Risk assessment in the mutual fund and insurance industry

PhD Thesis in Financial Mathematics, XXXI Cycle

Clemente De Rosa

Supervisors: Fabrizio Lillo, Elisa Luciano

November 26, 2020

Contents

Introduction

This Thesis is organized as a collection of two different parts. The first studies the problem of identification and estimation of peer effects in the Money Market Mutual Fund industry and their impact on its financial stability with respect to widespread redemption shocks. The second part focuses on the modeling of longevity risk in a multi-population setting and the assessment of basis risk in the context of dynamic hedging and international expansions.
While the first part is closely connected to the financial literature about mutual funds and fire-sales spillovers, the second part contributes instead to the actuarial literature on continuous time stochastic mortality models. Despite being different from many point of views, the two parts share a very simple and important common objective: modeling and assessing *Risks*.

0.1 Peer effects and financial stability

Money Market Mutual Funds (MMMFs) have attracted the attention of both academics and regulators due to the several runs they have suffered in recent years. In particular, the 2007-2008 financial crisis marked a significant shift in the opinion about the riskiness of the MMMF industry and its importance for the stability of the entire financial system. Money market funds were traditionally considered by market participants very safe investments, similar to bank deposits. This widespread belief was generally supported by the historical evidence - since no major money market fund had ever failed before the crisis - and by their regulatory constraints. Indeed, under Rule 2a-7 of the Investment

Company Act of 1940, MMFs were required to invest only in high quality short-term money market instruments, they could not be exposed for more than 5% to any single issuer, and were allowed to maintain a stable Net Asset Value (NAV) of $1.00 per share by valuing their portfolios at “amortized cost”.

However, in September 2008, the default of Lehman Brothers caused the NAV/Share of the Reserve Primary Fund to fall below the value of $ 0.995[1] triggering widespread redemptions in the entire industry. The failure of the Reserve Primary Fund showed firstly that MMFs were bearing significant risks - Kacperczyk and Schnabl [2013] highlights that in the year prior the crisis the excess yield of MMF with respect to US Treasuries rose from 15bp to 90bp with a substantial increase in the cross-sectional dispersion - and secondly that the entire industry was at risk for run-like behaviors. As described in Schmidt et al. [2016], during the week of September 15 2008 other prime funds suffered large outflows - many amounting to more than 10% of their asset under management in a single day - totaling about $300 billion. Given the already difficult conditions of the credit market, money market funds were struggling to sell securities in order to meet redemptions, showing they could be vulnerable to aggregate risks such as the risk of run-like behavior or the risk of fire sales spillovers.

Similar risks, although at a much smaller scale, materialized during the European Sovereign Debt Crisis of summer 2011 in what has been named the "slow-motion run" (see Chernenko and Sunderam [2014]). Investors feared the funds exposure to European banks and redeemed about $180 billion from the entire industry during the months from June to August 2011.

Motivated by these episodes, in Chapter 1 we study the contemporaneous flow dependence within the industry of US Money Market Mutual Funds and its resilience to widespread flow shocks. In particular, we empirically investigate the presence of peer effects within funds flows and their effect on the vulnerability of the industry.

Peer effects refer to the influence on an individual outcome of its peers outcomes, and

[1] The event of the NAV/Share of a money market mutual fund falling below $ 0.995 is usually referred to as “breaking the buck”.

have been studied in many different domains, such as school achievement (see Sacerdote [2001]). However, since the influential work of Manski [1993], peer effects have been divided in two categories: *endogenous* peer effects, i.e. the influence of peer outcomes, and *exogenous* peer effects, i.e. the influence of peer characteristics, leaving researchers with the complex challenge of distinguishing between the two effects in order to identify the drivers behind the peer outcome correlation. In the school achievement domain this amounts to understanding whether an individual test score is influenced by the achievement of its peers or by their exogenous characteristics such as social or economic background. This distinction becomes particularly important from a policy point of view because of the social multiplier entailed by endogenous peer effects: when endogenous peer effects are present, the average peer group outcome can be greater than the average outcome of the individuals if they where not interacting in groups.

