That's how we roll: an experiment on rollover risk

Ciril Bosch-Rosa *

* Technische Universität Berlin, Germany

BERLIN

ECONOMIC RISK

SFB 649

That's how we roll: an experiment on rollover risk

Ciril Bosch-Rosa*

May 23, 2014

Abstract

There is consensus that the recent financial crisis revolved around a crash of the short-term credit market. Yet there is no agreement around the necessary policies to prevent another credit freeze. In this experiment we test the effects that contract length (i.e. maturity mismatch) has on the market-wide supply of short-term credit. Our main result is that, while credit markets with shorter maturities are less prone to freezes, the optimal policy should be state-dependent, favoring long contracts and lower maturity mismatch when the economy is in good shape, and allowing for short-term contracts when the economy is in a recession. We also report the possibility of credit runs on firms with strong fundamentals, something that cannot be observed in the canonical static models of financial panics. Finally, we show that our experimental design produces rich learning dynamics, with a text-book bubble and crash pattern in the market for short-term credit.

Keywords: Experiment, Financial Crisis, Continuous Time, ABCP

JEL Codes: C92, C91, G01, GO2, G21

1 Introduction

While the literature agrees on placing a run on short-term credit at the center of the recent financial crisis (e.g. Brunnermeier (2009), Krishnamurthy (2010)), there is much less consensus on how to prevent another panic from happening. Brunnermeier et al. (2009) suggest that extending the maturity of short-term credits might help stabilize the market, while Farhi and Tirole (2012) advocate for putting a cap on the total amount of short-term debt issued by financial firms. In this paper we use experimental tools to test the effects that reducing maturity mismatch has on the market for short-term credit. To be more precise, we use experimental tools to investigate the effects of different maturity lengths on a market for Asset Backed Commercial Paper (ABCP), which is a specific type of short-term credit in which, if the issuing firm does not fulfill its promises, the holder of the ABCP can seize the posted collateral.

In order to study this type of markets we design a continuous time experiment with staggered credit maturities inspired on the model by He and Xiong (2012). Our results show that while on average markets with shorter credit maturities (i.e. larger maturity mismatch) have a lower probability of freezing, a more detailed analysis of the data indicates that the optimal policy should be state-dependent, favoring longer contracts when the economy is in good shape, and allowing for shorter ones during a recession. We also observe for the first time in a controlled lab experiment, a significant number of runs on firms with solid fundamentals[1], a direct result of the staggered maturities in the experimental design, and one of the most policy relevant predictions in He and Xiong (2012). Finally, we show that in our setup there are rich experimental dynamics with the formation of credit booms and busts across all sessions.

1.1 Why Run an Experiment on ABCP?

ABCP has been pointed out as the necessary transmitter of the housing bubble into the financial system, so much that while short-term credit was not a problem *per se*, ABCP played a central role in the financial meltdown and credit freeze of 2007 (Brunnermeier (2009)). The argument is that ABCP (usually supported by structured subprime mortgages) took over the more "traditional" credit market in the years before the crisis, and by virtue of being cheap and unregulated, it exposed the market to a credit bubble, and to "excessive mismatch in asset-liability maturities". In Figure 1 (borrowed from Brunnermeier (2009)) we see how the market for ABCP almost doubles in size from 2005 to 2007 (the final years of the housing bubble), to crash and drop from an outstanding $1,200 billion to $750 billion in just six months in 2007. Yet, what is most interesting about this graph is not the increase in ABCP, but rather how unsecured instruments were only slightly affected at the peak of the 2007 crisis. This

---

[1] We consider firms with solid fundamentals to be those that can post enough collateral to pay back all of its creditors.

Figure 1: Outstanding Asset backed and Unsecured Commercial Paper Comparison
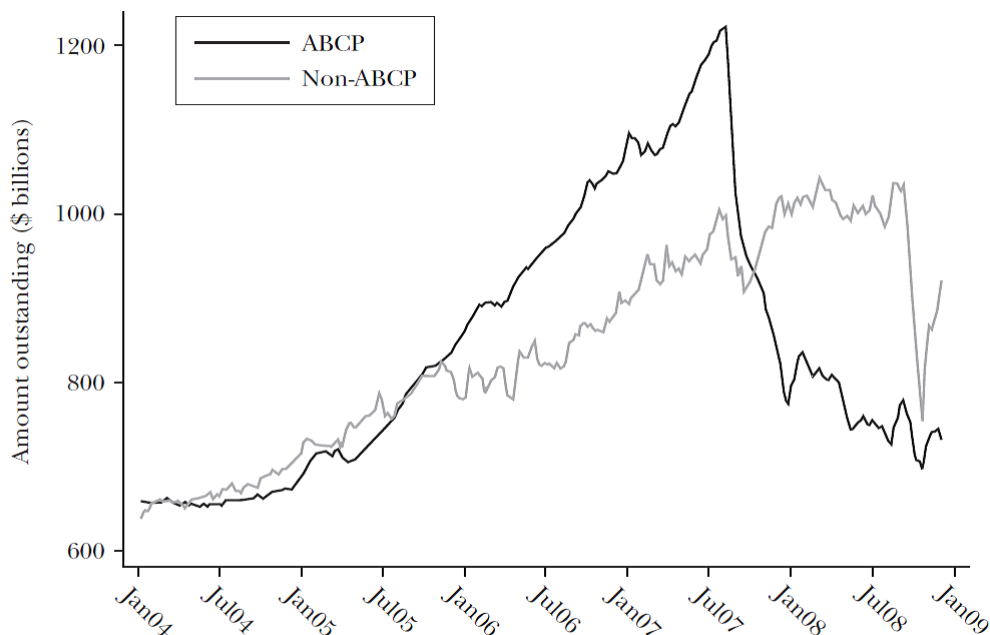
suggests that the problem revolved more around a change in the perceived value of collateral than around the use of short-term credit. In fact, even previous to the crash of the market, a literature had already developed around the idea of regulating collateral (e.g., Geanakoplos (2009, 1996)).

In a more specific analysis, Shin (2009) looks at the particular case of Northern Rock, and explains the shift in paradigm that modern financial markets have brought to our understanding of "bank runs". As he puts it, while we all remember the lines forming at the doors of Northern Rock, the real storm had occurred weeks before, when non-depository creditors (mostly of ABCP) decided not to roll over their loans to the bank.[2] The important question, according to the author is thus "not so much why banks depositors are so prone to running, but instead why the plentiful short-term funding (...) suddenly dried up".

Additionally, it appears that the Federal Reserve agrees with the necessity to control short-term credit, pointing it as one of the main reason for the recent crisis (Bernanke (2008, 2009a,b)), and modifying its policy to a "credit-easing strategy rather than a quantitative-easing approach" (Bernanke (2009a)). In fact, in a 2008 speech[3] , Bernanke expressed this shift in the paradigm of financial crises:

---

[2] As Shin (2009) puts it "The depositor run, although dramatic, was an event in the aftermath of the liquidity crisis at Northern Rock".

[3] Given at the Federal Reserve Bank of Atlanta Financial Markets Conference on May 13 2008.

3

"Bagehot defined a financial crisis largely in terms of a banking panic –that is, a situation in which depositors rapidly and simultaneously attempt to withdraw funds from their bank accounts. In the 19th century, such panics were a lethal threat for banks that were financing long-term loans with demand deposits that could be called at any time. In modern financial systems, the combination of effective banking supervision and deposit insurance has substantially reduced the threat of retail deposit runs. *Nonetheless, recent events demonstrate that liquidity risks are always present for institutions –banks and nonbanks alike–that finance illiquid assets with short-term liabilities.*" (Emphasis added)

2 Our Experiment in the Context of the Experimental Literature

No experimental literature exists on the topic we are covering, so our references consist on two strands of experimental research which are relatively close to our experimental design. The first one corresponds to continuous-time experiments, the second to "timing experiments", with a special emphasis on the experimental bank-runs literature. Continuous-time experiments, started a few years ago with Cheung and Friedman (2009) and Brunnermeier and Morgan (2010) (whose working papers appeared around 2003/04), but it has not been until recently that this experimental technique has taken off with Oprea et al. (2009) and Anderson et al. (2010) looking into strategic investment decisions, Oprea et al. (2011) studying the evolutionary equilibrium of the hawk and dove game, and Friedman and Oprea (2012) experimenting with the effects of response delay in a repeated prisoners dilemma game. While none of these papers directly address any of the questions of our paper, they are a good methodological reference for the design of our experiment.

The other relevant strand of literature for our paper deals with experimental bank runs. To our knowledge, the first paper on this topic is Madies (2006), which is based on the theoretical model of Diamond and Dybvig (1983) (DD henceforth). The results show that (partial) deposit insurance cannot avoid bank runs, and that the more experienced subjects are, the more often runs are observed. Garratt and Keister (2009) also test the DD setup but turn it into a repeated game by giving subjects the opportunity to exit several times per round. Schotter and Yorulmazer (2009) also adopt this technique. Both papers find that not only more experienced subjects are more prone to runs, but that the more opportunities to run within each round, the more likely runs are. Surprisingly, Garratt and Keister (2009) report having to exogenously force some subjects to exit, else no panics would occur. More recently Arifovic et al. (2013) look at how bank runs can be understood as a pure coordination problem. Also Brown et al. (2012) and Chakravarty et al. (2012) have looked at bank run contagion across independent banks. Finally, Klos and Sträter (2013) approach bank runs from a Global Games perspective.

In summary, while there exists some experimental literature studying banking panics, most of it is based on models of "classic" bank runs with simultaneous withdrawals and discrete timing, with no papers addressing the intricacies of modern financial markets, in particular the markets for short-term credit. For a summary of the existing literature on experimental finance see Heinemann (2012) or Dufwenberg (2012).

3   Theoretical Benchmark

Our experiment is inspired on the continuous time model by He and Xiong (2012) (HX henceforth). In it, a firm finances its long-term investment by issuing short-term credit to a continuum of creditors. Without loss of generality we will assume this credit to be of \$1. The value of the firm[4] follows a geometric Brownian motion and is perfectly observable by all agents. The Brownian motion can be written as:

$$\frac{dy}{y} = \mu dt + \sigma dZ \tag{1}$$

Where y_t is the value of the firm, μ is the drift, σ the volatility, and Z the standard Brownian motion.

Each creditor's debt matures with the arrival of an independent Poisson shock of intensity $\kappa > 0$, creating a uniform distribution of the maturities[5] , with all contracts having an expected duration of $1/\kappa$ at any point in time. This random-maturities system is a simplifying assumption akin to Calvo pricing (Calvo (1983)), and avoids agents having to keep track of all other maturities when making the rollover decision, while still capturing all of the first-order effects of other maturing contracts.

If within the time interval $[t, t+dt]$ enough creditors decide not to rollover their credit, then the firm draws from its cash reserves[6] (ϑ) and survives, on average, an extra $1/\vartheta\kappa$. Once the firm runs out of reserves it goes bankrupt and liquidates its assets at a discount value $\alpha < 1$, so the value of the asset is $\alpha F(y_t)$, where $F(y_t)$ is the present discounted value of the firm.

As payoffs, agents receive a stream of interests r until $\tau = \min(t_m, t_b, t_d)$ which is the earliest of three possible events. The first event (t_m) is the maturing of the long-term investment of the firm, in which case the agent gets back $\min(1, y_{t_m})$ and the firm ceases to exist. That is, the firm pays back the full principal of the credit if it can, or whatever it can pay back (never more than the original \$1 credit). The second possibility (t_b) is a bankruptcy of the firm, in which case the creditor gets back $min[1, \alpha F(y_t)]$. Finally, the short-term credit can mature (t_d), at which point the creditor will decide to rollover his credit if the continuation value $V(y_{t_d}; y*)$ is higher than getting

---

[4]We will assume that the firm's only investment is on the long-term asset. Therefore, the value of the long-term asset is the total value of the firm.

[5]Most firms spread out their maturities to avoid having large liquidity needs on a specific date.

