures of an underlying trait (Scheinkman and Soutter (2000)).

Control Variables

In this section, the control variables, written in italics, are presented. *Impulsiveness* could be argued to capture a situational form of impatience. This article controls for an individual's *Impulsiveness* by using a self-assessed SOEP measure. Socioeconomic background is an important determinant of career choice and success (e.g.,Ermisch and Francesconi (2001)). In order to capture that effect, *parents' education* measured as their highest school diploma is included in the analysis. In addition, the household's financial resources are also likely to have an impact on the schooling decision of secondary school graduates. Tertiary education signifies a long period of studying without any income (except for part-time jobs). Even though the German government supports students from poorer families financially, knowing that the parents can provide for the studies might influence an individual's decision. For, this reason the *household's wealth and income* are included as control variables. Caner and Okten (2010) have shown that career choice involves taking risks, since the returns in the labor market are uncertain. Therefore, *risk aversion* is controlled for by including a survey item, where respondents rate their willingness to take risks. An individual's schooling choice is restricted by the *Secondary School Grades.* The selection for some higher education and vocational training programs is very competitive, so that not all graduates have the opportunity to succeed. The secondary school grades for the subjects math, German and the first foreign language are averaged and incorporated in the analysis. Finally, *demographic variables*, namely *age* and *gender*, are included. In the setting of this article, it is very important to consider someone's age, since older individuals have had more time to make their choice. Controlling for gender-specific effects also makes sense, because the preferences for occupations might differ, so that different educational paths are chosen to obtain the necessary qualification for a job.

Methodology

In order to model the choice secondary school graduates make after their *Abitur*, a multinomial logit model[5] is applied. The starting point of the model is an individual i who makes a schooling choice j in order to maximize his utility U_{ij}. This can be represented by a random utility model, which assumes the utility U_{ij} to be a linear response,

$$U_{ij} = \boldsymbol{w}_i'\boldsymbol{\alpha}_j + \boldsymbol{\epsilon}_{ij} \quad , \tag{1}$$

[5] The derivations in this section are based on Greene (2008).

where $\boldsymbol{w}_i$ is a K-dimensional vector consisting of the characteristics of individual i, and $\boldsymbol{\alpha}_j$ yields the alternative-specific coefficients. The error term $\boldsymbol{\epsilon}_{ij}$ is purely stochastic and included to compensate for model imperfections. If j is chosen, U_{ij} is assumed to give the consumer the maximum utility among the J choices:

$$U_{ij} = \max(U_{i1}, \ldots, U_{iJ}) \quad . \tag{2}$$

Let Y be a random variable indicating the choice made. If and only if the error terms $\boldsymbol{\epsilon}_{ij}$ are assumed to be mutually independent and identically distributed with log-Weibull distribution

$$F(\boldsymbol{\epsilon}_{ij}) = \exp(-\exp(-\boldsymbol{\epsilon}_{ij})) \quad , \tag{3}$$

and α is normalized $\alpha_0 = 0$, this leads to:

$$\text{Prob}(Y_i = j) = P_{ij} = \frac{\exp(\boldsymbol{w}_i'\boldsymbol{\alpha}_j)}{1 + \sum_{k=1}^{J} \exp(\boldsymbol{w}_i'\boldsymbol{\alpha}_k)} \quad , j = 0, 1, \ldots, J;\ \alpha_0 = 0 \quad . \tag{4}$$

P_{ij} denotes the probability of individual i to choose one of the three alternatives j, tertiary education (category [2]), vocational training (category [1]) or other (category [0]). The latter comprises those individuals, for whom the decision is postponed or their status is not clear. The choice depends on an individual's characteristics $\boldsymbol{w}_i$. In the framework of this article, these consist of an individual's patience and a set of control variables. To allow for a better comparability and interpretation of the results, all continuous variables (except for age) are standardized to have a mean of 0 and a standard deviation of 1.

Results

In this section, the estimation results of the multinomial logit are presented. For the estimation, the patience indicator and the indirect patience measure are used. The former might be better suited to make predictions, since it is not as prone to scaling differences as the index, even though this comes at the cost of losing information from the initial measure.[6] Table 4 shows the results for the two patience measures, separately and conjointly. Vocational training [1] is taken as base outcome and the other two categories are compared to it. The focus lies on the comparison between vocational training [1] and tertiary education [2], since the category other [0] covers a heterogeneous group of respondents.

Using the patience indicator (column 1 and 2), a higher level of patience makes it more likely to enter a tertiary program than vocational training. The effect is significant at the 5% level and remains unchanged in the full specification with both patience measures. Looking at the indirect patience measure (column 3 and 4), there is no significant effect on schooling choice.[7]

What can be seen in all specifications is that higher average grades as well as a higher educational level of the father make the choice of a tertiary program significantly more likely. The other control variables are not significant. Only in the specification in column 4, being older make an individual more likely to enter tertiary education and is significant at the 10% level.

These results only allow for a “more likely” and “less likely” interpretation. To evaluate the effects of the regressors on a probability scale, one needs to calculate the marginal effects, which are presented in the next paragraph.

[6] It should be noted that due to unfavorable overlap in responses to schooling choice, the patience measures and some important control variables like parents’ education and average grades, the estimation sample is reduced to 198 and 194 observations respectively.

[7] The results of using the direct patience measure as a scale are not displayed. They do not show any significant effects of patience on schooling choices.

Marginal Effects Table 5 shows the marginal effects evaluated at the mean of the explanatory variables for the full specification with both patience measures. According to the direct measure, being more patient (i.e. ranking 7 or higher on the patience scale) decreases the probability of pursuing vocational training by 16.3% and increases the probability to enter tertiary education by 15.7%. The effects are significant at the 5% and 10% level respectively. As could be expected from the main results, the marginal effects of the indirect patience measure are not significant.

Taking a look at the control variables, one can see that a one standard deviation increase in average grades significantly increases the probability to enter tertiary education by 14.4%, which is significant at the 1% level. It decreases the probability to pursue vocational training by 6.8% being significant at the 10% level. The strongest effect in terms of magnitude can be seen for the level of father's education. A one standard deviation higher education decreases the probability to enter vocational training by 19.6% and increases the probability to pursue tertiary education by 20.3%. These effects are both significant at the 1% level. An age effect can also be observed. Being one year older significantly increases the probability that someone chooses tertiary education by 10.5%.

Discussion

Direct Patience Measure The findings of the patience indicator suggest that individuals who are more patient are significantly more likely to enter tertiary education than to pursue vocational training. These findings give support to the idea that a lower rate of time preferences increases the probability of individuals choosing longer educational paths. Also, the magnitude of around 16% is considerable. However, due to the small sample size, the results should at most be considered indicative. Using the item as a scale and not as an indicator fails to deliver the same results. This could be caused by the small sample, but it should be taken into account that a scale based on self-evaluation might be very prone to individuals' self-perception and their understanding of the scale. Additionally, one has to argue, whether the word patience in everyday language is the same as what an economist defines as discount rates or time preferences.

Indirect Patience Measure This article has also implemented an alternative approach, based on DellaVigna and Paserman (2005). Although the idea of collecting items revealing an individual's preferences for immediate or delayed gratification is very appealing, it has its drawbacks. The selection of the right items to include in a PCA or factor analysis requires careful considerations. When evaluating the results and identifying a component having the predicted loadings, it is not always clear, whether the underlying factor captures patience. Therefore, it is a very exploratory approach and requires a lot of reasoning. Unfortunately, the results obtained by PCA do not provide further insights on the relationship between patience and schooling choice. There might be several reasons for this. The item selection for the PCA might contain too many health-related items, so that a component "healthy lifestyle" is extracted. Although living healthily could imply a strong future-orientation, it might not adequately measure an individual's time preferences. The failure to obtain results could also be due to the small sample size. Regardless of the reasons, one can note that, if used as an indicator, the direct patience measure in the SOEP shows stronger predictive power than the self-created, indirect measure.

Control Variables The average grades obtained in secondary school and the father's educational level seem to be important determinants of the choice between vocational training and tertiary education. The latter implies that children, whose parents obtained a university degree, are also more likely to pursue higher education. Only some control variables being significant, could be due to the small sample or high correlations between the controls. For example, parents' education and income might be closely related.

Conclusion

In this article, the relationship between time preferences and schooling choice is analyzed using data from the years 2008 to 2012 of the SOEP. Applying a multinomial logit model, the effect of secondary school graduates' time preferences on their schooling decision in the subsequent years is estimated. By measuring patience before an educational path is chosen, the mutual interference of time preferences and schooling is taken into account. In the analysis, the predictive power of two different measures for patience is compared.