In the context of money market mutual funds we define peers as those funds holding similar portfolios and peer (flow) effects are interpreted as the impact of a fund flow on the flows of similar funds. A possible mechanism behind peer flow effects are fire-sales spillovers. For instance, Falato et al. [2016] investigate peer effects in open-end fixed-income mutual funds and provide a detailed accounting of the chain of events behind the transmission of a flow shock between peers. A fund receiving unexpected outflows may be forced to sell securities. When forced sales happen, security prices may be depressed (see for instance Coval and Stafford [2007], Greenwood and Thesmar [2011]). In turn, this may hurt the performance of the funds holding the same securities causing them to experience new redemptions, because of the positive flow-performance relationship, and leading to a second round of forced sales.

We abstract from the mechanism behind peer effects and, instead, take on the empirical and theoretical challenge of identifying endogenous and exogenous peer flow effects. In particular, using a rich and granular dataset of security-level holdings of US MMMFs constructed from regulatory N-MFP filings submitted to the SEC between 2010 and 2016, we propose a spatial dynamic model where the spatial structure is given by a similarity temporal network based on fund portfolios. Similarly to Bramoullé et al. [2009],

we prove the theoretical condition for peer effects identification in terms of topological properties of the underlying network structure and find evidence of positive and statistically significant peer flow effects, implying that an initial flow shock to a fund can potentially be endogenously amplified through indirect impacts on similar funds. These indirect impacts are statistically and economically significant for first and second order neighbors and their magnitude is decreasing with the network distance between the two funds.

We also study how peer effects change the resiliency of the MMMFs industry to widespread flow shocks. In order to simplify the analysis we assume that a run-like event is characterized by two dimensions: the magnitude of a flow shock and the percentage of funds shocked. A run-like event is said to generate fire-sales spillovers if the percentage of funds forced to sell less liquid assets is above a given threshold. The resilience of the MMFs industry to run-like events can then be defined as its capacity to absorb widespread flow shocks without triggering fire sales spillovers. Intuitively, without peer effects, if every fund has 10% holding in liquid assets, a flow shock less than 10% will not lead to the sale of less liquid assets, independently of the percentage of funds receiving the flow shock. However, when peer effects are present, we find that the endogenous amplification of flow shocks creates a non linear relationship between the initial percentage of funds receiving a flow shock and final percentage of funds forced to sell less liquid assets, with a non trivial impact on the ability of the MMMFs industry to absorb widespread flow shocks. For instance, a flow shock below the percentage of daily liquid assets can still force some funds to sell less liquid assets if the percentage of funds initially shocked is high enough.

0.2 Longevity risk modeling

Life-insurance companies and annuity providers such as pension funds are exposed to a number of risks, including market, credit, lapse, longevity, operational. For insurance companies, the Solvency II regulation has been addressing both the definition, measure-

ment, monitoring and disclosure of such risks. As a result of both realized losses in the financial sector at large and of a general concern about potential crises and sudden crashes in the future, even the consciousness of risks has been growing. The number of monitoring tools and organizational checks apt to keep risks under control, as well as the number of deals aiming at transferring those risks, has been steadily growing [Cummins and Weiss, 2009].

Our main focus is on the risks coming from the liabilities of insurance companies. These liabilities are mainly subject to longevity risk, or improvements in the survivorship of annuity beneficiaries, and interest rate risk, since the level of interest rates affects the fair value of liabilities. Longevity risk is the inability to correctly forecast improvements in the survivorship of a population or, equivalently, from the point of view of a life insurance company, longevity risk is the risk the policy holders live longer than expected when the company priced and reserved their policies. So, while increasing longevity is welcome from the social point of view, it is a risk for annuity providers because it may cause them to face higher than expected pay-outs. Longevity risk can be remarkable if we consider that, in the last century, the life expectancy of individuals has been underestimated by as much as 3 years [International Monetary Fund (IMF), 2012].

The actuarial literature has seen the development of two broad classes of stochastic mortality models, namely, discrete and continuous time models. Discrete time stochastic mortality models have the advantage of being straightforward to simulate but do not provide a closed formula expression for the survival probabilities. One of the earliest model, that is now still used as a benchmark model and applied to data from many countries and time-periods, was proposed by Lee and Carter [1992] and models the log mortality rate of an individual aged x at time t as a function of age and period effects. Over the years many other discrete time mortality models have been introduced. For instance, Renshaw and Haberman [2006] propose an extension of the Lee-Carter model that incorporates a cohort (or year-of-birth) effect, and Cairns et al. [2006] introduces a two factor model allowing for mortality improvements that are non perfectly correlated across ages. The detailed discussion of discrete time mortality models is outside the

scope of this thesis, but the interested reader can find a review of both discrete and continuous time models in Cairns et al. [2008].