[6]He and Xiong (2012) describe ϑ as unreliable credit lines that the firm may tap, which is why the extra time is a function of the contract length. We believe that describing ϑ as cash reserves is more intuitive for our experimental purposes.

his credit back ($1), where y_{t_d} is the value of the firm at the maturity point t_d, and $y*$ is the stopping threshold of other agents. The continuation value is thus written as:

$$
\begin{aligned}
V(y_t; y*) &= E_t \left\{ \int_t^\tau e^{(-\rho(s-t))} r \, ds + e^{(-\rho(\tau-t))} \left[\min(1, y_t) \mathbf{1}_{\tau=t_m} \right] \right\} + \\
&\quad \min\left(1, \alpha F\left(y_t\right)\right) \mathbf{1}_{\tau=t_m} + \max_{\text{rollover or run}} \left\{ 0, 1 - V\left(y_t; y*\right) \right\} \mathbf{1}_{\tau=t_d}
\end{aligned}
\tag{2}
$$

In equation (2) ρ is the discount value of the agent, and $\mathbf{1}_{\{.\}}$ is an indicator function which takes value 1 whenever the subscript is true, zero otherwise.

By evaluating the change in value of the continuation value (2) over a small time interval $[t, t+dt]$ the Hamilton-Jacobi-Bellman equation can be written as:

$$
\begin{aligned}
\rho V(y_t; y*) &= \mu y_t V_y + \frac{\sigma^2}{2} y_t^2 V_{yy} + r + \phi \left[\min\left(1, y_t\right) - V\left(y_t; y*\right) \right] + \\
&\quad \kappa \delta \mathbf{1}_{\{\mathbf{y_t} < \mathbf{y*}\}} + \delta \max_{\text{rollover or run}} \left\{ 0, 1 - V\left(y_t; y*\right) \right\}
\end{aligned}
\tag{3}
$$

The left hand side represents the required return to the creditor, the first two terms in the right hand-side evaluate the fluctuation in the value of the firm. The equation also contains the continuation values of each of the three outcomes (long-term maturity, bankruptcy, short-term maturity) weighted by the probability of each one.

Finally, from equation (3) HX show that agents will rollover the credit *if and only if* $V(y_t; y*) > 1$, that is, if the continuation value is greater than getting back the principal of the credit and "walking away". This results in a unique *symmetric* equilibrium determined by the condition $V(y*; y*) = 1$, where no subject rolls over the short-term loan to a firm whose value is below $y*$, and always does so for values above $y*$.

What we should understand from this model is that, unlike global games models, subjects do not get a noisy signal, but a precise one. The strategic uncertainty comes from the asynchronous structure of the maturities, and the frequent change in value of the firm. It is precisely from these two key elements that agents can coordinate on a unique equilibrium, and this is why we can have results that would never happen in classic static models.

4 Experimental Implementation

4.1 Basic Design

Our experiment considers groups of subjects where each member of the group provides a $1 short-term credit

to a firm which has made a time-varying long-term investment. Each group is composed of 4 subjects (which is a number close to the 5 individuals in the groups of Garratt and Keister (2009)), and the composition of the groups remains invariable during all of the 60 rounds that a session lasts. All subjects are informed of the size of their group and of its unchanged composition.

The time units of our experiment are ticks. Following Anderson et al. (2010), each tick is 1/5 of a second (i.e., 200 milliseconds). Each of the 60 rounds has a random end which is governed by a Poisson process, and has an expected length of 150 ticks (30 seconds), at which point the long-term investment matures and the firm ceases to exist. In HX, the value of the long-term investment (y_t) follows a geometric Brownian motion. Given the discrete nature of computer internal clocks, we will need to discretize this Brownian motion, and for that purpose we use the procedure described in Anderson et al. (2010).

In each of the 60 rounds we will ask subjects to make one and only one decision, namely, whether or not to continue rolling over their credit to the firm[7]. If at any time 2 subjects decide not to rollover their credit, then the firm will continue to run for a fixed θ of ticks before it goes bankrupt and has to liquidate its assets at a fire-sale value. This "extra time" θ is a linear function of the duration of short-term contracts and can be interpreted as the cash reserves of the firm[8]. The decision to choose 2 out of 4 subjects as the threshold for bankruptcy is again inspired by Garratt and Keister (2009), where the bank goes bankrupt if 3 out of 5 subjects decide to run.

The payoffs for each round will depend on the value of the firm at time $t, (y_t)$, and the decisions made by each subject in the group. To be precise, each round's individual payoffs will accrue from two different sources:

1. *Flow payoff*: For each tick that a subject keeps his investment in the firm she receives \$0.004 (i.e., \$0.6 for every 30 seconds invested).

2. *End of round status*: Depending on the decisions of the particular subject and the decisions of the other members of the group, the round could end in three different ways.

   (a) <u>Exit:</u> if a subject exits the project at time t_e, then she gets back her initial investment of \$1, independently of the value $y_{(t_e)}$ of the project at that point.

   (b) <u>Bankruptcy:</u> if at time t_b two subjects have stopped rolling over their credit, then the firm will run on its cash reserves for θ ticks, until finally going bankrupt at $t_{b+\theta}$, being forced to sell its assets in the secondary market at a value $\alpha(F(y_{t_{(b+\theta)}}))$, where $F(.)$ is the present discounted value of the firm and

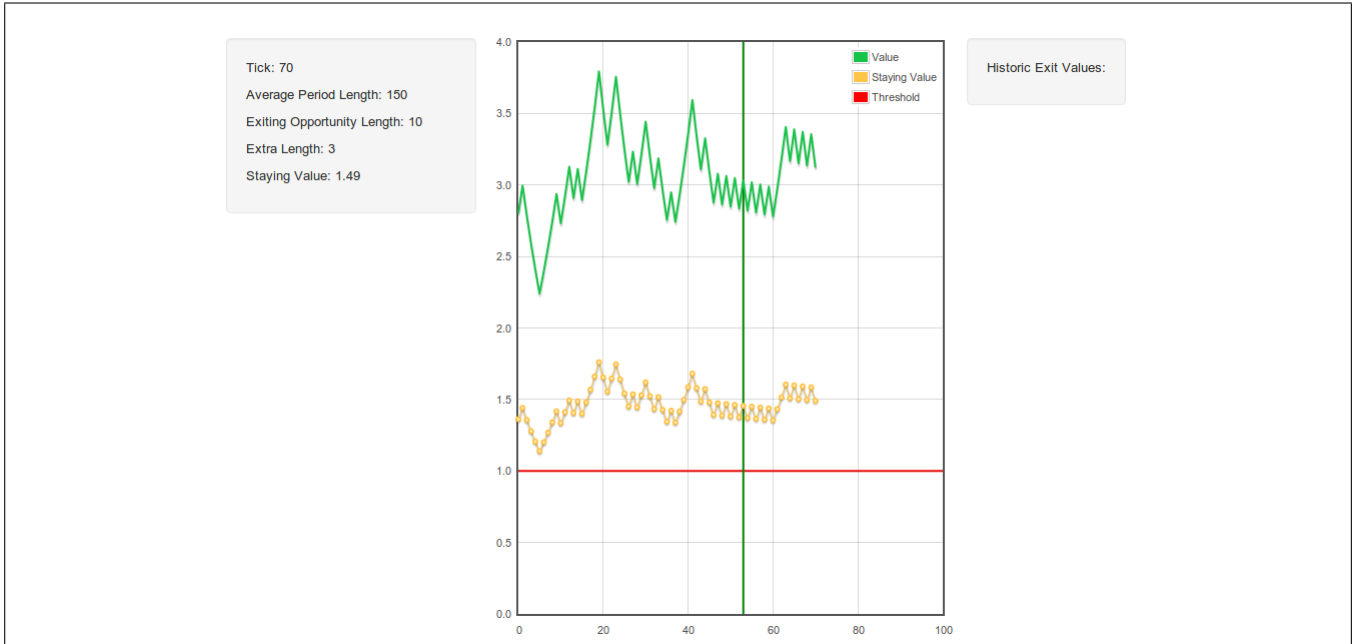---

[7]Once a subject decides to stop rolling over the credit that round is over for him (i.e. the decision to stop rolling over credit is final).
[8]This parameter comes directly from HX. In a future experiment we want to test the effects of changing θ.

Figure 2: Screen-shot

$0 < \alpha < 1$. At this point the firm will pay all subjects still invested $Min[1, \alpha F(y_{t_{b+\theta}})]$ and will cease to exist.

(c) <u>Natural Ending</u>: If the firm reaches its random "natural" ending t_n without going bankrupt, then all subjects still invested in the firm get $Min[1, y_{t_n}]$.

Subjects can keep track of both the firm's value (green jagged line in Figure 2), and of the fire-sale value (golden jagged line with dots in Figure 2) in the graphical interface on their screen. Other useful information appearing on the screen are the values at which subjects in the group decided to stop rolling over their credit in the previous 15 rounds (upper right box in Figure 2, the $1 threshold under which payoffs would be <$1 (horizontal red line in Figure 2), and the moment they had exited, if they had decided to do so (vertical green line in Figure 2).

4.2   Credit Rollover and Credit Maturities

The maturities system of this experiment is one of its unique aspects when compared to the experimental bank run literature, since in our game subjects decisions are not simultaneous. This creates several problems. The first one is how to keep the game flowing when all subjects have to decide asynchronously and every few seconds (see

8

Figure 3: Image of random maturity mechanism
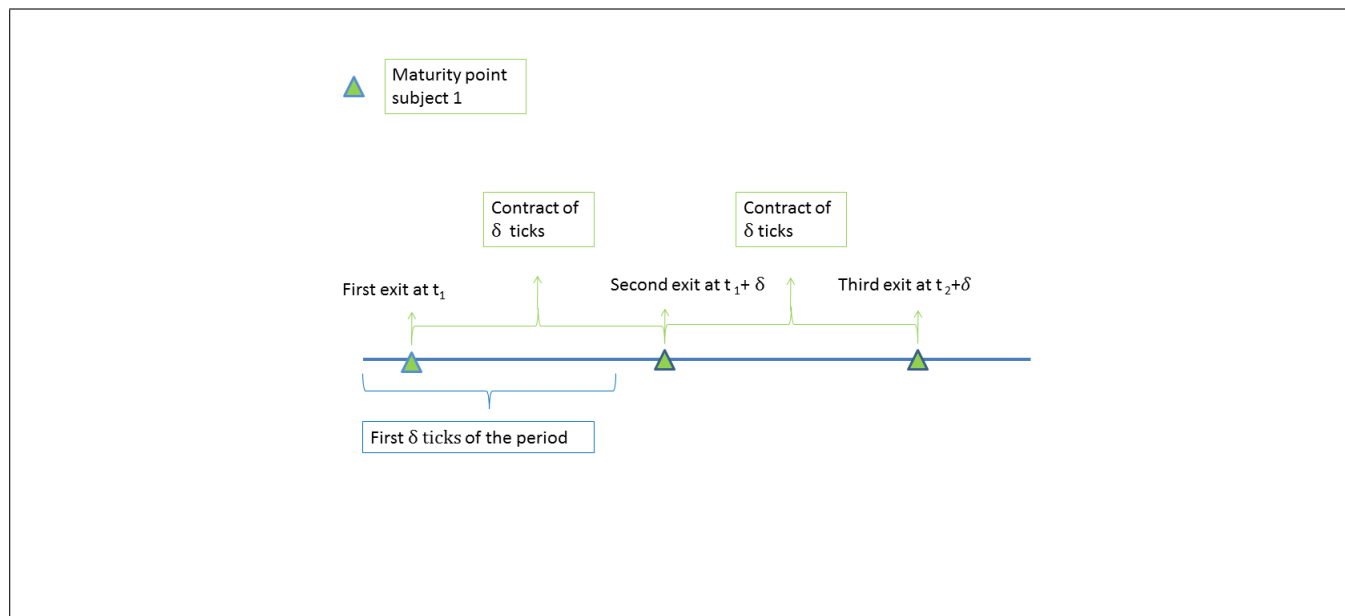
parameters below) whether or not to keep rolling over their credit. If we stopped the game at each point to have each individual subject ponder her decision, the experiment would be impossible to run. Our solution consists in having the credits roll over by default at each maturity point, unless the subject decides otherwise.

To stop this automatic rollover, subjects will have to "connect" three (consecutively) numbered buttons on the screen by hovering over them from left to right. The hovering idea comes from Brunnermeier and Morgan (2010), who introduce this mechanism to avoid subjects making inferences from clicking sounds coming from other terminals. We add the requirement that the hovering should be made following a certain gradient (from left to right), to avoid accidental stopping orders by subjects that inadvertently hover over the "stopping area", a prevalent problem reported in Brunnermeier and Morgan (2010).

Subjects can decide to stop rolling over their credit at any time during the experiment. If they decide so, at the next maturity point this decision will be implemented. In fact, the second problem we confronted with our staggered maturity system was how to avoid turning the maturity of credits into focal points. Because the proximity of a maturity point and the value of the collateral could interact in the decision-making of our subjects (and because we are only interested in decisions based on the value of the collateral), we hide the contract maturity points from our subjects.