The first one is a SOEP survey item in which an individual is asked to rate his patience on a Likert scale from 1 to 10. A dummy is constructed being 1 if someone rates himself 7 or higher and 0 otherwise. The measure predicts that being relatively more patient, significantly increases the likelihood of entering tertiary education and decreases the likelihood of entering vocational training by about 16%. When the same item is used as a scale, it does not deliver the same results hinting towards problems with individuals interpreting the scale differently.

The second measure is based on an approach by DellaVigna and Paserman (2005). The authors argue that an individual's time preferences are revealed by engaging in activities which involve immediate costs and delayed rewards or the other way around. By collecting items potentially revealing an individual's patience and performing a factor or principal component analysis, one can reduce the information of the different items and identify a common factor "patience". This approach has a very exploratory nature and it is difficult to evaluate, whether the identified factor measures patience or another underlying trait. In this article, a PCA is conducted and one component potentially capturing patience is used for the estimation. Not showing any significant effects, it fails to make the same predictions as the SOEP measure.

The article has also shown that it is difficult to find appropriate measures for time preferences in longitudinal data, where lottery choices might not be feasible to implement. The ultra-short survey item in the SOEP provides a straightforward measure, but using a self-assessed scale could be error-prone. The revealed preferences approach offers the possibility to construct a patience measure from different survey items. However, it requires a careful selection of the items and the interpretation can be complicated.

Due to its small sample size, this article has its limitations. An analysis with a larger sample would allow for more substantial claims. Nevertheless, the results obtained by using the direct patience measure could be seen as indicative. The relationship between time preferences and schooling is an interesting field of research and the interaction between the two requires further investigation. Therefore, the scientific community would highly benefit from a theory-based structural model allowing for mutual reinforcement.

References

ANDERSEN, S., HARRISON, G. W., LAU, M. I. and RUTSTRÖM, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, **76** (3), 583–618.

BECKER, G. S. and MULLIGAN, C. B. (1997). The endogenous determination of time preference. *The Quarterly Journal of Economics*, **112** (3), 729–758.

CADENA, B. C. and KEYS, B. J. (2012). Human capital and the lifetime costs of impatience. *Available at SSRN 1674068*.

CANER, A. and OKTEN, C. (2010). Risk and career choice: Evidence from turkey. *Economics of Education Review*, **29** (6), 1060–1075.

DELLAVIGNA, S. and PASERMAN, M. D. (2005). Job search and impatience. *Journal of Labor Economics*, **23** (3), 527–588.

ERMISCH, J. and FRANCESCONI, M. (2001). Family matters: Impacts of family background on educational attainments. *Economica*, **68** (270), 137–156.

FEDERAL STATISTICAL OFFICE (2008). Bildung und Kultur, Allgemeinbildende Schulen. *DESTATIS*, **Fachserie 11 Reihe 1** (Schuljahr 2007/08), Tabelle 6.6.

GOLSTEYN, B. H., GRÖNQVIST, H. and LINDAHL, L. (2014). Adolescent time preferences predict lifetime outcomes. *The Economic Journal*, **124** (580), F739–F761.

GREENE, W. (2008). *Econometric Analysis, Six Edition*. Pearson Education.

JAEGER, D. A., DOHMEN, T., FALK, A., HUFFMAN, D., SUNDE, U. and BONIN, H. (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics*, **92** (3), 684–689.

JOLLIFFE, I. (2005). *Principal component analysis*. Wiley Online Library.

MUNASINGHE, L. and SICHERMAN, N. (2006). Why do dancers smoke? smoking, time preference, and wage dynamics. *Eastern Economic Journal*, pp. 595–616.

PEREZ-ARCE, F. (2011). The effect of education on time preferences.

SAMUELSON, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, **4** (2), 155–161.

SCHEINKMAN, J. A. and SOUTTER, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, **115** (3), 811–846.

SCHMILLEN, A. and STÜBER, H. (2014). Lebensverdienste nach Qualifikation. Bildung lohnt sich ein Leben lang. *IAB-Kurzbericht*, (1), 1–8.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D. and Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *The American Economic Review*, **103** (1), 510–531.

Vischer, T., Dohmen, T., Falk, A., Huffman, D., Schupp, J., Sunde, U. and Wagner, G. G. (2013). Validating an ultra-short survey measure of patience. *Economics Letters*, **120** (2), 142–145.

Wagner, G. G., Frick, J. R. and Schupp, J. (2007). The German Socio-Economic Panel study (SOEP) - evolution, scope and enhancements.

Appendix

Schooling Choice	Freq.	Percent	Cum.
Other [0]	68	19.48	19.48
Vocational Training [1]	94	26.93	46.42
Tertiary Education [2]	187	53.58	100.00
Total	349	100.00	

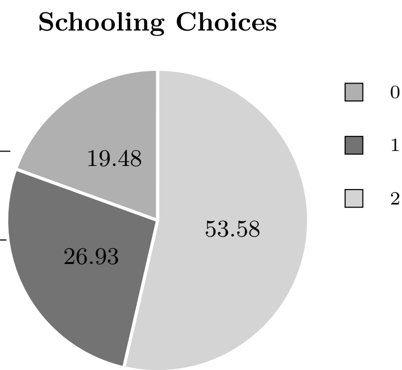


Figure 01: The figure shows the distribution of schooling choices within the selected sample in 2008. N=349; Source: SOEP 2008.

	(1)	(2)	(3)	(4)	(5)	(6)
Regular Alcohol Cons. (1)	1					
Smoker (2)	0.146	1				
	(0.009)					
BMI (3)	0.057	0.045	1			
	(0.315)	(0.432)				
Health-Conscious Diet (4)	-0.216	-0.227	-0.039	1		
	(0.000)	(0.000)	(0.487)			
Saving Attitude (5)	0.115	-0.177	-0.023	-0.050	1	
	(0.043)	(0.002)	(0.65)	(0.380)		
Active Sports (6)	0.099	-0.191	0.061	0.295	0.143	1
	(0.079)	(0.001)	(0.285)	(0.000)	(0.012)	
Artistic Activities (7)	-0.005	-0.057	-0.088	0.058	0.127	0.065
	(0.925)	(0.318)	(0.120)	(0.304)	(0.026)	(0.250)

Table 1: The table shows the pairwise correlations between the items used for the PCA. P-values are reported in parentheses. All values are rounded to three decimal places. N=316. Source: SOEP 2008.

Age [Years]	Female	Male	Total
17	27	21	48
18	52	57	109
19	48	42	90
20	38	36	74
21	15	13	28
Total	180	169	349

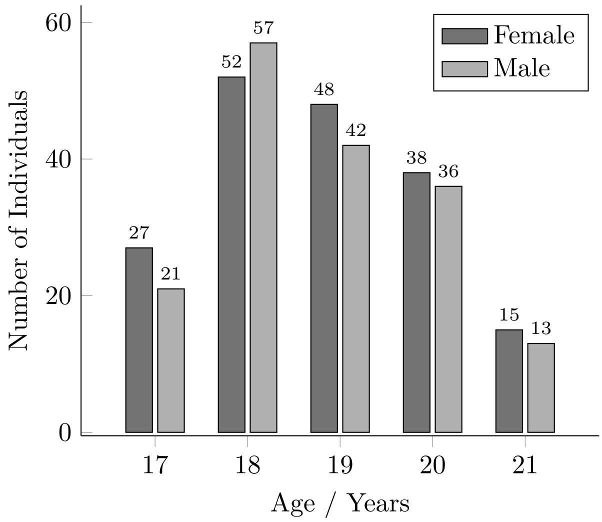


Figure 02: The figure displays the gender-specific age distribution within the selected sample in 2008. N=349; Source: SOEP 2008.