We place ourselves in the continuous time framework whose main advantage, over discrete time models, is the analytical tractability for the survival probability and the pricing of many insurance contracts. In particular, we model mortality risk by means of a Cox or a doubly stochastic counting process (see Milevsky and Promislow [2001] and Dahl [2004]). In particular the time-to-death of an individual, analogously to the time-to-default in credit risk literature, is assumed to be the first jump time of a Poisson process with stochastic intensity which is, on the other hand, modeled as a diffusion process. In addition, we assume the instantaneous mortality intensity process to be affine, leading to closed form expression for the survival probabilities and insurance contract prices [Luciano and Vigna, 2008].

In Chapter 2 we proposes dynamic hedging strategies of interest rate and longevity risk, when longevity basis risk is present. We model mortality intensity using a parsimonious model, that allows to disentangle the effects of basis risk. We consider the uncertainty in interest rates, that affects the fair-valuation of the liabilities of the annuity provider. We compare the effectiveness of different hedging strategies by comparing the risk margins that the annuity provider has to charge to keep the value-at-risk- (VaR-) based solvency of its hedged portfolio to a fixed 99.5% level. Our calibrated example, based on UK data, highlights that considering the uncertainty in interest rates affects the positions in longevity-linked instruments when hedging longevity exposures, and can significantly increase the variability of the hedging errors. However, we find that, at a long, 30-year horizon, 90% of the risk margin for an annuity contract on a 65-year old male is explained by the uncertainty in the longevity risk factor, while interest rate risk accounts only for 10% of the total risk margin.

In Chapter 3 we provide a method to assess the risk relief from a foreign expansion by a life-insurance company. We build a parsimonious continuous-time model for longevity risk, that captures the dependence across the different ages of two populations. We provide three measures of the diversification effects of expanding an annuity portfolio

internationally. The reduction in the risk margin, computed à la Solvency II, provides a regulation-consistent measure of the tail risk benefit. The change in the volatility of the average mortality intensity of a portfolio provides an intuitive measure of the change in its longevity risk. The Diversification Index provides a synthetic assessment of the diversification benefit of combining different populations. We calibrate the model to the case of a UK annuity portfolio expanding internationally towards Italian policyholders. Our application shows that the longevity risk diversification benefits of an international expansion are sizable, in particular when interest rates are low.

Part I

Peer effects and financial stability

Chapter 1

Peer Effects and Spillovers in Money Market Mutual Funds Flow[1]

1.1 Introduction

In recent years, prime money market mutual funds (MMMFs) have suffered several runs that were sector-wide and unfolded during the span of few weeks. For instance, during the week of September 15^{th} 2008, the default of Lehman Brothers caused the Reserve Primary Fund to break the buck, resulting in widespread redemptions - totaling about $300 billion - that affected also funds with no exposure to Lehman papers. One of the possible explanations for the vulnerability of the MMMF industry to the aggregate risk of run-like behaviours could be the amplification of an initial flow shock through fire-sales spillovers (Falato et al. [2016], Chernenko and Sunderam [2020]). Redemptions to one fund may induce the fund's portfolio managers to sell some of their assets, potentially depressing their price. This may, in turn, hurt the performance of other funds holding the same securities and lead to more redemptions because of the flow-performance rela-

[1]The material for this chapter is taken from De Rosa, La Spada, and Lillo [2020].

tionship. Due to the importance of MMMFs for financial stability, especially their role in providing short-term liquidity to financial institutions, these episodes have attracted the attention of many researchers (Schmidt et al. [2016], Kacperczyk and Schnabl [2013], Brady et al. [2012], Chernenko and Sunderam [2014]).