To hide the maturity points, we fix the length of credits to be δ ticks and have the computer randomly assign in each round j, and for each subject i, a random starting point t_{1ij} within the first δ ticks. From this initial

9

Figure 4: Semi-Strategy Method Screen
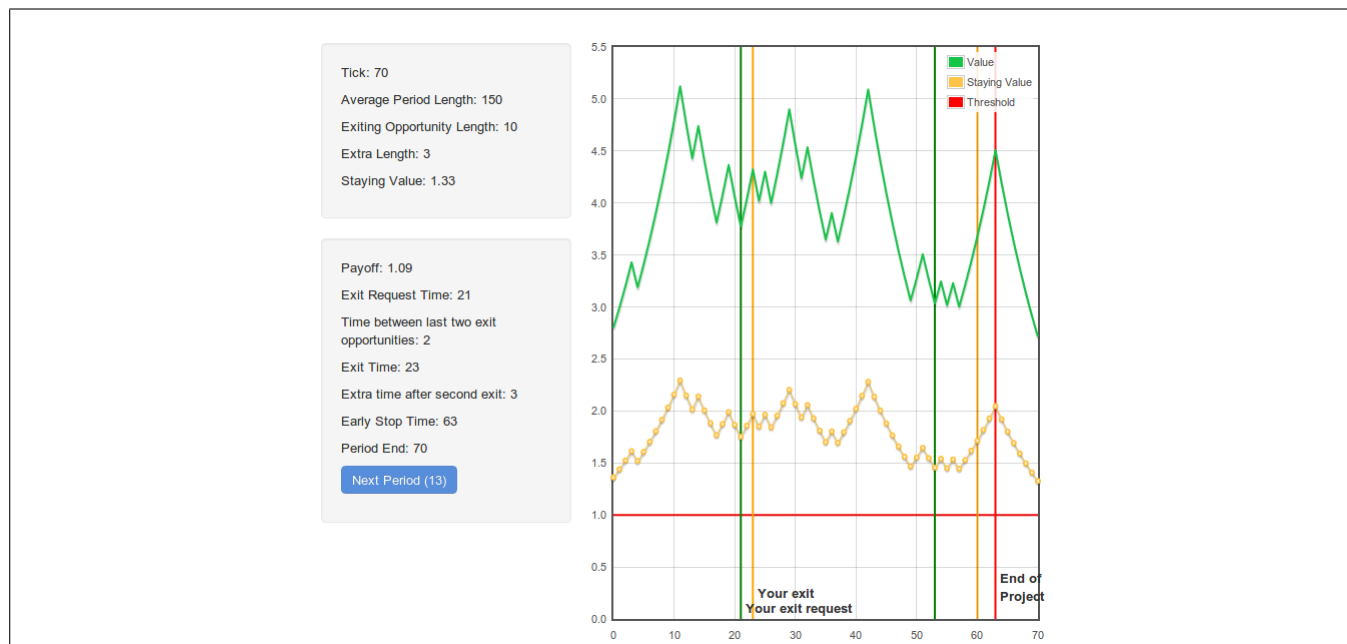
(individual) point, maturities will happen every δ ticks. So, for example, for subject i in round j his first maturity point will be at $t_{1ij}\epsilon[0,\delta]$, his second maturity point t_{2ij} at $t_{2ij} = t_{1ij} + \delta$, the third maturity t_{3ij} at $t_{3ij} = t_{2ij} + \delta$, etc. (see Figure 3). As a result, at every point in time the expected maturity of every subject is $\delta/2$ ticks away, akin to a Poisson shock of intensity $\delta/2$ in the HX model, and avoiding a focal point problem.

Finally, we borrow the idea of the "semi-strategy method" from Anderson et al. (2010) and let all rounds play until their random ending without providing any information to subjects of what other members of their group are doing. Once the round ends, all subjects are informed about all the events in the round (Figure 4), including other subjects (and own) requests to exit (green vertical lines), other subjects (and own) actual exit (orange vertical line), as well as round length and final payoffs, and a bankruptcy point (if there was one) shown as a red vertical line. This informational system allows us to gather more information than ending a round whenever a bankruptcy occurs.

4.3 Parameters and Hypotheses

As mentioned above, the goal of this experiment is to test the effects that maturity lengths have on the market for short-term credit. Therefore, we implement two treatments:

- <u>Long treatment:</u> Each contract is 8 seconds long (i.e., $\delta = 40$ ticks), and cash reserves last for 15 extra ticks

10

Table 1: Parameter Values

Parameter	Long Contract	Short Contract	Comment
δ	40 ticks	10 ticks	Contract Length
θ	15 ticks	3 ticks	Cash reserves
μ	0.0024	0.0024	Drift of the GBM
r	$0.004 per tick	$0.004 per tick	Per-tick flow payoff
σ^2	1.1	1.1	Volatility

after 2 subjects exit the market.

- Short treatment: Each contract is 2 seconds long (i.e., $\delta = 10$ ticks), and cash reserves last for 3 extra ticks after 2 subjects exit the market.

Plugging these parameters into the He and Xiong (2012) model, it predicts that the optimal stopping threshold in the Short treatment will be higher than in the Long treatment. Therefore, we make the following prediction:

- Prediction 1: Subjects will stop rolling over their credit at higher values of the firm in the Short treatment than in the Long treatment.

Our second prediction is that we will see subjects stopping their rollover at values of the firm where, in case of a fire-sale, creditors would get back all of their investment (i.e., $Min[1, \alpha F(y_{(t_{(b+\theta)})})] \geq 1$). Such a prediction is also in He and Xiong (2012).

- Prediction 2: Credit freezes will happen for values of the collateral such that, even at a fire-sale prices, the firm would be able to pay back all creditors in full.

5   Experimental Results

All session were run at the LEEPS lab of the University of California Santa Cruz, and all subjects were under-graduates from this institution. In total 92 subjects participated in the experiment, spread into 7 different sessions, and no subject played the game twice. In each session we had either 12 or 8 subjects for a total of 5,520 decisions

(60 rounds × 92 subjects)[9]. Among these observations we will ignore all stopping decisions for a value above $4, a value for which it is impossible to lose money in the Short treatment, and for which the probability of losing money in the Long treatment is <0.1%. In total we end up with 5,274 observations. In Figure 5 we plot the CDF for both treatments along with the theoretical stopping threshold from HX (vertical dotted lines) where a horizontal orange line indicates the median value of the distribution. As we can see in the left pane of Figure 5, the observed stopping values in the Long treatment are much lower than those in the Short treatment (Kolmogorov-Smirnov p-value = 0.000), which is in line with Prediction 1.

Figure 5: CDF plots for both treatments
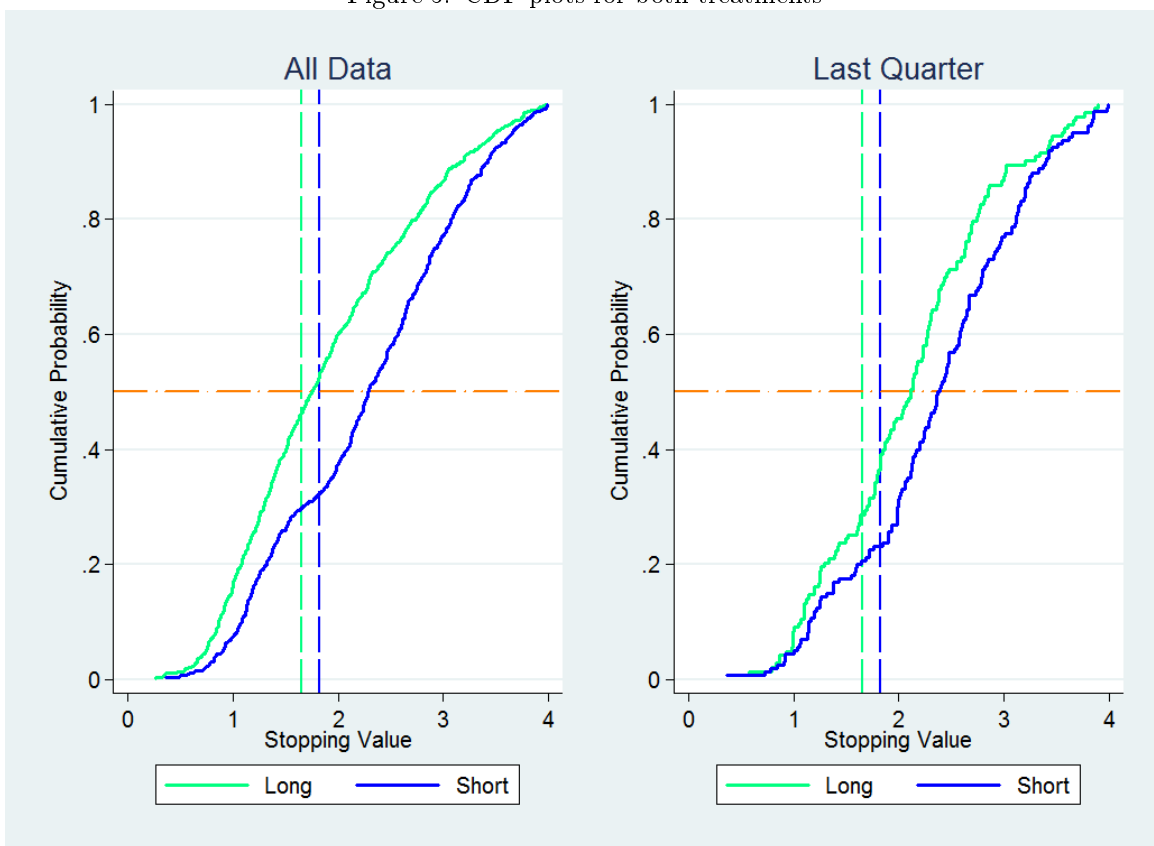
In addition, since we expect some learning from our subjects we break the experiment into quarters (15 periods in each quarter) to analyze the evolution of the stopping thresholds in the game. In the right pane of Figure 5 we show the CDF's for both treatments in the last quarter. As in the left pane, the stopping values across the two treatments continue to be widely separated in the fourth quarter (Kolmogorov-Smirnov corrected p-value = 0.008). But most interestingly, the stopping values for both treatments in the last quarter are significantly shifted to the right, which means that in the last quarter subjects stop rolling over their credit to the firm at much higher values

---

[9]Each session begins with some practice rounds whose results are not used in the analysis.

Table 2: Regression of stopping values and treatment

	(1) Stopvalue	(2) Stopvalue
1.short	0.363**	0.928***
	(0.066)	(0.116)
2.quarter	-0.234***	-0.2395***
	(0.0013)	(0.0457)
3.quarter	-0.218***	-0.200**
	(0.0022)	(0.0502)
4.quarter	0.0320***	-0.0786
	(0.0015)	(0.0532)
_cons	2.010***	1.807***
	(0.0336)	(0.0729)
N	1533	1533
adj. R^2	0.059	0.197
Group dummies	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of the collateral[10].

To have a more precise idea of how the behavior of subjects varies across quarters, we run a regression of the stopping value (Stopvalue), on our treatment variable ("1.short") and dummies for each quarter (Table 2). The results confirm that the length of the contract has an effect on the stopping decisions of our subjects, and show that there are statistically significant differences in behavior across quarters[11].

Table 2 confirms Prediction 1, namely, that Short contracts have significantly higher stopping values than Long ones. Yet, more interestingly, the second and third quarters dummies are significant *and negative* (i.e. subjects are taking higher risks in the second and third quarters by stopping their rollover at very lower values of the collateral), while fourth quarter stopping decision are not significantly different from those in the first quarter[12].

Plotting the kernel density estimates for each quarter in both treatments (Figure 6) we see a text-book boom and crash pattern for the credit market[13]. In the Long treatment the first quarter is similar to a left skewed plateau, while in the Short treatment, the initial stopping decisions follow a bimodal distribution with a higher right peak. Yet, if we look at the second quarter kernel density estimates, it appears that subjects in the first quarter follow

---

[10]If we compare the CDF's for the last quarter to aggregate stopping values of the rest of the session we get Kolmogorov-Smirnov p-value $= 0.022$ in the Short treatment, and p-value $= 0.000$ in the Long treatment.

[11]In Model 1, group dummies are not included in the regression, so clustering is at the quarter level. In Model 2, we include group dummies and cluster at the quarter-group interaction.