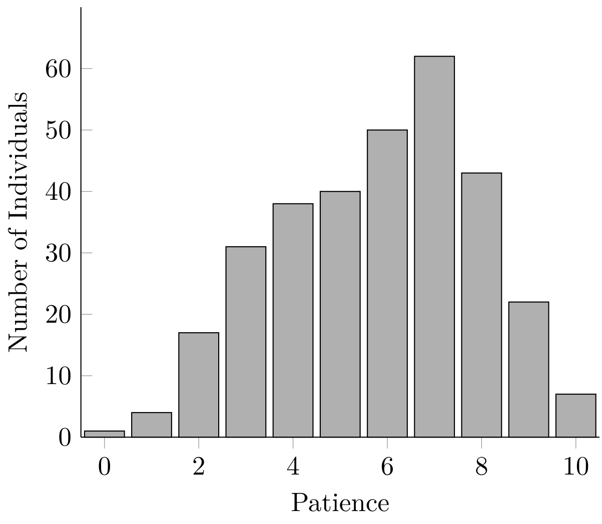


Figure 03: The figure shows the distribution of the direct patience measure within the selected sample in 2008. The patience index ranges from “very unpatient” [0] to “very patient” [10]. N=315; Source: SOEP 2008.

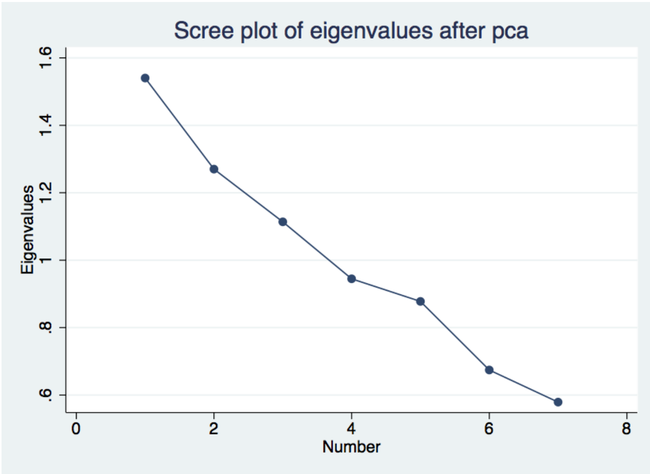


Figure 04: The figure displays the scree plot for the selected sample. The x-axis shows the components and the y-axis their corresponding eigenvalues. No rotation used; N=306. Source: SOEP 2008.

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	1.5402	0.270 122	0.22	0.22
Comp2	1.270 08	0.156 332	0.1814	0.4015
Comp3	1.113 75	0.168 916	0.1591	0.5606
Comp4	0.944 831	0.067 108	0.135	0.6956
Comp5	0.877 723	0.203 396	0.1254	0.8209
Comp6	0.674 327	0.095 235	0.0963	0.9173
Comp7	0.579 092	.	0.0827	1

Table 2: The table shows the PCA components for the selected sample. No rotation used; N=306. Source: SOEP 2008.

Variable	Comp1	Comp2	Comp3	Comp4
Regular Alcohol Consumption	−0.2276	0.6308	0.1159	0.4068
Smoker	−0.5425	0.0038	0.0616	0.4345
BMI	−0.1054	0.1972	0.6348	−0.4108
Health-Conscious Diet	0.5474	−0.3079	0.2355	0.2818
Saving Attitude	0.2487	0.5795	−0.208	−0.4388
Active Sports	0.4827	0.3039	0.3696	0.4039
Artistic Activities	0.2197	0.2008	−0.5869	0.2045

Table 3: The table displays the eigenvectors for the first four components explaining 70% of the total variance. No rotation used; N=306. Source: SOEP 2008.

	(1) Other	(2) Tertiary	(3) Other	(4) Tertiary	(5) Other	(6) Tertiary
Patience (Dir. Measure)	0.799	0.894**			0.699	0.922**
	(0.512)	(0.414)			(0.531)	(0.432)
Patience (Ind. Measure)			0.273	0.220	0.230	0.201
			(0.306)	(0.199)	(0.248)	(0.166)
Impulsiveness	0.0666	0.0524	-0.0969	-0.0322	-0.0372	0.0386
	(0.290)	(0.217)	(0.285)	(0.219)	(0.287)	(0.234)
Risk Attitude	-0.212	-0.363	-0.108	-0.325	-0.141	-0.358
	(0.366)	(0.228)	(0.363)	(0.234)	(0.368)	(0.232)
Average Grades	-0.0922	0.560**	-0.274	0.466**	-0.237	0.512**
	(0.310)	(0.223)	(0.341)	(0.228)	(0.340)	(0.224)
Father's Education	0.697**	1.162***	0.717***	1.086***	0.748***	1.131***
	(0.277)	(0.241)	(0.270)	(0.229)	(0.280)	(0.245)
Mother's Education	-0.154	-0.325	-0.130	-0.298	-0.130	-0.316
	(0.308)	(0.259)	(0.294)	(0.230)	(0.308)	(0.255)
1 if Male	0.103	0.0847	0.0160	0.145	-0.000272	0.0929
	(0.502)	(0.401)	(0.513)	(0.406)	(0.513)	(0.409)
Age	-0.290	0.353	-0.229	0.381*	-0.265	0.341
	(0.292)	(0.218)	(0.312)	(0.223)	(0.308)	(0.222)
HH Income	0.148	0.205	0.0874	0.183	0.0682	0.167
	(0.369)	(0.305)	(0.368)	(0.295)	(0.374)	(0.300)
HH Wealth	-0.823	0.520	-0.788	0.597	-0.910	0.497
	(1.373)	(0.601)	(1.423)	(0.662)	(1.482)	(0.596)
Constant	4.139	-7.200*	3.264	-7.360*	3.677	-6.950*
	(5.406)	(4.139)	(5.696)	(4.201)	(5.688)	(4.202)
Observations	198	198	194	194	194	194
Pseudo R-squared	0.1847	0.1847	0.1738	0.1738	0.1876	0.1876

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: The table reports the estimation results of the multinomial logit model. The reference category is vocational training.
(1)-(2): Direct Patience Measure (Patience Indicator)
(3)-(4): Indirect Patience Measure
(5)-(6): Both Measures
Sample: SOEP 2008 - 2012.

	(1) Other	(2) Vocational	(3) Tertiary
Patience (Dir. Measure)	0.005	-0.163**	0.157*
	(0.056)	(0.077)	(0.087)
Patience (Ind. Measure)	0.011	-0.038	0.027
	(0.027)	(0.030)	(0.034)
Impulsiveness	-0.008	-0.004	0.013
	(0.033)	(0.040)	(0.050)
Risk Attitude	0.014	0.058	-0.073
	(0.043)	(0.042)	(0.052)
Average Grades	-0.076**	-0.068*	0.144***
	(0.033)	(0.041)	(0.047)
Father's Education	-0.007	-0.196***	0.203***
	(0.029)	(0.043)	(0.048)
Mother's Education	0.012	0.052	-0.064
	(0.033)	(0.045)	(0.053)
1 if Male	-0.008	-0.014	0.022
	(0.058)	(0.072)	(0.087)
Age	-0.064**	-0.041	0.105**
	(0.032)	(0.039)	(0.044)
HH Income	-0.006	-0.027	0.034
	(0.040)	(0.054)	(0.060)
HH Wealth	-0.160	-0.041	0.201
	(0.170)	(0.113)	(0.164)
Observations	194	194	194

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: The table reports marginal effect estimates of the multinomial logit model evaluated at the mean of the explanatory variables. Sample: SOEP 2008 - 2012.

Model-Independent Bounds for Option Prices

Kilian Russ *

Introduction

The 2008 financial crisis and its dramatic consequences fuelled an ongoing debate on the principles and methods underlying the financial industry. Not least, derivative trading drew a lot of attention as the stakes at play are typically humongous. Critical voices questioned the standard approach to model future behaviour based on past observation. A drastic view on this was expressed by Nassim Taleb.

> He [Robert C. Merton] starts with rigidly Platonic assumptions, completely unrealistic - such as the Gaussian probabilities, along with many more equally disturbing ones. Then he generates "theorems" and "proofs" from these. The math is tight and elegant. [...] He assumes that we know the likelihood of events. The beastly word equilibrium is always present. But the whole edifice is like a game that is entirely closed, like Monopoly with all of its rules.
>
> Taleb (2010), The Black Swan, p. 282

*Kilian Russ received his degree in Economics (B.Sc.) from the University of Bonn in 2015. The present paper refers to his bachelor thesis under supervision of Prof. Dr. Klaus Sandmann.

A more constructive stream of criticism regarded the main problem as not lying within the models itself, but rather in the naive and excessive reliance on such. The consequential attempts and suggestions for improvement are well summarised in the following statement.