We empirically investigate the vulnerability of the MMMFs industry stemming from the endogenous amplification of an initial flow shock. Following Falato et al. [2016], we borrow from the applied literature on *peer effects* in order to study the contemporaneous dependence between the flow of a fund and the flows of its peers, defined as the funds with the most similar portfolios. Differently from Falato et al. [2016], however, we abstract from the mechanisms or chain of events that could give rise to peer effects[2] and focus, instead, on a two-fold objective. First, we construct a temporal network based on portfolio similarities that provides the spatial structure for our model and allow us to disentangle the impact of peer flows (endogenous peer effects[3]) from the impact of peer characteristics (exogenous peer effects), such as past yield. Second, we study how the presence of endogenous peer effects changes the MMMFs industry vulnerability with respect to run-like behaviours, where we say that the industry is vulnerable to the combination of an initial flow shock and a percentage of funds shocked if the final percentage of funds forced to sell less liquid assets is above a given threshold. In particular, we perform a simulation study to characterize the vulnerability region for different values of this threshold.

Using a dataset of detailed portfolio level holdings of US MMMFs coming from regulatory N-MFP filings, we find evidence of positive and statistically significant peer effects implying that an initial flow shock to a fund can potentially be endogenously amplified through indirect impacts on similar funds. These indirect impacts are statistically and economically significant for first and second order neighbors and their magnitude is decreasing with the network distance between the two funds. A fund may suffer a 0.75% flow shock from the indirect impact of a 1% shock to its peers. We also find that the

[2]For instance, fire sales spillovers or herding behaviors between investors with asymmetric information.

[3]If not specified, in the following we will denote endogenous peer effects simply with the term peer effects.

presence of peer effects creates a non linear relationship between the initial percentage of funds receiving a flow shock and final percentage of funds forced to sell less liquid assets, with a non trivial impact on the ability of the MMMFs industry to absorb widespread flow shocks. For instance, a flow shock below the percentage of daily liquid assets can still force some funds to sell less liquid assets if the percentage of funds initially shocked is high enough.

These results are robust to different methodologies and measures for the construction of the similarity temporal network as well as to the introduction of fund and time fixed effects. Moreover, we employ different strategies to address the endogeneity concerns typical in peer effects models. Specifically, we solve the reflection problem [Manski, 1993] regarding the identification of endogenous and exogenous peer effects by providing, similarly to Bramoullé et al. [2009], an identification condition based on the topological properties of the underlying similarity temporal network. In order to rule out the possibility of common portfolio characteristics driving our results, we also test their robustness with respect to the inclusion of controls for the portfolio weighted average maturity and credit risk exposure. Finally, we tackle the possibility that unobserved common components could be driving our results by performing a quasi-random experiment that exploits, as a source of exogenous variation, the different exposures of money market funds to European banks during summer 2011.

Our work is related to two main strands of the literature. First, we contribute to growing literature on peer effects and fire-sales spillovers in the mutual fund industry. Coval and Stafford [2007], Ellul et al. [2011] and Feldhütter [2012] find evidence of significant price impact in equity and corporate debt markets for instruments under selling pressure by large mutual funds or insurance companies. Chernenko and Sunderam [2020] provide evidence of meaningful fire sales externalities in the US domestic equity mutual fund industry, showing that funds internalizing more the pricing impact of their trading use their cash holdings more aggressively in order to manage their inflows and outflows. Falato et al. [2016] investigate peer effects in open-end fixed-income mutual funds and provide a detailed accounting of the chain of events behind the transmission of a flow

shock between peers. We contributes to this literature by extending the evidence on peer flow effects to the MMMFs industry and by tackling the identification of endogenous and exogenous peer effects which, to the best of our knowledge, has not been studied yet in the context of mutual funds. Second, we contribute to the recent literature on the riskiness of the MMMFs industry (see Schmidt et al. [2016], Strahan and Tanyeri [2015], McCabe [2010]) by investigating how the endogenous amplification of flow shocks caused by peer effects changes the vulnerability of the industry to run-like behaviours.

The chapter is organized as follows. Section 1.2 provides a general description of Money Market Mutual Funds and their regulatory framework while Section 1.3 details the source and characteristics of the data set used. Section 1.4 introduces our model and Section 1.5 describe our methodology for the construction of the similarity temporal network. Section 1.6 describes the main empirical results while Section 1.7 analyzes their financial stability implications. Finally, Section 1.8 concludes.

1.2 Money Market Mutual Funds

Money market mutual funds (MMMFs) are open-ended mutual funds that invest in short-term, high credit-quality, money-market instruments. MMMFs are an important part of the financial system. As of the end of June 2020, MMMFs had about $5 trillion in assets under management (AUM). They are a key source of short-term financing for financial institutions [Hanson et al., 2015] and are the largest cash lender in the repo market [Afonso et al., 2020].