[12]A Kolmogorov-Smirnov comparing the CDF's of first and fourth quarter confirms this; p-value $= 0.248$, and p-value $= 0.182$ in the Short and Long treatments respectively.

[13]Appendix A has kernel densities broken into finer slices for a more precise description of the timings in the experiment

13

a "tâtonnement" learning process trying to get better acquainted with both the mechanics of the game and the strategy of other members of the group, leading them to realize after the first quarter that by taking higher risks they can earn more money. This significantly shifts the distribution of stopping values from their original starting points to much lower values (i.e. to the left).

This excessive risk-taking results in a series of costly bankruptcies that start a preemption race between subjects in the same group who, in order to avoid being caught in a costly fire-sale are start running for higher values as the experiment progresses, resulting in a shift of the overall stopping decisions to the right. So, as we can see in the third and fourth quarters, there is a "rebound effect", where subjects, after having taken excessive risk in the second quarter, are now being excessively cautious and sub-optimally stopping their credit too early.

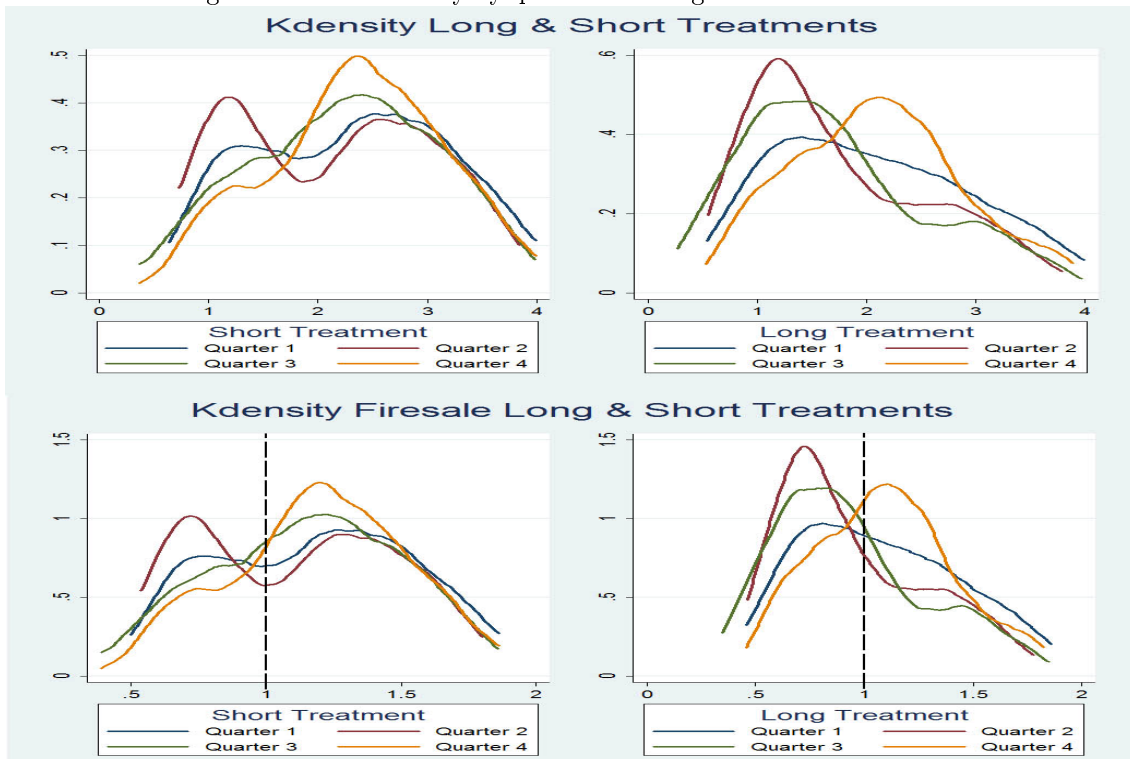Figure 6: Kernel density by quarter for Long and Short treatments

Table 3: Mean and SD of observed stopping values for each quarter by treatment

Quarter	Mean Long	Mean Short
1	1.92±0.898	2.26±0.901
2	1.35±0.887	2.11±0.924
3	1.48±0.839	2.23±0.893
4	2.02±0.774	2.35±0.851

In Table 3 we report the mean stopping values for each quarter and treatment. This table confirms the bubble

and crash story, and suggests why we didn't see a significant difference across the first and fourth quarter (Table 2); the "rebounding" leaves the final stopping threshold very close to the starting average[14].

Given the parabolic evolution of the stopping values, we call these dynamics the "boomerang effect". We can state now:

- Result 1: *The dynamics of stopping values does not move in one direction, but rather we show a bubble followed by a crash of the credit market; a "boomerang effect".*

And, while this first result is an interesting starting point for research in the dynamics of market bubbles, it is beyond of the scope of this paper to go into more detail, leaving the analysis of its potential applications to detect and prevent bubbles to future research.

A second important result of the experiment is the high number of credit runs on firms with strong fundamentals. We can see this at the bottom panels of Figure 6, where all stopping decision to the right of the dashed line are decisions to stop running even when the firm can pay all of its creditors in case of a fire-sale (i.e. cases where $\alpha F(y_t) \geq \$1$). These runs are called "frantic runs" in HX, and are a direct result of the staggered maturities. To our knowledge this is the first time that frantic runs are reported in any laboratory experiment.

- Result 2: *Overall, 60% of the decisions to stop rolling over the credit are made when firms have strong fundamentals, and can pay back the entire investment to all subjects even in the case of a fire-sale. In other words, the frantic runs predicted by HX occur in a laboratory setting with staggered maturities.*

Both Result1 and Result 2, as well as Table 2 and Table 3 are a first approximation to the results of our experiment. So far we have only used the information contained in the decisions to stop rolling over the credit. Yet, if a subject decides not to stop rolling over her credit, she is in fact telling us that her "stopping threshold" for that round is below the minimum value achieved by the collateral (for that round). Therefore, by only using the "observed" actions of subjects, we will only be reporting the decisions of those subjects more prone to run (i.e., those with the highest running thresholds), biasing or results upwards.

The unobserved stopping actions are what the survival literature would consider "censored observations". To overcome this censoring problem and avoid the bias of our estimates, we will have to use the product limit estimator[15] (Kaplan and Meier (1958)), a technique used to estimate survival functions that incorporates also the censored data.

---

[14] A Mann-Whitney test confirms that there is no significant difference between the first and last quarter, yet second and fourth are significantly different in both treatments.

[15] Also known as the Kaplan-Meier estimator.

5.1 Hazard Rates and the Product Limit Estimator

The product limit (PL) estimator is a non-parametric Maximum Likelihood Estimator of the distribution which is adapted to dealing with censored data. We will have a censored observation when for a given threshold of subject i in round j, (t_{ij}), the sequence of values of the firm for that round (y_j) never gets below that threshold (i.e. when $Min[y_j] \geq t_{ij}$).

Usually censored data are right-censored, and using the PL estimator with such data is pretty straightforward. But, our data set is left-censored, requiring us to "flip" the data to work with their mirror image. To do this we need to find a constant, S, large enough so that $S > max[y_j]$. As explained in some detail in Appendix C, we decided to give this constant a value 4, so $S = 4$.

In Figure 7 we present the hazard and the cumulative hazard estimates for the flipped data of each treatment. The hazard function can be understood as the "probability"[16] that a subject that has not stopped rolling her credit, decides to do so. To be more precise, the instantaneous hazard rate $(h(y))$ is a measure of the probability that a subject will decide to stop rolling over her credit within the (limiting) interval $\triangle y$ of collateral values, conditional on her not having already stopped rolling over her credit. Formally:

$$h(y) = \lim_{\triangle y \to 0} \frac{e[y, y + \triangle y]/N(y)}{\triangle y} \tag{4}$$

Where $e[y, y + \triangle y]$ is the number of observed rollover stops in the interval $[y, y + \triangle y]$, and $N(y)$ is the number of subjects at risk[17] for the value of the collateral y. From Equation 4 it is clear that if we do not take into account the censored observations, then $N(y)$ would be too high, bringing down the real hazard rate for the infinitesimal value of the collateral $\triangle y$, consequently biasing the hazard curve.

The Nelson-Aalen cumulative hazard estimate is an estimation of the cumulative hazard of the collateral for each of these limiting collateral values[18]. As shown in Figure 7 there is a strong interaction between the treatment hazard ratios and the value of the collateral. While for high values of the collateral Long contracts are more likely to be rolled over, when the collateral value falls, then Short contracts reduce the probability of a credit dry-up. This interaction between contract length and the value of the collateral seems to indicate that reducing the maturity mismatch in the market for short-term credit will not necessarily reduce the odds of sudden dry-ups. Rather, that the effects of this decision will differ depending on the state of the economy (in our case the value of the collateral). Therefore, our results suggests that, whenever feasible, the optimal approach to regulating short-term credit should

---

[16] Actually it is the ratio between the probability density function of the event (stopping the rollover) and the survivor function.

[17] Subjects at risk are all subjects that can still decide to stop rolling the credit, that is those that have either not stopped rolling or have not been censored.

[18] Note that because this is an *accumulated* measure, the estimate can go above the value of 1. The interpretation is that for those values above one, subjects would stop rolling their credit more than once if that were possible.

Figure 7: Cumulative hazard and hazard estimates for both treatments
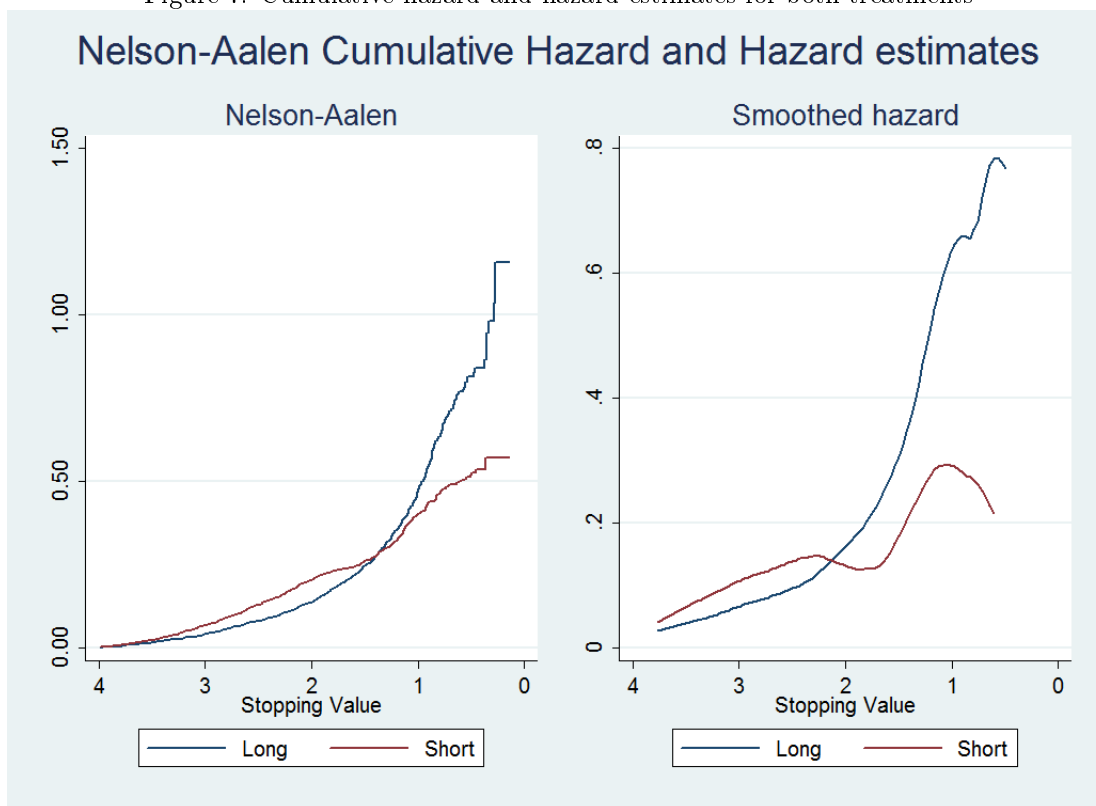
follow a *dynamic policy* favoring longer contracts (therefore lower maturity mismatch) when the economy is in good shape, while allowing for shorter-term contracts (therefore higher maturity mismatch) when the economy is in a recession. In summary:

- Result 3: Our data show crossing hazard functions for the different treatments, with a higher hazard estimate for the Short treatment when the value of the collateral is high, but a (much) higher hazard estimate for the Long treatment when the value of the collateral is low. This suggests that regulating the maturity mismatch of financial institutions could have opposite effects depending on the state of the economy.