> For banks the only way to avoid repetition of the current crisis is to measure and control all their risks, including the risk that their models give incorrect results.
>
> Shreve (2008)

With this statement in mind we want to recall a concept originally introduced by Knight. Knight was the first to formally make a distinction between risk and uncertainty. While the former is referred to as the "known unknown" and captures the randomness within a specified probabilistic framework and is hence measurable within the model, the latter, or "unknown unknown", accounts for the possibility that the model is not a satisfactory abstraction of reality (Obłój (2013)). In the context of financial or economic modelling this is also known as the problem of model misspecification. As derivative pricing traditionally requires strong modelling assumptions the described problem is severe, meaning that small variations in the modelling choice may cause drastic changes in the results. In the following we present an alternative approach for the pricing of exotic options, which is robust to model misspecification.

This work attempts to provide an introduction to the Skorokhod embedding technique and how it can be employed in the context of derivative pricing to obtain model-independent price bounds. For illustrative purposes we focus on one exemplary type of exotic option, namely digital double no-touch barrier options. Section 2 gives a brief recap on the standard theory of mathematical finance and explains how the new framework relates to it. Just as in standard financial modelling the principle of replication is the centrepiece of the new approach. Pricing is done through replication or hedging. However, in the new framework it is not possible to perfectly replicate the exotic option's payoff in general. Nevertheless, one can construct sub- and superhedging portfolios which pay out at most or at least as much as the exotic option, respectively. Section 4 derives the sub- and superhedging strategies in full detail. By simple no-arbitrage arguments the prices of such portfolios yield a lower and upper bound on the price of the option. The challenging goal is to find the most expensive and the cheapest sub- and superhedging strategy, respectively. While the constructed portfolios in section 4 are rather simple, the difficult part is to prove that the obtained price bounds are sharp, i.e. that there always exists a possible price process, such that the bounds are attained. For this we resort to the so called Skorokhod embedding technique, which is introduced in section 3.[1] Section 5 translates our results to more realistic setups and briefly summarises an assessment of their performance compared to traditional hedging methods conducted by Ulmer (2010). All proofs are omitted.

Double Barrier Options

Digital double no-touch barrier options belong to the class of path-dependent contingent claims as their payoff depends on the current and past evolution of the underlying. Concretely, the option gives its holder a standardised payoff of 1 if and only if the underlying stays strictly within a pre-specified interval until some maturity time T. Since there is no closed form solution for the price of such options available, practitioners rely on numerical approximations to perform hedging strategies. Formally, we refer to $(S_t)_{t\in[0,T]}$, short (S_t), as the price process of the underlying which is assumed to be continuous throughout this work. We denote by $\underline{b} < S_0 < \overline{b}$ the lower and upper bound, which must not be

[1]For a comprehensive survey of the Skorokhod embedding problem see Obłój (2004).

undercut or exceeded, respectively. The payoff can be written as $\mathbb{1}_{\underline{b}<\underline{S}_T, \overline{S}_T<\overline{b}}$.[2]

A Robust Framework for Derivative Pricing

Standard Modelling in Mathematical Finance

Contemporary mathematical finance has its roots back in Louis Bachelier's thesis "Théorie de la Spéculation" in 1900. However, it was not until the early 1970s with the contributions of Black and Scholes (1973) and Merton (1969), when the field began to expand rapidly. Following the description by Obłój and Hobson (Obłój (2013)) the standard approach in mathematical finance, in particular in derivative pricing, is to postulate a model by specifying a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and an adapted stochastic process (S_t). The fair price of a contingent commodity claim with underlying (S_t) is then calculated as a discounted expected value under an equivalent martingale measure. The existence of such martingale measures turns out to be equivalent to one of the major hypotheses in financial economics, namely the principle of market efficiency. This result is known as the First Fundamental Theorem of Asset Pricing. At the very heart of this result lies the premise that prices can be derived through hedging and duplication – a principle we will not retreat from in the robust approach. The elegance and simplicity of solutions obtained under these specific modelling assumptions, for instance the prominent Black-Scholes formula for the price of a European call option, made it the dominant framework to work with over the past decades.

As Obłój (2013) points out, there are at least three major shortcomings of the standard approach. First, it entirely ignores information currently available in the market, such as prices of other traded options. Second, specifying the

[2] $\mathbb{1}_A$ denotes the indicator function of the set A. Running maxima and minima of (S_t) are defined as $\overline{S}_T := \sup_{u\le T} S_u$ and $\underline{S}_T := \inf_{u\le T} S_u$, respectively.

process to follow a certain distributive law, e.g. Geometric Brownian motion, both is responsible for the unique and mathematically elegant solutions, but also constitutes a strong modelling assumption. Third, markets are evidently not frictionless and transaction costs, liquidity constraints and risk of default pose considerable limitations on the implementability of trading and hedging strategies.

The Robust Framework

Before we introduce the new framework we want to take a step back and briefly reflect on how financial models are built following the description by Obłój (2013). The main mechanism of a model is to transform given inputs into outputs according to some reasoning doctrine. One can distinguish three types of inputs: beliefs, information and rules. Beliefs refer to the assumption on the evolution of the underlying, i.e. the chosen structure of randomness. Information means all market data, which we would like our model to treat consistently. Lastly, rules define how the modelled interaction takes place. What is traded when by whom. Especially market frictions, if any, are incorporated here. The reasoning doctrine in our context is market efficiency, resulting in absence of arbitrage opportunities. Obviously, this outline of model building is very much tailored towards financial market models, nevertheless, it provides a nice intuition on what the key ingredients are one has at his dispense. It is worth noticing how the classical Black-Scholes model fits into this general description of a model. The beliefs are very strict in the sense that the evolution of the stochastic process of the underlying is assumed to be of a known Geometric Brownian motion type. On the contrary, the information only consists of today's asset price. The rules are kept as simple as possible, in particular, abstracting from any kind of market frictions.

The main idea for the robust framework is to *milden* beliefs which are encoded through a choice of space of possible paths of the risky asset and to *increase* the amount of information used in the model.

In the rest of this section we introduce the setting of this thesis sticking to the formalism of Cox and Obłój (2011). Throughout we work in finite time $t \in [0, T]$. Denote by $\mathscr{P}$ the set of all continuous, non-negative functions on $[0, T]$ with initial value $S_0 > 0$. As mentioned earlier we require $(S_t) \in \mathscr{P}$, corresponding to rather mild beliefs. Additionally, we assume that (S_t) is the price of an asset, which does not incur any cost to its holder.[3] To compensate for the degree of freedom of the underlying we require our model to treat more information consistently. This is done through a linear pricing operator defined on the set of traded assets. In particular, we assume that call options with maturity T of arbitrary strike K are liquidly traded in the market at prices $\{C(K), K \in \mathbb{R}_+\}$. Furthermore forward transactions are feasible and costless and constant payoffs are available at their trivial price. In line with Cox and Obłój (2011) define

$$\mathcal{X} := \left\{ A, (S_T - K)^+, (S_T - S_\rho)\mathbb{1}_{\rho \leq T} \middle| A \in \mathbb{R}, K \in \mathbb{R}_+, \rho \in \mathscr{I} \right\}, \tag{1}$$

where $\mathscr{I}$ is the set of all admissible trading dates. We denote by H_x the first hitting time of x, formally $H_x := \inf\{t > 0 | S_t = x\}$. As all sub- and superhedging strategies will be semi-static and involve trades in forward contracts upon hitting the barriers we assume:

$$\left\{0, H_{\underline{b}}, H_{\overline{b}}, \inf\{t < T \wedge H_{\overline{b}} | S_t = \underline{b}\}, \inf\{t < T \wedge H_{\underline{b}} | S_t = \overline{b}\}\right\} \subset \mathscr{I}$$

i.e. that such strategies are feasible. We further resort to digital call options, or

[3] For example (S_t) might be a forward price, or pay out dividends equal to the riskless rate or (S_t) could be the exchange rate between two economies with the same interest rate (Cox and Obłój (2011)).

cash-or-nothing call options at strike price $\underline{b}$ and $\bar{b}$. Define the set of all traded assets as:

$$\mathcal{X}_{\mathcal{D}} = \mathcal{X} \cup \left\{ \mathbb{1}_{\underline{b}<S_T}, \mathbb{1}_{\bar{b}\leq S_T} \right\}. \tag{2}$$

Let $Lin(\mathcal{X}_{\mathcal{D}})$ be the set of finite linear combinations of elements of $\mathcal{X}_{\mathcal{D}}$. On $Lin(\mathcal{X}_{\mathcal{D}})$ we define a pricing operator, which encodes market prices available today. Markets are assumed to be frictionless which is translated in the condition that the pricing operator is linear. We assume the existence of a normalised pricing operator, which treats all information, namely the prices of call options, forward contracts and digital call options, consistently. Linearity of $\mathcal{P}$ implies the classical Put-Call parity (B.1), which justifies the use of put options in our hedging strategies. Formally, we impose the following assumption.