MMMFs can be categorized into four types according to their investment universe. Treasury MMMFs can invest only in US Treasuries or repurchase agreements (repos) collateralized by Treasuries; Government MMMFs can also invest in agency debt or repos collateralized by agency debt; Prime MMMFs can also buy private debt, both secured (e.g., repos) and unsecured (e.g., certificates of deposit); and Municipal MMMFs can only invest in short-term debt issued by states and other local authorities. Finally, based on the profile of their investors, MMMFs share classes can be divided into institutional and

retail.

Similarly to other mutual funds, MMMFs levy fees as a fixed percentage of their AUM; as a result, they are subject to the tournament-like incentives generated by the positive flow-performance relationship observed in the data (Kacperczyk and Schnabl [2013]; La Spada [2018]). However, in contrast to other mutual funds, until October 2016, all MMMFs were allowed to maintain a stable net asset value (NAV) of $1.00 per share; they did so by valuing assets at amortized cost and distributing daily dividends as securities approach their maturity date. Since their shares are not insured by the government and are daily redeemable, this stable-NAV feature makes MMMFs susceptible to runs. If a fund "breaks the buck," i.e., the market value of the fund's NAV drops below $0.995, the fund manager has to disclose it to investors and reprice the fund shares. Such event gives investors a strong incentive to redeem their shares en masse to preserve the value of their capital, as it happened on September 16, 2008, when the Reserve Primary Fund broke the buck after writing off Lehman Brothers debt. The 2008 run on Reserve Primary Fund quickly spread to the whole prime and municipal money market industry: investors redeemed more than $300 billion within a few days after Lehman's default. In the summer of 2011, another run hit the prime fund industry; amid the European debt crisis, investors became concerned of funds' exposure to troubled European banks and redeemed prime fund shares for more than $170 billion in less than two months. Both episodes caused a severe shortage of credit to the banking sector (both in the US and in Europe) and large disruptions in the money markets more generally.

In response to the 2008 financial crisis, the SEC adopted in 2010 a first set of amendments to Rule 2a-7 of the Investment Company Act that were mainly focused on improving the quality and the risk exposure of MMMFs portfolios. The main characteristics of this new regulation can be summarized into four main points. First, in order to reduce the credit risk exposure of MMMFs, the limit of their holdings in second tier securities has been reduced from 5% to 3% of the portfolio, and the concentration limit in any issuer of such securities set to 0.5% of the fund assets. Second, the maximum weighted average maturity and life of money market funds' portfolios have also been reduced to 60

and 120 days, respectively. Third, taxable MMMFs were required to maintain at least 10% of their holdings in daily liquid assets and all funds were required to maintain at least 30% of their total assets in weekly liquid assets. Fourth, in order to enhance the SEC oversight ability, money market funds were required to provide a monthly electronic filing of detailed portfolio holdings information using a new form called N-MFP Form, through the EDGAR system. All of these regulatory constraints were in effect during our analysis period, that goes from November 2010 to September 2015. However, it is worth noticing that in July 2014 a new set of structural changes to the MMMFs industry was approved, but came in effect only 2016. These changes introduced the ability, by MMMFs, to impose liquidity fees and redemption gates in certain circumstances. Under the new regulation, if the weekly liquid assets of a fund fall below 30% of its total assets, the fund can[4] impose a 2% liquidity fee or suspend redemptions for up to 10 business days in a 90 day period. Moreover, if the percentage of weekly liquid assets falls below 10%, then the fund is required[5] to impose a 1% liquidity fee. Furthermore, MMMFs other than Government and Retail funds were required to transact at a Floating Net Asset Value rather than a Stable NAV and to be valued at market value rather than amortized cost. The shift towards a floating net asset value has the aim of eliminating, for the affected funds, the risk of breaking the buck and, therefore, reducing the incremental redemption incentive for investors.

1.3 Data Set description

Our data come from the regulatory N-MFP filings that all US MMMFs have to submit monthly to the Securities and Exchange Commission (SEC). The dataset covers the whole universe of US MMMFs and contain detailed fund-level, share-class level and portfolio-level information as of the end of each month.