Next we want to test whether or not both hazard functions are different across treatments. To do so we cannot simply compare the mean of each treatment through a t-test, as the survival curves might behave very similarly for some values of the collateral, but very differently for some others. To overcome this problem, survival analysis generally uses a weighted difference between expected and observed hazard rates across all the possible values of the collateral. These weights are used to give more or less importance in the statistical test to the different areas according to the priorities of the research.

In general, the most common test is the logrank test, which gives the same weight to all observations independently of the value of the collateral. Yet Fleming et al. (1980) show that this test is not appropriate when the proportional hazards assumption is not satisfied[19], or when there is heavy data censoring. Because the hazard functions of our two treatments cross, and because of the heavy right-hand censoring of our data[20], we cannot use the logrank test, leaving us with no systematic way to decide which are the best weights for our data analysis (Suciu et al. (2003)).

Yet, the selection of the appropriate weights and weight functions is central for the correct analysis of the survival data when the proportional hazards assumption is not met. In fact, Suciu et al. (2003) report that 96% of the papers appearing in major medical journals between 1999 and 2001, and using survival analysis to test crossing survival curves, use "inappropriate or questionable tests" for their tests of equality[21]. That is why we follow Gaugler et al. (2007) to select the best weight function for our crossing survival curves (see Appendix C for details).

We observe that the weight function that has more consistent p-values across the whole span of tested weight combinations is a modified Peto and Peto (1972) and Prentice (1978) function, which is exactly the same result reached by Gaugler et al. (2007)[22]. Using this weight function we find that for any combination of weights that emphasizes the low values of the collateral, the hazard curves of our treatments are significantly different (p-value=0.000)[23].

Unfortunately, if there is too much censoring in the data set at study, then both the hazard functions and the Kaplan-Meier survivor function estimates will be biased in the direction of the censored data points (Moeschberger and Klein (1985); Klein and Moeschberger (2003), Miller (1983)). Moreover, as shown in Moeschberger and Klein (1985) if there are censored points beyond the last precise observation (i.e., the last observed stopping value), then this right-hand bias will be even stronger. With almost 70% of our data censored, and with some of these censored values below the lowest observed stopping value, we will need to find a rule to reduce the censoring in our sample as in Anderson et al. (2010).

---

[19]In our case, the proportional hazards assumption would hold if the relation between both hazard curves in our experiment could be described *for all values of the collateral* as $H_L = \theta H_S$ where θ is any constant, and H_L and H_S the hazard functions for the Long and Short treatments respectively. That is, if the ratio between both hazard curves were the same across all values of the collateral. As it is clear from Figure 7, this is not the case in our data. For a lengthier discussion on the proportional hazards assumption see Suciu et al. (2003).

[20]See section 5.1.2.

[21]They suggest that the main reason for this misuse of statistical tests might be due to the fact that most statistical analysis software packages offer the logrank test as their default tool for survival analysis.

[22]The Peto and Peto (1972) and Prentice (1978) weight function uses an estimate of the survivor function to distribute the weights of the test statistic. Two parameters allow us to put or less weight at the beginning or the end of the estimated survivor functions. For a formal definition of the function see Appendix C.

[23]In fact, only when we put all the weight on the observations for high values of the collateral can we not reject the null hypothesis of equality between both hazard function estimates. See appendix C for a breakdown of different p-values.

5.1.1 Subset of the data

We start by dropping all observations that are censored below the latest observed stopping value in all sessions (34 observations). Then, following Anderson et al. (2010), we eliminate all observations in the sessions where the minimum value of the collateral was (strictly) greater than $0.9 (in total 3,921 observations). Finally, we also need to drop the subjects who run only two or fewer times in the whole session (180 observations). So, in total, we are left with a sub-sample of 1,128 observations, where 563 are stopping decisions and 565 are censored observations (almost a 50% ratio of censored data).

A graphical comparison between the survival curves of the full data set and the subset shows that the right-hand censoring bias in the PL estimator has been eliminated, with an increase in the steepness, and a shift to the left, of both treatment survival curves (see Appendix D for a graphical comparison of the survival curves between the full and partial data sets).

The results show that when using the modified Peto-Peto weight function for our statistical test we continue to have significant differences between the hazard curves (p-value=0.000 for weights on low values of the collateral, see Appendix E for a graph of the weight functions).

- <u>Result</u> 4: There is a clear treatment effect, with the null being rejected for all the tested p and q tuples. Therefore, changing the length of the contracts has a significant effect on the behavior of subjects.

Together, Result 5 and Result 3 show that while changing the maturity mismatch in the credit markets can have a large impact on the subjects behavior, such change should be done with extreme care since there is a strong correlation between the state of the economy and the effects it might have on the credit market. Specifically, while it might be a good idea to reduce the maturity mismatch when the economy is booming, the same intervention in the midst of a crisis could have extremely counter-productive effects.

Table 4: Mean and Bootstrapped SE of the estimated mean stopping value

Quarter	Mean Long	Mean Short
1	1.16±0.046	1.10±0.050
2	1.00±0.042	1.08±0.034
3	0.92±0.071	0.85±0.058
4	0.93±0.774	0.84±0.851

Finally, we use the PL estimator on the full data-set to calculate the mean stopping value for each quarter. The Short treatment has mean stopping values that are (in general) lower than in the Long treatment. 4 presents the results of the sub-sample estimates with their bootstrapped standard errors, indicating that if a dynamic policy like

19

the one suggested in Result 3 were not possible to implement, then a market with short contracts appears to be, on average, less prone to credit dry-ups.

- Result 5: The PL estimator shows that Short contracts have, on average, a lower mean stopping value than Long contracts.

6 Conclusion

Al Roth, in the introduction to the *Handbook of Experimental Economics* Kagel and Roth (1995) describes several reasons for running experiments. Among them, to offer policy advice ("whisper in the ear of princes"). This is exactly what we try to do in this paper.

To be specific, we try to compare the effects that maturity lengths have on the functioning of ABCP markets. Building on the theoretical model of He and Xiong (2012), we compare two markets where the only difference is the maturity of ABCP. Our main result is that while, on average, ABCP markets with short contracts are less prone to freezes (Result 6), the optimal policy is state-dependent, favoring long contracts when the economy is in good shape and short-term contracts when the economy is in a recession (Result 4). This latter result comes from estimating the hazard curves of the two different markets, and observing that while the value of the collateral remains high, the hazard curve for the Short markets is above that of the Long market, while, for low values of the collateral, not only does the hazard estimate for Long contracts overtake the Short market estimate, but it shoots up spectacularly, indicating an extreme risk of credit dry-up in the former market when the value of the collateral is low.

Our second important result (Result 2) is having shown, for the first time in a laboratory setting, runs on firms that can pay all their debts even in a fire-sale (i.e., runs on firms with strong fundamentals); what He and Xiong (2012) call "frantic runs". This is an important outcome, since it cannot be observed in the canonical static models of financial panics. As Bernanke (2008) puts it, the loss of access too secured borrowing during the 2007 crisis was "surprising" and firms had no contingency plans prepared to face this situation. Being possible now to reproduce them in the lab is a first step to better understand how they work, and how they can be prevented.

Finally, we also report rich experimental dynamics, with a consistent bubble-and-crash pattern across our sessions (Result 1). And, while studying the learning dynamics behind these results is beyond the scope of this paper, we believe that our experiment can be an interesting starting point to analyze sophisticated learning in any kind of dynamic markets.

It is our hope that a future buildup of experimental papers similar to ours will allow economists to whisper with

more confidence in the ear of financial princes.

References

Anderson, Steven T., Daniel Friedman, and Ryan Oprea, "Preemption Games: Theory and Experiment," *American Economic Review*, 2010, *100* (4), 1778–1803.

Arifovic, Jasmina, Janet Hua Jiang, and Yiping Xu, "Experimental evidence of bank runs as pure coordination failures," *Journal of Economic Dynamics and Control*, 2013.

Bernanke, Ben S., "Financial regulation and financial stability: a speech at the Federal Deposit Insurance Corporation's Forum on Mortgage Lending for Low and Moderate Income Households, Arlington, Virginia, July 8, 2008," *Speech*, 2008.

_ , "Financial regulation and supervision after the crisis: the role of the Federal Reserve : a speech at the Federal Reserve Bank of Boston's 54th Economic Conference, Chatham, Massachusetts, October 23, 2009," *Speech*, 2009.

_ , "Reflections on a year of crisis: a speech at the Federal Reserve Bank of Kansas City's Annual Economic Symposium, Jackson Hole, Wyoming, August 21, 2009," *Speech*, 2009.

Brown, Martin, Stefan Trautmann, and Razvan Vlahu, "Contagious Bank Runs: Experimental Evidence," DNB Working Paper 363, Netherlands Central Bank, Research Department 2012.

Brunnermeier, Markus K., "Deciphering the Liquidity and Credit Crunch 2007-2008," *Journal of Economic Perspectives*, 2009, *23* (1), 77–100.

_ **and John Morgan**, "Clock games: Theory and experiments," *Games and Economic Behavior*, 2010, *68* (2), 532–550.

_ , **Andrew Crockett, Charles Goodhart, Avi Persaud, and Hyun Shin**, *The fundamental principles of financial regulation*, Geneva London: International Center for Monetary and Banking Studies Centre for Economic Policy Research, 2009.

Calvo, Guillermo A., "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 1983, *12* (3), 383–398.

Chakravarty, Surajeet, Miguel A. Fonseca, and Todd Kaplan, "An Experiment on the Causes of Bank Run Contagions," Discussion Paper 1206, Exeter University, Department of Economics 2012.

Cheung, Yin-Wong and Daniel Friedman, "Speculative attacks: A laboratory study in continuous time," *Journal of International Money and Finance*, 2009, *28* (6), 1064–1082.

Diamond, Douglas W. and Philip H. Dybvig, "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 1983, *91* (3), 401–19.

Dufwenberg, Martin, "Banking on Experiments?," Technical Report, Mimeo 2012.

Farhi, Emmanuel and Jean Tirole, "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts," *American Economic Review*, February 2012, *102* (1), 60–93.

Fleming, Thomas R., Judith R. O'Fallon, Peter C. O'Brien, and David P. Harrington, "Modified Kolmogorov-Smirnov Test Procedures with Application to Arbitrarily Right-Censored Data," *Biometrics*, December 1980, *36* (4), 607.

Friedman, Daniel and Ryan Oprea, "A Continuous Dilemma," *American Economic Review*, 2012, *102* (1), 337–63.

Garratt, Rod and Todd Keister, "Bank runs as coordination failures: An experimental study," *Journal of Economic Behavior & Organization*, 2009, *71* (2), 300–317.

Gaugler, T., D. Kim, and S. Liao, "Comparing Two Survival Time Distributions: An Investigation of Several Weight Functions for the Weighted Logrank Statistic," *Communications in Statistics - Simulation and Computation*, 2007, *36* (2), 423–435.

Geanakoplos, John, "Promises Promises," Cowles Foundation Discussion Paper 1143, Cowles Foundation for Research in Economics, Yale University 1996.

_ , "The Leverage Cycle," Cowles Foundation Discussion Paper 1715, Cowles Foundation for Research in Economics, Yale University 2009.

Harrington, David P. and Thomas R. Fleming, "A Class of Rank Test Procedures for Censored Survival Data," *Biometrika*, December 1982, *69* (3), 553.

He, Zhiguo and Wei Xiong, "Dynamic Debt Runs," *Review of Financial Studies*, 2012, *25* (6), 1799–1843.

Heinemann, Frank, "Understanding financial crises: The contribution of experimental economics," *Annals of Economics and Statistics*, 2012.

Kagel, John H and Alvin E Roth, *The handbook of experimental economics*, Princeton, N.J.: Princeton University Press, 1995.

Kaplan, Edward L. and Paul Meier, "Nonparametric estimation from incomplete observations," *Journal of the American statistical association*, 1958, *53* (282), 457–481.

Klein, John P. and Melvin L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*, Springer, February 2003.

Klos, Alexander and Norbert Sträter, "How Strongly Do Players React to Increased Risk Sharing in an Experimental Bank Run Game?," Technical Report, QBER DISCUSSION PAPER No. 6/2013 2013.

Krishnamurthy, Arvind, "How Debt Markets Have Malfunctioned in the Crisis," *Journal of Economic Perspectives*, February 2010, *24* (1), 3–28.