Assumption 2.1 (Existence and consistency of a linear pricing operator)**.** *Assume there exists a linear pricing operator $\mathcal{P} : Lin(\mathcal{X}_{\mathcal{D}}) \to \mathbb{R}_+$ which satisfies:*

(i) $\mathcal{P}(1) = 1$,

(ii) $\mathcal{P}((S_T - K)^+) = C(K) \quad \forall K \in \mathbb{R}_+$,

(iii) $\mathcal{P}((S_T - S_\rho)\mathbb{1}_{\rho \leq T}) = 0 \quad \forall \rho \in \mathscr{I}$,

(iv) $\mathcal{P}(\mathbb{1}_{\{\underline{b}<S_T\}}) = -C'_+(\underline{b}) =: \underline{D}(\underline{b})$ *and* $\mathcal{P}(\mathbb{1}_{\{\bar{b}\leq S_T\}}) = -C'_-(\bar{b}) =: \overline{D}(\bar{b})$,

where C'_+ and C'_- denote the right and left derivative of the function $C(\cdot)$, respectively. We also allow for unlimited short-selling. Further, we introduce the following general and intuitive notion of arbitrage.

Definition 2.2 (Absence of model-independent arbitrage; Cox and Obłój (2011))**.** *A pricing operator $\mathcal{P}$ admits no model-independent arbitrage on $\mathcal{X}_{\mathcal{D}}$ it holds that: if* $\forall X \in Lin(\mathcal{X}_{\mathcal{D}}) X \geq 0 \;\Rightarrow\; \mathcal{P}(X) \geq 0$.

Any asset with a non-negative payoff must not have a price below zero. We assume that $\mathcal{P}$ admits no model independent arbitrage on $\mathcal{X}_\mathcal{D}$ and our goal is to ensure that this also holds on $\mathcal{X}_\mathcal{D} \cup \{\mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\bar{b}}\}$, i.e. that no arbitrage opportunities are created by including our exotic double barrier option.

The First Fundamental Theorem of Asset Pricing states that in the classical framework the pricing operator coincides with taking expectations with respect to an equivalent martingale measure, if and only if there are no arbitrage opportunities. Cox and Obłój (2011) derive a similar dichotomy for the robust framework and for the remainder of this text it is convenient to regard the absence of model-independent arbitrage and the existence of a suitable martingale measure as equivalent.

The Skorokhod Embedding Problem

The Problem

The Skorokhod embedding problem (henceforth SEP) was first formulated and solved by Anatoli Skorokhod (1961):

Problem 3.1 (Skorokhod embedding problem)**.**
Given a standard real-valued Brownian motion $(B_t)_{t\geq 0}$ with initial value B_0 and a probability measure μ on $\mathbb{R}$ such that $\int_\mathbb{R} x\mu(dx) = B_0$ and $\int_\mathbb{R} |x|\mu(dx) < \infty$, find a stopping time τ such that $B_\tau \sim \mu$ and $(B_{\tau\wedge t})_{t\geq 0}$ uniformly integrable.

For a brief discussion of the problem see appendix. Here we focus on how the problem and in particular its solution can be of use to construct model-independent price bounds. As discussed in section 2.2 the assumption of no model-independent arbitrage translates into (S_t) being a martingale under an equivalent measure. By a theorem of Monroe (1978) any semimartingale and, hence, any martingale is equivalent to a time-changed Brownian motion. It is worth emphasising that this does not mean that we are in any sense back in the Black Scholes modelling setup. It is rather the universal property of the Brownian motion that we use to rewrite any martingale process conveniently. More precisely, let τ be a solution to the SEP for $(B_t)_{t\geq 0}$, hence $B_\tau \sim \mu$, then $S_t := B_{\tau \wedge \frac{t}{T-t}}$ is a continuous martingale process on $[0,T]$ with terminal distribution μ. This is the first link to the problem of constructing model-independent price bounds.

Secondly, Breenden and Litzenberger's formula allows us to back out a unique risk neutral distribution of S_T which is consistent with the given call price function $C(\cdot)$ (Breeden and Litzenberger (1978); B.2). This determines a target measure μ which our price process should attain at maturity T. In other words, the measure μ in the SEP is derived from the call price function via the Breeden and Litzenberger formula.

Thus far, we have seen how to find a martingale on $[0,T]$ which takes the right distribution at maturity T. What we would like is that these martingale processes are in some sense extreme. Intuitively, we are looking for the process which is most likely to stay between the barriers and still takes the correct terminal distribution. This would correspond to the most favourable market model. Conversely, we also need to construct the solution process which is most likely to leave the interval $(\underline{b}, \overline{b})$ at some point. Fortunately, such processes have already been constructed and arise as special solutions to the SEP.

Extremal Solutions

The stopping time τ_P defined in (A.2) is due to Perkins (1986) and thus refered to as Perkins solution. Apart from solving the SEP it satisfies the following optimality condition.

Lemma 3.2 (Maximality of Perkins solution; Cox and Obłój (2011)).
For any $\underline{b} < B_0 < \overline{b}$ and for any stopping time τ solving the SEP for μ:

$$\mathbb{P}\Big(\underline{b} < \underline{B}_\tau \text{ and } \overline{B}_\tau < \overline{b}\Big) \leq \mathbb{P}\Big(\underline{b} < \underline{B}_{\tau_P} \text{ and } \overline{B}_{\tau_P} < \overline{b}\Big)$$

where τ_P denotes the Perkins solution defined in A.2.

Lemma 3.2 implies that $S_t^P := B_{\tau_P \wedge \frac{t}{T-t}}$ is the continuous martingale process on $[0, T]$, which satisfies $S_T \sim \mu$ and is most likely to stay within the interval $(\underline{b}, \overline{b})$. Fortunately, there is the analogous result for the mirror case due to Jacka (1988) and Cox (2004, 2008). The Tilted-Jacka solution τ_J of the SEP corresponds to the continuous martingale process with terminal distribution μ, which is most likely not to stay the interval $(\underline{b}, \overline{b})$. The following lemma formalises its optimality condition.

Lemma 3.3 (Minimality of Tilted-Jacka solution; Cox and Obłój (2011)).
For any $\underline{b} < B_0 < \overline{b}$ and for any stopping time τ solving the SEP for μ:

$$\mathbb{P}\Big(\underline{b} < \underline{B}_{\tau_J} \text{ and } \overline{B}_{\tau_J} < \overline{b}\Big) \leq \mathbb{P}\Big(\underline{b} < \underline{B}_\tau \text{ and } \overline{B}_\tau < \overline{b}\Big)$$

where τ_J denotes the Tilted-Jacka solution defined in A.3.

Robust Hedging and Pricing

After constructing suitable sub- and superhedging strategies in the first part of this section we establish model-independent price bounds for our exotic option (theorem 4.2).

Robust Hedging Strategies

In the following section we present the sub- and superhedging strategies originating from Cox and Obłój (2011). Recall that any portfolio must consist of assets in $\mathcal{X}_{\mathcal{D}}$ exclusively, i.e. plain vanilla call, put or binary options, forward contracts and cash holdings.

Subhedging Strategies

(i) The first subhedge is rather trivial as it simply consists of doing nothing. Consequently, it serves as a non-negativity constraint on the lower bound. Let formally $\underline{H}^1 \equiv 0$. Obviously, $\underline{H}^1 \leq \mathbb{1}_{\underline{b}<\underline{S}_T, \overline{S}_T<\overline{b}}$ holds.
(ii) The second subhedging strategy is a combination of cash-holdings, calls, puts and conditional forward contracts. More precisley, $\underline{H}^2$ corresponds to initially holding one unit of cash, selling $(\overline{b} - K_2)$ calls with strike K_2, selling $(K_1 - \underline{b})$ puts with strike K_1 and, conditionally on the stock price, hitting the barrier $\underline{b}$ or $\overline{b}$, taking a long or short position in $(\overline{b} - K_2)$ or $(K_1 - \underline{b})$ forward contracts, respectively. The portfolio makes use of the following weak inequality which is graphically illustrated in figure 01:

$$\begin{aligned} \mathbb{1}_{\underline{b}<\underline{S}_T, \overline{S}_T<\overline{b}} &\geq \underline{H}^2(K_1, K_2) \\ \textit{where } \underline{H}^2(K_1, K_2) &:= 1 - \frac{(S_T - K_2)^+}{\overline{b} - K_2} - \frac{(K_1 - S_T)^+}{K_1 - \underline{b}} \\ &\quad + \frac{S_T - \overline{b}}{\overline{b} - K_2} \mathbb{1}_{H_{\overline{b}} < H_{\underline{b}} \wedge T} - \frac{S_T - \underline{b}}{K_1 - \underline{b}} \mathbb{1}_{H_{\underline{b}} < H_{\overline{b}} \wedge T} \end{aligned}$$

with strikes $\underline{b} < K_1 < K_2 < \overline{b}$.