The fund level data contains information such as total net asset value of the fund,

[4] If the fund's board of directors determines that such choice would be in the fund's best interest.

[5] Unless the fund's board of directors determines that imposing such liquidity fee would not be in the fund best interest.

fund name, adviser name, the annualized 7-day gross yield for the month, the fund type, the end of month fund portfolio weighted average maturity and dummy variables indicating whether a fund is merging, liquidating or in a master-feeder relationship. Each fund can have different share classes for which the fund reports additional information such as the share class minimum initial investment, the share class monthly redemptions and subscriptions and a variable indicating whether the share class is for institutional or retail investors.

For each filing date of a given fund, we also have access to the full portfolio composition of the fund. For each security in the portfolio, we know the cusip, the issuer name, the instrument category, its principal, its yield, its maturity, its fair value and amortized cost valuation. The total number of unique securities contained in the dataset is $246,676$ corresponding to $37,933$ unique borrowers.

The dataset covers a total of 715 unique money market mutual funds. Within our observation period, funds are closed, reclassified, merged and new funds are opened. Hence, the number of funds reporting each month is not constant. For instance, in January 2011 we observe reporting for 667 unique funds, while in November 2016 this number is 414.

1.4 Peer Effects

Run-like events are sudden and the bulk of redemptions may unfold in the span of very few days (see Schmidt et al. [2016]). Hence, in order to assess the potential vulnerability of the MMMFs industry to the risk of runs, we focus on the study of the contemporaneous dependence among fund flows, that we call *endogenous peer (flow) effects* because we borrow from the methodologies of the applied literature on peer effects.

To identify the simultaneous dependence among MMMF flows, we propose a spatial dynamic model where the spatial structure is given by a similarity network based on fund portfolios. The idea is that a fund's flow may have a stronger impact on MMMFs with similar portfolios. As shown in Falato et al. [2016], this may happens because funds

under flow pressure may be forced to sell securities, depressing their prices and hurting the performances of other funds holding the same instruments which, in turn, could suffer new outflows because of a positive flow-performance relationship.[6]

More precisely, we study endogenous peer effects by investigating the relationship between the flow at time t of fund i, y_{it}, and the average of contemporaneous flows of similar funds. The average peer flow is computed among the first degree neighbors of i given the network W_t. Since there is empircal evidence supporting a dependence between the previous performance of a fund the its current flow (see La Spada [2018]), we use the lagged own performance as an exogenous regressor, but we also include an *exogenous peer group effect* given by the average past performance of fund i first degree neighbors a time $t-1$.

Hence, our baseline regression is:

$$\boldsymbol{y}_t = \alpha_t \boldsymbol{\iota} + \boldsymbol{\zeta}_n + \lambda W_t \boldsymbol{y}_t + \beta_1 W_{t-1} \boldsymbol{x}_{t-1} + \beta_2 \boldsymbol{x}_{t-1} + \boldsymbol{\mathcal{C}}_t \boldsymbol{\delta}' + \boldsymbol{\epsilon}_t, \tag{1.1}$$

where $\boldsymbol{y}_t$ is the $N \times 1$ vector containing the cross-sectional flows at time t, W_t is the $N \times N$ adjacency matrix at time t, $\boldsymbol{x}_{t-1}$ is the $N \times 1$ vector containing the prior month funds' yield, $\boldsymbol{\mathcal{C}}_t$ is a $N \times K$ matrix of control variables, α_t and $\boldsymbol{\zeta}_n$ are time and fund fixed effects, while $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{it}, \ldots, \epsilon_{Nt})'$ is a vector of disturbances. As standard in the literature, we use as controls the logarithm of the fund size and fund's family size expressed in millions. Moreover, as a robustness check, we also consider an extended version of model (1.1) which includes as exogenous regressors the past flow of a fund and the average flow at time $t-1$ of similar funds, where the average is computed among the first degree neighbors given by the network W_{t-1}:

$$\boldsymbol{y}_t = \alpha_t \boldsymbol{\iota} + \boldsymbol{\zeta}_n + \lambda W_t \boldsymbol{y}_t + \rho \boldsymbol{y}_{t-1} + \gamma W_{t-1} \boldsymbol{y}_{t-1} + \beta_1 W_{t-1} \boldsymbol{x}_{t-1} + \beta_2 \boldsymbol{x}_{t-1} + \boldsymbol{\mathcal{C}}_t \boldsymbol{\delta}' + \boldsymbol{\epsilon}_t. \tag{1.2}$$