Madies, Philippe, "An Experimental Exploration of Self-Fulfilling Banking Panics: Their Occurrence, Persistence, and Prevention," *The Journal of Business*, 2006, *79* (4), 1831–1866.

Miller, Rupert G., "What Price Kaplan-Meier?," *Biometrics*, December 1983, *39* (4), 1077–1081. ArticleType: research-article / Full publication date: Dec., 1983 / Copyright © 1983 International Biometric Society.

Moeschberger, M. L. and John P. Klein, "A Comparison of Several Methods of Estimating the Survival Function When There is Extreme Right Censoring," *Biometrics*, March 1985, *41* (1), 253–259. ArticleType: research-article / Full publication date: Mar., 1985 / Copyright © 1985 International Biometric Society.

Oprea, Ryan, Daniel Friedman, and Steven T. Anderson, "Learning to Wait: A Laboratory Investigation," *Review of Economic Studies*, 2009, *76* (3), 1103–1124.

_ , Keith Henwood, and Daniel Friedman, "Separating the Hawks from the Doves: Evidence from continuous time laboratory games," *Journal of Economic Theory*, 2011, *146* (6), 2206–2225.

Peto, Richard and Julian Peto, "Asymptotically Efficient Rank Invariant Test Procedures," *Journal of the Royal Statistical Society. Series A (General)*, 1972, *135* (2), 185.

Prentice, R. L., "Linear Rank Tests with Right Censored Data," *Biometrika*, April 1978, *65* (1), 167.

Schotter, Andrew and Tanju Yorulmazer, "On the dynamics and severity of bank runs: An experimental study," *Journal of Financial Intermediation*, 2009, *18* (2), 217–241.

Shin, Hyun Song, "Reflections on Northern Rock: The Bank Run that Heralded the Global Financial Crisis," *Journal of Economic Perspectives*, January 2009, *23* (1), 101–119.

Suciu, Gabriel P, Stanley Lemeshow, and Melvin Moeschberger, "Statistical Tests of the Equality of Survival Curves: Reconsidering the Options," in N. Balakrishnan and C.R. Rao, ed., *Handbook of Statistics*, Vol. Volume 23 of *Advances in Survival Analysis*, Elsevier, 2003, pp. 251–262.

Appendix A: Detailed Kernel Density

Figure 8: Density estimates for both treatments
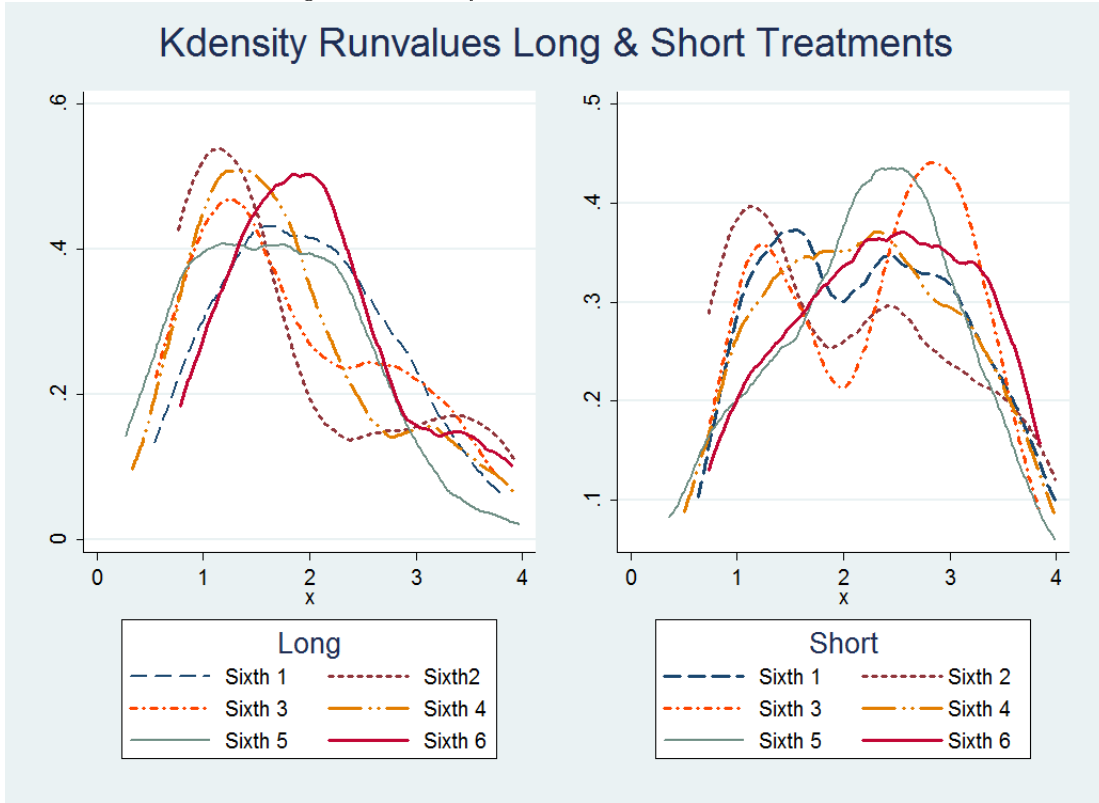
If we break up the sessions into sixth's (1/6) of a session, we have a more cluttered graph, but a more precise description of the dynamics of the experiment. As we can see, the underlying dynamics are the same, with the "boomerang effect" taking place. The difference now is that we can better see the timing of events. While the first sixth is again of a plateau-like shape, we see that risk taking actually starts by the second sixth (earlier than we expected) taking place by the second sixth. On the other hand, our first guess of a panic taking place by the end of the second quarter, and across the next periods was correct, as we can see from the fourth and fifth sixth of the data.

A very interesting result of this more detailed breakdown of estimated CDF's is that we can better appreciate how polarized are the stopping thresholds in the Short treatment. This is clear when observing that most sixths have a bimodal shape, especially in the "panic" (third) one, but also in the final ending distribution. While in the Long treatment this distribution is a sharp hill, in the Short treatment we see a more spread-out result.

Appendix B: Data Flipping

Flipping the data is a simple procedure where we just need to find a constant, S, large enough such that $S > maxy_{ij}$ for all subject i and round j. Since all data are in the interval $[0, 4)$, $S = 4$ can be used to flip the data. Therefore, for every subject i and round j, we can define z_{ij} such that $z_{ij} = 4 - y_{ij}$. A graphical explanation of the process is found in the Appendix B figure.

Figure 9: Left to right censoring switch
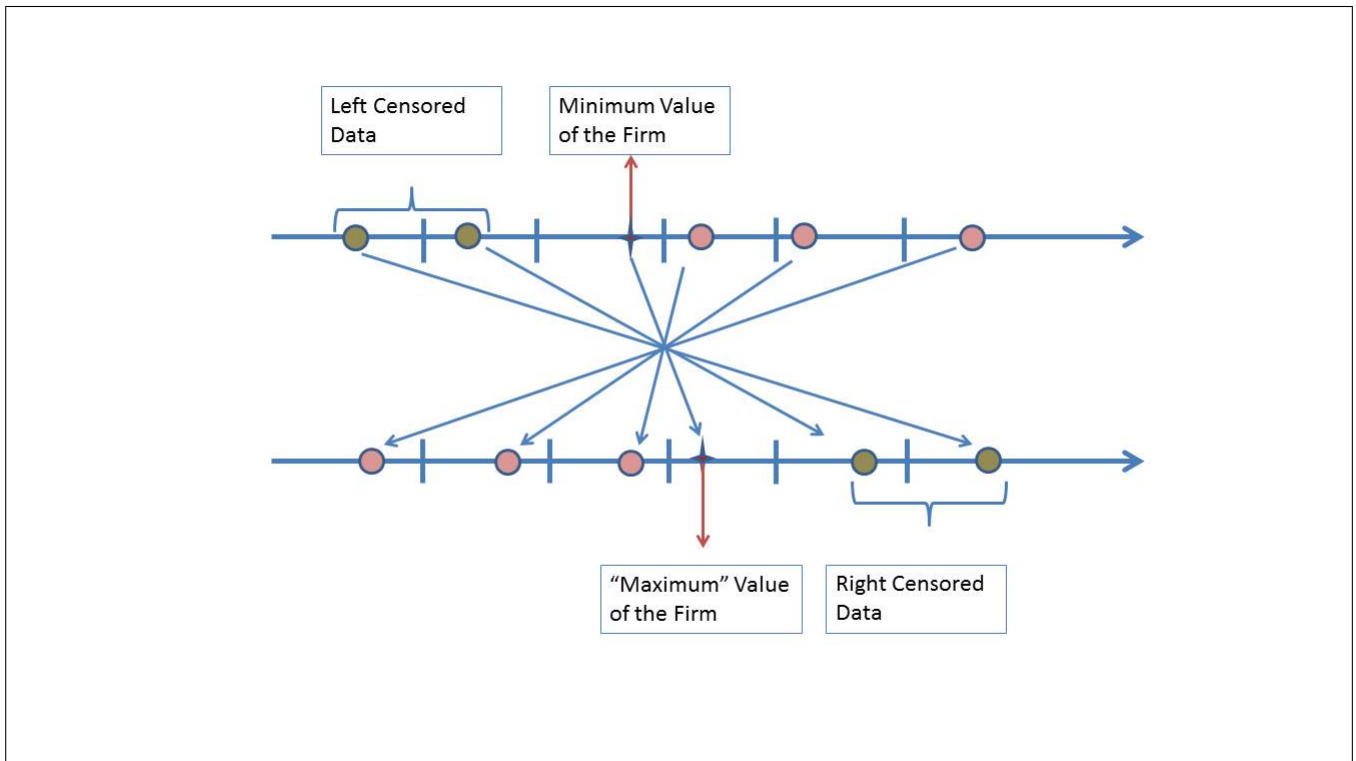
Appendix C: Picking the Correct Weight Function

Because the hazard functions cross each other, we need to take a "search and find" approach to choose the best non-parametric comparison method for our analysis. To do so we will need a set of different weight functions to compose our test statistic, and compare the "smoothness" of the resulting p-values for the different possible weights using bootstrapping techniques.

Following Gaugler et al. (2007) notation, we define the logrank statistic comparing the data from both Long and Short treatments as:

$$A_w = \sum_{i=1}^{l} W_i \left[d_i - \frac{n_i D_i}{N_i} \right] \tag{5}$$

Where d_i is the number of subjects that stopped rolling over their credit in the Long treatment at a value of the collateral $t = i$, n_i is the number at risk at $t = i$ in the same group, D_i the *pooled* (both Long and Short treatments) number of stopping decisions *until* $t = i$, and $N = i$ is the number of *pooled* subjects at risk at $t = i$. Finally, W_i is the weight function for the statistic at $t = i$. The weight functions are critical in determining the results of the test, and should be used in accordance with the needs of the researcher. Some examples are the Logrank which uses $W_i = 1$, where all observations have the same weight, the Gehan ($W_i = N_i$), or the Tarone-Ware $\left(W_i = N_i^{1/2} \right)$, both of which are designed to give more emphasis to the regions that contain more observations.

In our case we will study variations of an extremely versatile and well known weight function, the Fleming-Harrington weight function (Harrington and Fleming (1982)) which includes many other well-known weight functions as special cases[24]:

$$W_i = \left[\hat{S} \left(t_{i-1} \right) \right]^p \left[1 - \hat{S} \left(t_{i-1} \right) \right]^q \tag{6}$$

The Fleming-Harrington weight function for the statistic at $t = i$ is a function of the Kaplan-Meier survivor function estimate[25] at $t = i - 1$, ($\hat{S} \left(t_{i-1} \right)$), and two parameters, p and q which are used to give more or less importance to the different areas of study. In particular, when $q = 0$ and $p > 0$ more weight is given to rollover stops for high values of the collateral, and when $q > 0$ and $p = 0$ more weight is assigned to stopping decisions for low values of the collateral.

In the following we will study different weight functions as suggested in Gaugler et al. (2007) by modifying the time dependence of the Fleming-Harrington weight function at $t = i$ from $t = i - 1$ to $t = i$, so, from $\hat{S} \left(t_{i-1} \right)$ to

---

[24]For example, when $p = 0$ and $q = 0$ the Fleming Harrington weight function turns into the the logrank test ($W_i = 1$).

[25]The K-M survivor function estimate is defined as $\hat{S} = \prod_{t_i \leq t} \left(1 - \frac{b_i}{m_i} \right)$.