These two subhedging strategies are enough for our purpose and we proceed with superhedging strategies.

Superhedging Strategies

(i) Again, we include one rather obvious superhedge. Namely let $\overline{H}^1 := \mathbb{1}_{S_T \in (\underline{b}, \bar{b})}$ consist of a single digital option with payoff 1, if and only if the stock price at time T lies within the interval $(\underline{b}, \bar{b})$. The digital option clearly constitutes a superhedge for our exotic option.

The other superhedging strategies are obtained by treating our double no-touch barrier option as if it was only a simple digital barrier option with payoff $\mathbb{1}_{\underline{b}<\underline{S}_T}$ and $\mathbb{1}_{\overline{S}_T<\bar{b}}$.

(ii) We first focus on the second strateg where we treat our option as if it was a digital barrier option with payoff $\mathbb{1}_{\underline{b}<\underline{S}_T}$. The superhedging portfolio consists of initially buying a digital option paying 1 if and only if $\underline{b} < S_T$, buying $\frac{1}{K-\underline{b}}$ call options with strike K, selling the same amount of put options with strike $\underline{b}$ and, conditionally on hitting the lower level $\underline{b}$, selling again the same amount of forward contracts with forward price $\underline{b}$. Formally, the payoff of this portfolio corresponds to the right hand side of the following inequality which is graphically represented in figure 02:

$$\mathbb{1}_{\underline{b}<\underline{S}_T, \overline{S}_T<\bar{b}} \leq \overline{H}^2(K) := \mathbb{1}_{\underline{b}<S_T} + \frac{(S_T - K)^+}{K - \underline{b}} - \frac{(\underline{b} - S_T)^+}{K - \underline{b}} - \frac{S_T - \underline{b}}{K - \underline{b}} \mathbb{1}_{\underline{S}_T \leq \underline{b}}$$

with strike $\underline{b} < K$.

We now have all three superhedging strategies at hand and turn to the first intermediate result.

Model-Independent Price Bounds

An immediate consequence of the above is the following lemma. If we want to impose absence of model-independent arbitrage, the price of any sub- and superhedge can never exceed and never fall below that of the option, respectively.

Lemma 4.1 (Most expensive subhedge & cheapest superhedge of such kind; Cox and Obłój (2011)).
For $\underline{b} < S_0 < \overline{b}$. If $\mathcal{P}$ admits no model-independent arbitrage on $\mathcal{X} \cup \mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\overline{b}}$, then

$$\mathcal{P}(\mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\overline{b}}) \geq \sup_{K_1,K_2\in\mathbb{R}_+,\underline{b}<K_1<K_2<\overline{b}} \left\{\mathcal{P}\big(\underline{H}^1\big), \mathcal{P}\big(\underline{H}^2(K_1,K_2)\big)\right\} \tag{3}$$

$$\mathcal{P}(\mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\overline{b}}) \leq \inf_{K_3,K_4\in\mathbb{R}_+,\underline{b}<K_3,K_4<\overline{b}} \left\{\mathcal{P}\big(\overline{H}^1\big), \mathcal{P}\big(\overline{H}^2(K_3)\big), \mathcal{P}\big(\overline{H}^3(K_4)\big)\right\} \tag{4}$$

Lemma 4.1 provides necessary conditions for the absence of arbitrage opportunities. Note that it is not obvious why it should suffice to restrict attention to the presented hedging strategies. Nevertheless, the following theorem states that this is indeed the case. Below we construct possible price processes which attain the bounds given above. For this we resort to the Skorokhod embedding problem and, in particular, to the optimal solutions introduced in section 3.2.

Theorem 4.2 (Robust price bounds; Cox and Obłój (2011)).
For $0 < \underline{b} < S_0 < \overline{b}$. Under assumption 2.1 if $\mathcal{P}$ admits no model-independent arbitrage on $\mathcal{X}_{\mathcal{D}}$ and $C(K) \to 0$ for $K \to \infty$, then the following statements are equivalent:

1. *$\mathcal{P}$ admits no model-independent arbitrage on $\mathcal{X}_{\mathcal{D}} \cup \{\mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\overline{b}}\}$*

2. *(3) and (4) in Lemma 4.1 hold.*

3. *with μ defined via the Breeden and Litzenberger formula (B.2),*

$$\max\left\{0, 1 - \frac{C(\Psi_\mu^{-1}(\overline{b}))}{\overline{b} - \Psi_\mu^{-1}(\overline{b})} - \frac{P(\Theta_\mu^{-1}(\underline{b}))}{\Theta_\mu^{-1}(\underline{b}) - \underline{b}}\right\} \leq \mathcal{P}(\mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\overline{b}}) \tag{5}$$
$$\leq \min\left\{\underline{D}(\underline{b}) - \overline{D}(\overline{b}), \underline{D}(\underline{b}) + \frac{C(\gamma_\mu^-(\underline{b})) - P(\underline{b})}{\gamma_\mu^-(\underline{b}) - \underline{b}}, 1 - \overline{D}(\overline{b}) + \frac{P(\gamma_\mu^+(\overline{b})) - C(\overline{b})}{\overline{b} - \gamma_\mu^+(\overline{b})}\right\}$$

where γ^+, γ^-, Ψ_μ and Θ_μ are defined in A.2 and A.3. Furthermore, the upper bound in (5) is attained for the market model $S_t^P := B_{\tau_P \wedge \frac{t}{T-t}}$

whereas the lower bound is attained for $S^J_t := B_{\tau_J \wedge \frac{t}{T-t}}$ *where* τ_P *is the Perkins solution (A.2),* τ_J *the Tilted-Jacka solution (A.3) and* $(B_t)_{t\geq 0}$ *a standard Brownian motion with* $B_0 = S_0$*.*

Its straightforward to see that *1.* implies *2.* which in turn implies *3.*. The difficult part is to prove that *3.* indeed implies *1.* We refer to Cox and Obłój (2011) for details. In essence, theorem 4.2 states that we can find arbitrage opportunities using assets in $\mathcal{X}_{\mathcal{D}} \cup \{\mathbb{1}_{\underline{b}<\underline{S}_T, \overline{S}_T<\bar{b}}\}$, if and only if the price of our digital double no-touch barrier option lies within the boundaries. As alluded to earlier the extremal solutions to the SEP, namely the Perkins and the Tilted-Jacka solution, correspond to the price processes which attain the bounds constructed through sub- and superhedging portfolios.

Applications

We now focus on how to practically utilise the theoretical results of the previous section. Unfortunately, the robust construction is useless for pricing or finding arbitrage opportunities. Almost all observed bid-ask spreads are too narrow and the model-independent price bounds too wide to yield riskless profits.
Yet, not all of the above is worthless for practitioners. The first part of this section introduces the necessary extensions of our theory to more realistic settings. The second part briefly summarises an assessment by Ulmer (2010) of performance of the robust hedging strategies compared to classical delta/vega hedging techniques. Interestingly, the new robust approach turns out to be quite competitive in terms of hedging.

Finite Number of Call Options and Non-zero Interest Rate

At the very core of determining model-independent bounds for option prices stands the Breeden and Litzenberger formula (B.2). It requires a given call price function, i.e. a call price for any possible strike, to uniquely back out a target distribution. For any application however, there is always only a finite number of call prices available. One way to apply the robust construction would be to fit a function to all given points. It turns out that extending the given information to a function is, indeed, a suitable idea.