Table 5: Asymptotic p-values

p	q	$\hat{W}_{pq}(t_{i-1})$	$\hat{W}_{pq}(t_i)$	$\tilde{W}_{pq}(t_{i-1})$	$\tilde{W}_{pq}(t_i)$
1.00	0.00	0.244	0.245	0.244	0.245
0.97	0.03	0.140	0.159	0.139	0.158
0.90	0.10	0.038	0.043	0.038	0.043
0.80	0.20	0.003	0.003	0.003	0.003
0.70	0.30	0.000	0.000	0.000	0.000
0.50	0.50	0.000	0.000	0.000	0.000
0.30	0.70	0.000	0.000	0.000	0.000
0.00	1.00	0.000	0.000	0.000	0.000
0	0	0.013	0.013	0.013	0.013

Table 6: Bootstrapped p-values

p	q	$\hat{W}_{pq}(t_{i-1})$	$\hat{W}_{pq}(t_i)$	$\tilde{W}_{pq}(t_{i-1})$	$\tilde{W}_{pq}(t_i)$
1.00	0.00	0.246	0.227	0.237	0.259
0.97	0.03	0.159	0.139	0.133	0.170
0.90	0.10	0.049	0.042	0.045	0.050
0.80	0.20	0.003	0.007	0.006	0.003
0.70	0.30	0.000	0.000	0.000	0.000
0.50	0.50	0.000	0.000	0.000	0.000
0.30	0.70	0.000	0.000	0.000	0.000
0.00	1.00	0.000	0.000	0.000	0.000
0	0	0.016	0.014	0.017	0.016

$\hat{S}(t_i)$, and by changing the Kaplan-Meier estimate $\left(\hat{S}(\cdot)\right)$ by the Peto-Peto estimate $\left[\tilde{S}(\cdot)\right]$ [26]. This will leave us with four different weight functions; the original F-H $\left(\hat{W}_{pq}(t_{i-1})\right)$, a modified F-H $\left(\hat{W}_{pq}(t_i)\right)$, the original P-P $\left(\tilde{W}_{pq}(t_{i-1})\right)$), and a modified P-P $\left(\tilde{W}_{pq}(t_i)\right)$ (Equation 7):

$$Kaplan-Meier = \begin{cases} \hat{W}_{pq}(t_{i-1}) = \left[\hat{S}(t_{i-1})\right]^p \left[1 - \hat{S}(t_{i-1})\right]^q \\ \hat{W}_{pq}(t_i) = \left[\hat{S}(t_i)\right]^p \left[1 - \hat{S}(t_i)\right]^q \end{cases} \quad Peto-Peto \begin{cases} \tilde{W}_{pq}(t_{i-1}) = \left[\tilde{S}(t_{i-1})\right]^p \left[1 - \tilde{S}(t_{i-1})\right]^q \\ \tilde{W}_{pq}(t_i) = \left[\tilde{S}(t_i)\right]^p \left[1 - \tilde{S}(t_i)\right]^q \end{cases}$$

$$(7)$$

In Figure 10 we present the plots, for different p and q values, for the four weight functions[27]. As we can see the logrank weight function gives the same weight (1) to all observations in the data, while all other weight functions seem to converge at giving a weight of 0.5 to those observations at the lowest values of the collateral. This convergence is due to the heavy right-hand censoring observed in our data (see section 5.1.2). Notice also that for the weight functions that use the survivor estimate evaluated at $t = i - 1$ we see a jump at $t = 0$. This jump could be problematic if we were interested in the differences of our hazard curves for high values of the collateral, yet we are interested in the lower values of the collateral, so our choice should *a priori* favor the cases where $p < 0.5$ and $q > 0.5$ which are not affected by the jump. For a longer discussion on the implication of these weight jumps see Gaugler et al. (2007).

Next we compare the evolution of p-values across the weight functions, for the different values of p and q (Table 6)[28]. In addition we also compare the bootstrapped p-values to the asymptotic p-values (Table 5).

---

[26] The Peto-Peto as survivor function estimate is defined as $\tilde{S} = \prod_{t_i \leq t} \left(1 - \frac{b_i}{m_i+1}\right)$.

[27] Following Gaugler et al. (2007) we have limited the values of p and q to those where $p + q = 1$.

[28] To find the p-values we follow Gaugler et al. (2007) and create a 1000 bootstrap synthetic data-sets to then calculate for each of them the test statistic $(A_1^*, ..., A_{1000}^*)$ following Equation 5. Finding the bootstrap p-value as $p* = \sum_{i=1}^{1000} \mathbf{I}\left\{A_I^* \geq A_{org}\right\}/1000$ where A_{org} is the value of the test statistic calculated from the original data set.

Figure 10: Four different weight functions
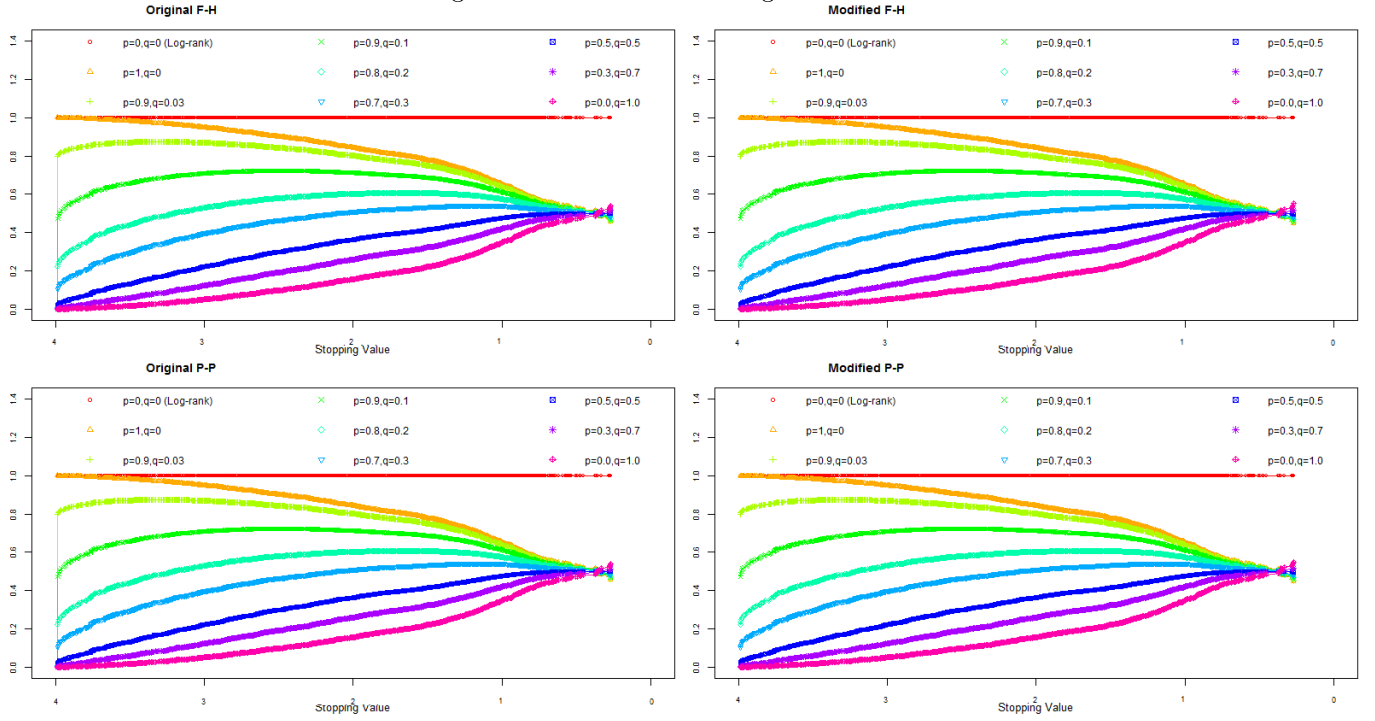
As we can see, the bootstrapped p-values are similar to those coming from the asymptotic theory. In both cases, the results show that we cannot reject the null hypothesis of equality between both survivor curves when we place all the weight on the early stopping decisions, but that once we move away from $p = 1.00$ and $q = 0.00$, there is a sharp drop in p-values with significant differences for all weight functions where $p \leq 0.9$. This abrupt drop in p-values is consistent with 7, where the divergence between both hazard functions is clearly higher for the lower values of the collateral. Therefore, not only are the bootstrapped p-values of all weight functions aligned with their asymptotic counterparts, but the results match a graphical inspection of the data.

Like in Gaugler et al. (2007), the jump in p-values is much smoother in both $\hat{W}_{pq}(t_i)$ and $\tilde{W}_{pq}(t_i)$ than in the cases where the weight function is based on the survivor estimate at t_{i-1}. And between the two, the best choice for is the Peto-Peto weight function with no time lag ($W_{pq}(t_i)$), as it has the least variation in p-values across all the tested weight combinations. This turns out to be the same conclusion that Gaugler et al. (2007) reach in their own analysis of all the above weight functions.

- <u>Result</u>: Like in Gaugler et al. (2007) the weight function that seems most appropriate for our test if the modified Peto-Peto: $\tilde{W}_{pq}(t_i) = \left[\tilde{S}(t_i)\right]^p \left[1 - \tilde{S}(t_i)\right]^q$.

Appendix D: Full Data Set and Subset Comparison

As the PL estimates graphed below show, our dropping of some data does not affect the underlying relationship of the hazard functions but, as predicted by the theory, it shifts the estimates in the opposite direction of the censoring

Figure 11: KM Survival estimates for all data and subset
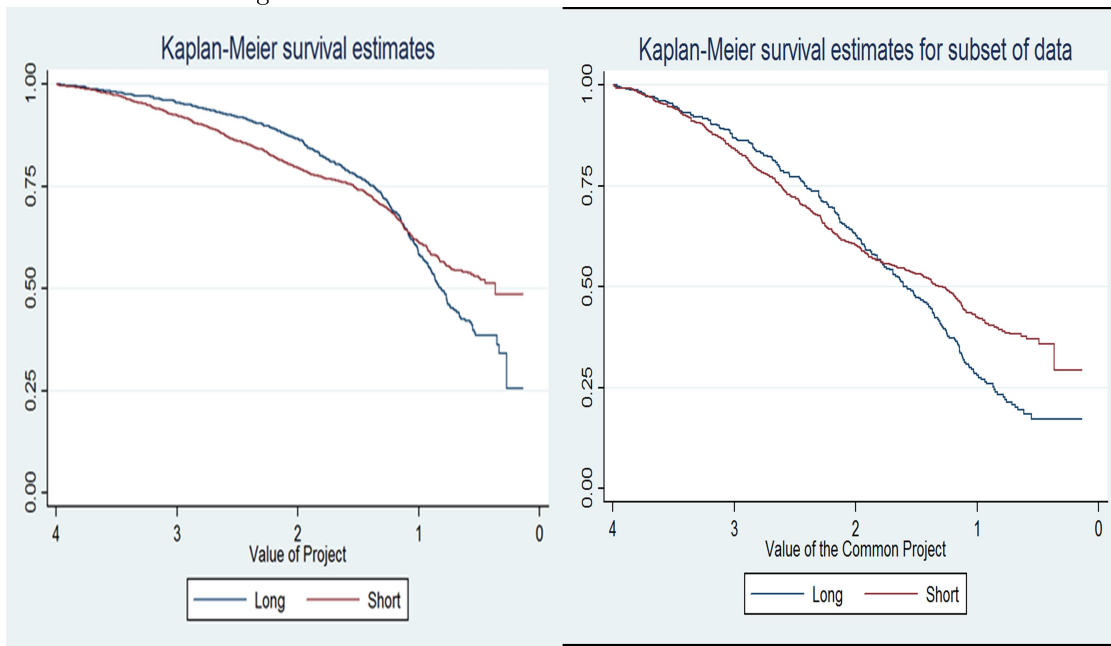
Appendix E: Weight Function Graphs for Full Data and Subset of the Data

Plotting the four weight functions (Figure 12) we can observe that the weight functions behave differently than when we use the full data-set, and look very much like the similar weight functions studied in Gaugler et al. (2007), Suciu et al. (2003), or Klein and Moeschberger (2003), confirming that the our data subset has corrected for the bias in the PL estimator and behaves much more the typical right-censored data set.

Figure 12: Four Different Weight Functions
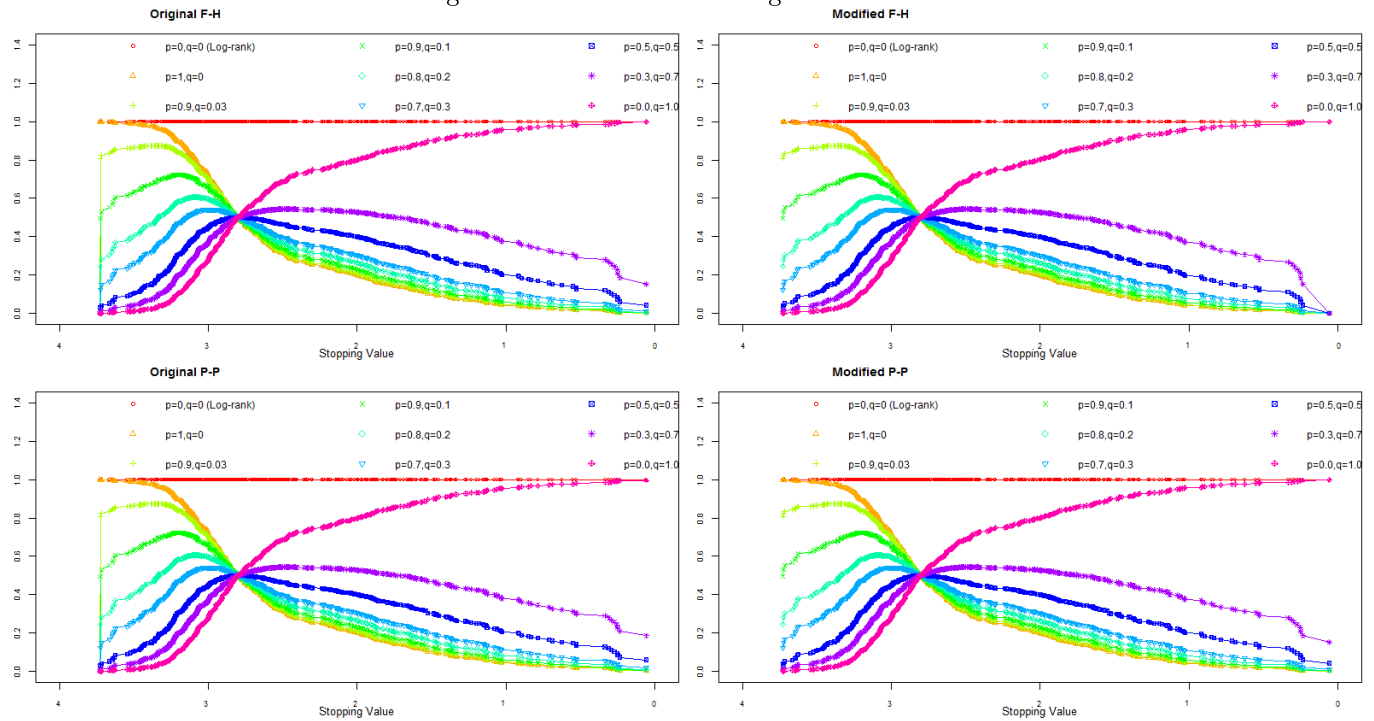
Instructions

<u>**Timing of the Experiment**</u>:

The session we will be running today has 60 rounds. At the beginning of the session you will be grouped with 3 other subjects <u>with whom you will play all 60 rounds of the session</u>.

The time units of the round are "ticks" (1/5 of a second). Each round has a probability of 1/150 per tick of maturing; this means that on average each round will last 30 seconds.

<u>**The Common Project**</u>:

In each round, everyone in your group will start by investing 1 florin (lab currency) into a common project. Every tick the value of the common project will change. To be precise:

- The value of the common project will go up with probability: 0.5001, and down with probability: 0.4999.

- The change in value (whether up or down) will always be 7% of the current value of the investment.

You will be able to track the value of the firm on your screen:

[Image on projector]

<u>**Your Decision**</u>:

In each round you will make only **ONE** decision:

- To stay in the common project

- To exit the common project

<u>How to exit a project</u>: To exit the common project you will need to slide (not click) your mouse over the counter at the bottom of your screen and connect the numbers 3, 2, and 1

[Image on projector]

Once you have done so a green line will appear on your screen. This green line marks your "**exit request**" and you will exit at the next "**exit gate**" after your exit request.

Exit gates are individual (so no two players share the same exit gate), and happen every 8 seconds. To be more precise:

- In each round, every member of a group is assigned a first "exit gate" within the first 8 seconds

- After that, his next exit opportunities will happen every 8 seconds.

- Example: imagine your first exit opportunity is in second 2 of the round, then your next exit opportunity will be in second 10, then 18, then 26 etc.

[Image on projector]

To stay in the project you do not have to do anything.

Overview:

1. In this experiment you are grouped with three other subjects across 60 rounds.

2. In each round you all start with an investment of 1 florin in the common project

3. Each round you are asked to make one decision: whether or not to stay invested in the common project

4. To exit you need to swipe your mouse over the 3,2,1 countdown area.

5. This swiping will record an exit request and you will exit at your next exit gate

6. To stay you do not need to do anything

Payoffs:

Your payoff in each round will come from two different sources:

- Constant Return

- Original investment return

How much you make from each income source will depend on your decision to stay or to exit, and on the staying or leaving decisions of the other investors in your group.

Constant Return: For every "tick" that you keep your investment in the project, you will get a constant return. This constant return is of 0.004 florins per tick. This means that if you keep your investment for 30 seconds you will get 0.6 florins from the constant return (so a 60% return for every 30 seconds).

Original investment of 1 florin: Of the original investment of 1 florin that you made at the beginning of the round you can get back either the original florin you invested, or a part of the florin you invested, but never more.

This payoff will depend on:

1. Your decision to stay or to exit

2. The decisions of the other investors in your group

34

3. When and how the round ends.

The round can end in three different ways:

1. You Exit the project: if, at some point, you decide to exit the project, and are able to do so, you will get your 1 florin back independently of the value of the common project. On the other hand, you will stop getting paid the constant return per tick for the rest of the round.

2. Premature end of the project: if 2 investors in your group exit the project, then the project will continue running for 2 extra seconds before it "ends early" and pays all of the remaining investors a "staying value". How much the staying value pays back will depend on where the jagged yellow line is at the moment of the premature ending: a) If the jagged yellow line is above 1, then you will be paid 1 florin. b) If the jagged yellow line is below 1, then you will be paid the value of the line at that point.

3. Maturation of the project: as mentioned, the common project has a probability of 1/150 per tick of maturing. If the common project matures before an early stop happens, then all investors will be paid depending on the value of the common project (green jagged line): a) If the jagged green line is 1 or greater than 1, then all players that are still invested get their 1 florin back b) If the value of the common project is below 1, then all players that are still invested will get back the value of the common project at that point.

You can track both the value of the project and the premature ending value of the project on your screen.

[Image on projector]

Overview of the payoffs:

1. Your payoffs come from two different sources: a. Constant payoff b. Individual end of the round

2. The constant payoff gives you 0.004 florins per tick as long as you are invested and the round has not finished (there has not been a premature ending or a maturation of the project)

3. Individual end of projects has 3 different ways of taking place: a.

- You withdraw your investment and get back your entire 1 florin independent of the value of the common project

- The project has a premature ending, in which case those investors that are still in the project get back 1 florin if the yellow jagged line is above 1, or the value of the jagged line if it is below the value of 1

- The project matures, at which point all those still invested get back 1 florin if the green line was above 1, the value of the green line if it was below 1

Important things to notice:

All rounds will continue ticking until the project's maturation, so even if there are premature ending, you will not be given this information until the end of the round. You will also not be told when other investors are leaving the common project nor will you be told where your exit gates are. The information that you will see while the round is ticking will be:

- Value of the project

- Staying value

- Past exit requests by all investors in your group (upper right corner of screen)

[Image on projector]

Once the project has matured, then a screen will appear showing the whole unraveling of the round which includes:

- The exit requests made by all players (green lines)

- The actual exits at each individual exit gate (yellow lines)

- You will also be informed about your exit request and your allowed exit tick.

- Finally, if there was an premature ending it will be shown as a red line.

[Image on projector]

In summary:

Your goal each round is to decide whether you leave or not the project balancing the advantages and disadvantages of staying invested, the probability of a natural end and the behavior of other investors.

But not all rounds are paid. Not all rounds will count for your final payoffs. Although you will see how much you made at the end of each round, 10 of the 60 rounds will count towards your final payoffs. These 10 rounds are randomly chosen by the computer.

Practice: Before the session properly begins, we will have 6 practice rounds so that you get used to the mechanics of the session, so you should practice exiting. These rounds will be shorter than the rounds during the experiment.

While the instructions are somewhat long and complex, it is very important that you understand how the game works. You don't need to really understand all of the probabilities and numbers that we give you, as you can learn from experience, but you should make sure that you understand the mechanics of the game.

FAQ:

1) Is there a pattern in the change of value of the common project? No, we really tried to make it random. No matter what is the history of values that the common project took the probabilities of going up or down on value are always the same.

2) If values over the threshold of 1 always pay me back 1 Florin, why do you show them to me? We show you these values because we think you might be interested in knowing how far away you are from the 1 florin threshold.

Please feel free to ask as many questions as necessary to make sure that you have a full understanding of the instructions. To ask a question, just raise your hand to call my attention.

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649,
please visit http://sfb649.wiwi.hu-berlin.de.

001 "Principal Component Analysis in an Asymmetric Norm" by Ngoc Mai
     Tran, Maria Osipenko and Wolfgang Karl Härdle, January 2014.
002 "A Simultaneous Confidence Corridor for Varying Coefficient Regression
     with Sparse Functional Data" by Lijie Gu, Li Wang, Wolfgang Karl Härdle
     and Lijian Yang, January 2014.
003 "An Extended Single Index Model with Missing Response at Random" by
     Qihua Wang, Tao Zhang, Wolfgang Karl Härdle, January 2014.
004 "Structural Vector Autoregressive Analysis in a Data Rich Environment:
     A Survey" by Helmut Lütkepohl, January 2014.
005 "Functional stable limit theorems for efficient spectral covolatility
     estimators" by Randolf Altmeyer and Markus Bibinger, January 2014.
006 "A consistent two-factor model for pricing temperature derivatives" by
     Andreas Groll, Brenda López-Cabrera and Thilo Meyer-Brandis, January
     2014.
007 "Confidence Bands for Impulse Responses: Bonferroni versus Wald" by
     Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, January
     2014.
008 "Simultaneous Confidence Corridors and Variable Selection for
     Generalized Additive Models" by Shuzhuan Zheng, Rong Liu, Lijian Yang
     and Wolfgang Karl Härdle, January 2014.
009 "Structural Vector Autoregressions: Checking Identifying Long-run
     Restrictions via Heteroskedasticity" by Helmut Lütkepohl and Anton
     Velinov, January 2014.
010 "Efficient Iterative Maximum Likelihood Estimation of High-
     Parameterized Time Series Models" by Nikolaus Hautsch, Ostap Okhrin
     and Alexander Ristig, January 2014.
011 "Fiscal Devaluation in a Monetary Union" by Philipp Engler, Giovanni
     Ganelli, Juha Tervala and Simon Voigts, January 2014.
012 "Nonparametric Estimates for Conditional Quantiles of Time Series" by
     Jürgen Franke, Peter Mwita and Weining Wang, January 2014.
013 "Product Market Deregulation and Employment Outcomes: Evidence
     from the German Retail Sector" by Charlotte Senftleben-König, January
     2014.
014 "Estimation procedures for exchangeable Marshall copulas with
     hydrological application" by Fabrizio Durante and Ostap Okhrin, January
     2014.
015 "Ladislaus von Bortkiewicz - statistician, economist, and a European
     intellectual" by Wolfgang Karl Härdle and Annette B. Vogt, February
     2014.
016 "An Application of Principal Component Analysis on Multivariate Time-
     Stationary Spatio-Temporal Data" by Stephan Stahlschmidt, Wolfgang
     Karl Härdle and Helmut Thome, February 2014.
017 "The composition of government spending and the multiplier at the Zero
     Lower Bound" by Julien Albertini, Arthur Poirier and Jordan Roulleau-
     Pasdeloup, February 2014.
018 "Interacting Product and Labor Market Regulation and the Impact of
     Immigration on Native Wages" by Susanne Prantl and Alexandra Spitz-
     Oener, February 2014.

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649,
please visit http://sfb649.wiwi.hu-berlin.de.

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.