Throughout the text we made the assumption that the underlying asset incurs zero cost of carry. Relaxing this assumption renders the problem considerably harder since the boundaries become time-dependent. Consider the constant, continuously compounded interest rate example from Ulmer (2010). Let (S_t) be a spot price on a foreign exchange rate and denote by r_d and r_f the domestic and foreign interest rate, respectively. The natural candidate for our martingale price process is the forward price process $S^* = (S_t^* := S_t \cdot e^{-(r_d - r_f)(T-t)})_{t\in[0,T]}$ which is obviously a martingale. Applying the SEP technique to (S^*) is not possible because of the time-depending barriers. If $\underline{b}$ and $\overline{b}$ are the constant barriers for the spot exchange rate, we see that

$$\mathbb{1}_{\underline{b}<\underline{S}_T,\overline{S}_T<\overline{b}} = \mathbb{1}_{\forall u,v\leq T:\underline{b}\cdot e^{-(r_d-r_f)(T-u)}<S_u^*,S_v^*<\overline{b}\cdot e^{-(r_d-r_f)(T-v)}}. \tag{6}$$

There are at least two ways of dealing with the problem. The first one is a worst case consideration originating from Cox and Obłój (2011), the second is a hybrid approach pursued by Ulmer (2010).

The latter simply decomposes (6) into two different payoffs

$$\mathbb{1}_{\forall u,v\leq T:\underline{b}\cdot e^{-(r_d-r_f)(T-u)}<S_u^*,S_v^*<\overline{b}\cdot e^{-(r_d-r_f)(T-v)}} = \mathbb{1}_{\underline{b}<\underline{S^*}_T,\overline{S^*}_T<\overline{b}} + R, \tag{7}$$

where R denotes the residual term, which is hedged using standard delta/vega hedging. Thus the barriers' time-dependence is essentially neglected and corrected for by an additional error term. Obviously, the efficiency of this hybrid approach depends on the impact and size of the residual term and the associated hedging error and costs.

Performance Comparison of Hedging Strategies

This section summarises a detailed assessment of Ulmer (2010) analysing the performance of classical hedging strategies compared to the presented robust approach. In general standard hedging strategies, such as delta/vega hedging, are subject to several severe difficulties, such as transaction costs and model misspecification. Contrarily, the robust approach does not only avoid the risk of model misspecification, but also has very little exposure to transaction costs, because all hedges are semi-static.

In the following, we focus on one particular part of the study conducted by Ulmer (2010) in which the author compares the performance of a modified delta/vega hedging strategy and the presented robust hedging approach for a one month double no-touch digital barrier option in foreign exchange spot rate markets. He considers a low and a high transaction cost scenario for both the EUR/USD and AUD/USD foreign exchange spot rates and compares four different set of barriers for the exotic option. For the conducted Monte Carlo simulation both exchange rates are assumed to evolve according to a Variance Gamma process with CIR stochastic clock (VGSV) with parameters calibrated from market data.

The classical standard delta/vega hedging strategy is adjusted to a world with transaction cost using a rather simple but widespread practical approach, that is to not rebalance the portfolio if the delta and vega is within a certain bandwidth away from the Black-Scholes target values. Targeting the theoretical Black-Scholes values obviously created a hedging error due to model misspecification. Both hedging strategies, modified delta/vega and robust SEP, are conducted in two different frequencies of monitoring and rebalancing.

The performance of all hedging strategies was assessed on the basis of several risk- and performance measures such as: mean hedging error, standard deviation, Value-at-Risk, Conditional Value-at-Risk and maximum loss. To sum up, the SEP method outperforms adjusted delta/vega hedging strategies on a risk adjusted basis in spot markets with no risk-neutral drift and high transaction costs. Even in markets with non-zero cost of carry the SEP method might outperform classical hedging strategies if the risk-neutral drift is not too high as to become the main source of hedging errors. The shorter the maturity of the option and the higher the volatility of the underlying, the less likely this problem occurs and the more promising the hybrid SEP approach.

Conclusion

This thesis attempts to introduce the Skorokhod embedding technique to a broader audience by providing a concise overview of the approach covering its mathematical foundation, illustrating the construction by means of an example, sketching a suitable translation to a more realistic setup and summarising an assessment of its performance compared to existing hedging strategies.

The new approach proves to be useful both from a theoretical and practical point of view. Theoretically, the Skorokhod embedding technique offers an alternative framework for dealing with model uncertainty in a coherent and exact abstract manner. Practically, even though the obtained bounds are not useful for pricing, they provide competitive hedging

References

Black, F., and M. Scholes (1973): "The pricing of options and corporate liabilities," *The journal of political economy*, pp. 637–654.

Blumenthal, R. M., and R. K. Getoor (1968): "Markov processes and potential theory," *Pure and Applied Mathematics*, 29.

Breeden, D. T., and R. H. Litzenberger (1978): "Prices of state-contingent claims implicit in option prices," *Journal of business*, pp. 621–651.

Cox, A. (2008): "Extending Chacon-Walsh: Minimality and generalised starting distributions," in *Séminaire de probabilités XLI*, pp. 233–264. Springer.

Cox, A. M., and J. Obłój (2011): "Robust pricing and hedging of double no-touch options," *Finance and Stochastics*, 15(3), 573–605.

Cox, A. M. G. (2004): "Skorokhod Embeddings: Non-Centred Target Distributions, Diffusions and Minimality," Ph.D. thesis, University of Bath.

Hobson, D. (2011): "The Skorokhod embedding problem and model-independent bounds for option prices," in *Paris-Princeton Lectures on Mathematical Finance 2010*, pp. 267–318. Springer.

Hobson, D. G. (1998): "Robust hedging of the lookback option," *Finance and Stochastics*, 2(4), 329–347.

Jacka, S. (1988): "Doob's inequalities revisited: A maximal H1-embedding," *Stochastic processes and their applications*, 29(2), 281–290.

Knight, F. B. (1962): "On the random walk and Brownian motion," *Transactions of the American Mathematical Society*, 103(2), 218–228.

Merton, R. C. (1969): "Lifetime portfolio selection under uncertainty: The continuous-time case," *The review of Economics and Statistics*, 51(3), 247–257.

Monroe, I. (1978): "Processes that can be embedded in Brownian motion," *The Annals of Probability*, pp. 42–56.

Obłój, J. (2004): "The Skorokhod embedding problem and its offspring," *Probability Surveys*, 1, 321–392.

——— (2013): *On some aspects of the Skorokhod Embedding Problem and its applications in Mathematical Finance*Preliminary version of lecture notes edn.

Perkins, E. (1986): "The Cereteli-Davis solution to the H1-embedding problem and an optimal embedding in Brownian motion," in *Seminar on stochastic processes, 1985*, pp. 172–223. Springer.

Shreve, S. (2008): “Don’t blame the quants,” *Forbes Magazine*, August 2008.

Skorokhod, A. V. (1961): “Issledovaniya po teorii sluchainykh processov in Russian; English translation 1983: Studies in the theory of random processes,” .

Taleb, N. N. (2010): *The Black Swan - The Impact of the Highly Improbable Fragility*. Random House Digital, Inc.

Ulmer, F. (2010): “Performance of robust model-free hedging via Skorokhod embeddings of digital double barrier options,” *Mathematical Finance*.

Appendix

A - Skorokhod Embedding Problem

The Problem

The additional requirement of uniformly integrability of the stopped process $B_{\tau\wedge t}$ is necessary since there would exist a trivial solution otherwise.

Example A.1 (Necessity of uniformly integrability).
Let (B_t) be a standard real-valued Brownian motion with initial value B_0 and μ a probability measure on $\mathbb{R}$ such that $\int_{\mathbb{R}} x\mu(dx) = B_0$ and $\int_{\mathbb{R}} |x|\mu(dx) < \infty$. Let further Φ denote the cumulative distribution function of the standard normal distribution with mean B_0 and F_μ^{-1} its right-continuous inverse. Then

$$\tau_\Phi := \inf\{t \geq 1 \mid B_t = F_\mu^{-1}(\Phi(B_1))\} \tag{8}$$

solves the SEP for (B_t) and μ, since $\forall a \in \mathbb{R}$

$$\begin{aligned}\mathbb{P}[B_{\tau_\Phi} \leq a] &= \mathbb{P}[F_\mu^{-1}(\Phi(B_1)) \leq a] = \mathbb{P}[\Phi(B_1)) \leq F_\mu(a)] \\ &= \mathbb{P}[B_1 \leq \Phi^{-1}(F_\mu(a))] \overset{B_1\sim\mathcal{N}(B_0,1)}{=} \Phi(\Phi^{-1}(F_\mu(a))) \\ &= F_\mu(a) = \mu((-\infty, a])\end{aligned}$$

However, $\mathbb{E}[\tau_\Phi] = \infty$ *because* $\mathbb{P}[\tau_\Phi = \infty] > 0$ *which motivates the additional requirement.*

It is worth noticing that uniformly integrability of $(B_{\tau\wedge t})$ implies that μ must have mean B_0 since

$$\begin{aligned}(B_{\tau\wedge t}) \text{ uniformly integrabe} \Longrightarrow \lim_{t\to\infty} B_{\tau\wedge t} = B_\tau \text{ exists in } L^1 \text{ and}\\ B_{\tau\wedge t} = \mathbb{E}[B_\tau \mid \mathcal{F}_{\tau\wedge t}] \text{ almost surely}\end{aligned} \tag{9}$$

Where (9) for $t = 0$ implies $\mathbb{E}[B_\tau] = B_0$ almost surely and, hence, μ has mean B_0.

Perkins Solution

Proposition A.2 (Perkins solution; Perkins (1986)). *Let*

$$\gamma_\mu^+(x) := \sup\left\{y < B_0 : \int_{(0,y)\cup(x,\infty)} (w-x)\mu(dw) \geq 0\right\} \text{ for } x > B_0$$
$$\gamma_\mu^-(y) := \inf\left\{x > B_0 : \int_{(0,y)\cup(x,\infty)} (w-y)\mu(dw) \leq 0\right\} \text{ for } y < B_0.$$

A solution to the SEP is given by the Perkins solution, which is defined as

$$\tau_P := \inf\left\{t : B_t \notin \left(\gamma_\mu^+(\overline{B}_t), \gamma_\mu^-(\underline{B}_t)\right)\right\}.$$

Furthermore, the Perkins solution has the property that for any other solution τ we have:

$$\mathbb{P}\Big(\underline{b} < \underline{B}_\tau\Big) \leq \mathbb{P}\Big(\underline{b} < \underline{B}_{\tau_P}\Big) \text{ and } \mathbb{P}\Big(\overline{b} < \overline{B}_\tau\Big) \leq \mathbb{P}\Big(\overline{B}_{\tau_P} < \overline{b}\Big)$$

Tilted-Jacka Solution

Proposition A.3 (Tilted-Jacka Solution; Cox (2004), Cox (2008), Jacka (1988)). *Let*

$$\Psi_\mu(K) := \begin{cases} \frac{1}{\mu([K,\infty))}\int_{[K,\infty)} x\mu(dx) & \text{if } \mu([K,\infty)) > 0 \\ \infty & \text{if } \mu([K,\infty)) \leq 0 \end{cases}$$
$$\Theta_\mu(K) := \begin{cases} \frac{1}{\mu((-\infty,K])}\int_{(-\infty,K]} x\mu(dx) & \text{if } \mu((-\infty,K]) > 0 \\ -\infty & \text{if } \mu((-\infty,K]) \leq 0 \end{cases}$$

A solution to the SEP is given by the Tilted-Jacka solution which is defined for $K \in (0, \infty)$ as

$$\begin{aligned}
\tau_1 &= \inf \left\{ t \geq 0 : B_t \notin \left(\Theta_\mu(K), \Psi_\mu(K) \right) \right\}, \\
\tau_\Psi &= \inf \left\{ t \geq \tau_1 : \Psi_\mu(B_t) \leq \overline{B}_t \right\}, \\
\tau_\Theta &= \inf \left\{ t \geq \tau_1 : \Theta_\mu(B_t) \geq \underline{B}_t \right\}, \\
\tau_J(K) &= \tau_\Psi \mathbb{1}_{\{\tau_1 = \Psi_\mu(K)\}} + \tau_\Theta \mathbb{1}_{\{\tau_1 = \Theta_\mu(K)\}}
\end{aligned}$$

Furthermore, the Tilted-Jacka solution has the following property.
For $K \in (0, \infty)$, if $\underline{b} < \Theta_\mu(K)$ and $\Psi_\mu(K) < \overline{b}$, then for any solution τ we have

$$\mathbb{P}\left(\underline{b} < \underline{B}_{\tau_J(K)} \right) \leq \mathbb{P}\left(\underline{b} < \underline{B}_\tau \right) \text{ and } \mathbb{P}\left(\overline{B}_{\tau_J(K)} < \overline{b} \right) \leq \mathbb{P}\left(\overline{B}_\tau < \overline{b} \right)$$

We now briefly describe the choice of a suitable K for the construction of the Tilted-Jacka solution. Take an arbitrary function f increasing for $x > B_0$ and decreasing for $x < B_0$. As $\Theta_\mu(\cdot)$ and $\Psi_\mu(\cdot)$ are both increasing, there exists a $K \in \mathbb{R}$ such that $f(\Theta_\mu(K)) = f(\Psi_\mu(K))$. Cox Cox (2004) proves that the Tilted-Jacka solution constructed using such K indeed maximises $\mathbb{P}(\sup_{t \leq \tau} f(B_t) \geq z)\ \forall z$. For the consideration of our double no-touch barrier options the function of interest is $f(x) = \mathbb{1}_{x \notin (\underline{b}, \overline{b})}$. Note that the choice of K is generally not unique, however, for our purpose it suffice to notice that a suitable K exists and we implicitly assume that an appropriate K has been chosen when writing $\tau_J(K)$ for the Tilted-Jacka solution.

B - Put-Call Parity

Proposition B.1 (Put-Call parity).

Under assumption 2.1, it holds that

$$\mathcal{P}((K - S_T)^+) = K - S_0 + C(K).$$

Breeden and Litzenberger Formula

Lemma B.2 (Distribution implied by option prices; Breeden and Litzenberger Breeden and Litzenberger (1978)).

Fix $T \in (0, \infty)$*. Suppose call prices with maturity* T *are known for every strike* $K \in (0, \infty)$*. Then assuming call prices are calculated as the discounted expected payoff, so that*

$$C(K) = \mathbb{E}_\mu[e^{-rT}(S_T - K)^+]$$

we have

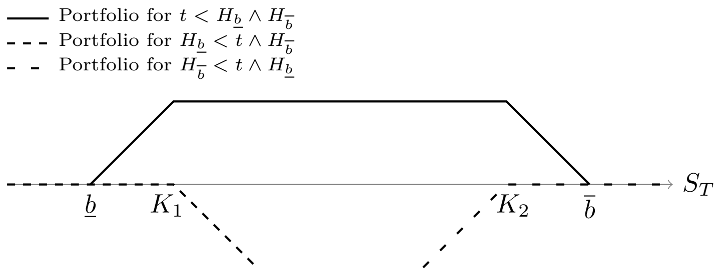
$$\mu(S_T > K) = e^{rT}\left|\frac{\partial}{\partial K}C(K)\right|, \tag{10}$$

and, provided $C(\cdot)$ *is twice-differentiable in* K*,*

$$\mu(S_T \in dK) = e^{rT}\frac{\partial^2}{\partial K^2}C(K). \tag{11}$$

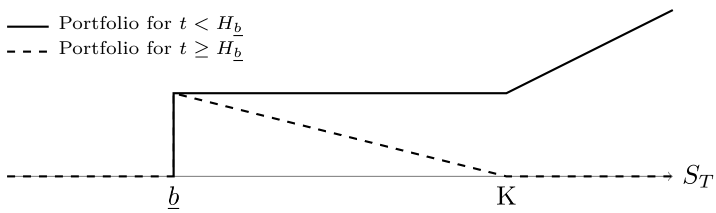
If the law of S_T *and* μ *has atoms, then* $\mu(S_T > K)$ *is given by the right derivative in (10) and* $\mu(S_T \geq K)$ *by the left derivative. In this case (11) must be understood in a distributional sense.*

Figures

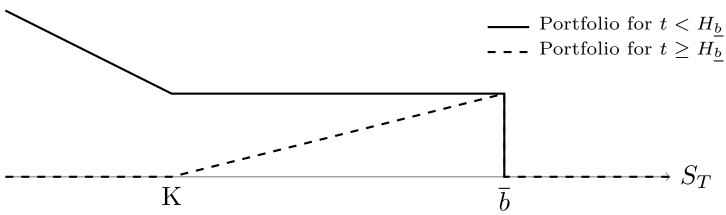


(a) The value of the portfolio $\underline{H}^2$

Figure 01: Payoff diagram of subhedging strategy $\underline{H}^2$



(a) The value of the portfolio $\overline{H}^2$



(b) The value of the portfolio $\overline{H}^3$

Figure 02: Payoff diagrams of superhedging strategies