Defining $P_{i,t}$ as the set of fund i's neighbors derived from the network structure at time

[6] It is worth noticing that a different, but possible, explanation for this phenomenon could be investors following the behavior of other investors in similar funds because of incomplete and asymmetric information. However, we do not aim at disentangling this two mechanisms.

t, and assuming that the element w_{ij}^t of W_t is $\frac{1}{|P_{i,t}|}$ $\forall j \in P_{i,t}$ and zero otherwise, then the i-th component of the *endogenous peer effect* term $W_t \boldsymbol{y}_t$ can be rewritten as

$$\frac{\sum_{j \in P_{i,t}} y_{j,t}}{|P_{i,t}|}, \tag{1.3}$$

and interpreted as the average simultaneous flows of the peers. Similarly, the *exogenous peer effects* $W_{t-1}\boldsymbol{y}_{t-1}$ and $W_{t-1}\boldsymbol{x}_{t-1}$ can be read as the average $t-1$ flow and yield of the peers, respectively.

1.4.1 Identification: Reflection Problem

It is well known that identification of endogenous peer group effects may be hindered by the so called *reflection problem* [Manski, 1993]. One instance of the reflection problem is that it may be difficult to distinguish between endogenous and exogenous peer effects, due to the perfect collinearity introduced by the simultaneous dependence between the flows. In other words, it may be difficult to assess whether the "similarity" between the flows of a fund i and the flows of its peer group is due to the influence of the peers or due to the fact that similar funds may have similar characteristics driving the flows. To prove identification of our model we follow the approach of Bramoullé et al. [2009], which studies the identification of models with peer-group effects when links follow a social network structure. They consider a simple model with contemporaneous endogenous and exogenous peer group effects and prove that identification is achieved if simple conditions on the network structure are met. These identification conditions cannot be directly applied to our case, since in our model endogenous and exogenous peer group effects are determined by two different network structures or, more precisely, by two different time instants of the temporal network. Hence, we prove new sufficient conditions for identification based on the topological properties of the underlying temporal network.

Model (1.2) can be rewritten in reduced form as :

$$\begin{aligned} \boldsymbol{y}_t = \alpha_t(\boldsymbol{I} - \lambda W_t)^{-1}\boldsymbol{\iota} + (\boldsymbol{I} - \lambda W_t)^{-1}\boldsymbol{\zeta}_n + \\ + (\boldsymbol{I} - \lambda W_t)^{-1}(\rho\boldsymbol{I} + \gamma W_{t-1})\boldsymbol{y}_{t-1} + \\ + (\boldsymbol{I} - \lambda W_t)^{-1}(\beta_2\boldsymbol{I} + \beta_1 W_{t-1})\boldsymbol{x}_{t-1} + \\ + (\boldsymbol{I} - \lambda W_t)^{-1}\boldsymbol{\mathcal{C}}_t\boldsymbol{\delta}' + (\boldsymbol{I} - \lambda W_t)^{-1}\boldsymbol{\epsilon}_t, \end{aligned} \tag{1.4}$$

and we say that the structural model (1.2) is not identifiable if we can find two sets of different structural parameters that lead to the same reduced form model (1.4). To prove identificability of the model, we need the following definitions:

Definition 1. *We say that the nodes (i, j, k) form an Intransitive Temporal Triad at time t if j is influenced by k at time $t-1$, i is influenced by j at time t, but i is not influenced by k neither at time $t-1$ nor at time t (see Figure 1.1).*

Definition 2. *We say that the temporal network $\{W_t\}$ is row-normalized if $\forall t$ the row sums of W_t are 1.*

Definition 3. *We say that the temporal network $\{W_t\}$ has no isolated fund if each fund always belongs to a non empty peer group or, equivalently, W_t has no zero rows $\forall t$.*

Definition 4. *We say that the temporal network $\{W_t\}$ is non trivial if $\exists t$ such that $W_t \neq W_{t-1}$.*

A path between two funds i and j is defined as a sequence of distinct funds

$$k_1, \ldots, k_s, \ldots, k_m$$

where k_1 is influenced by i, k_s is influenced by k_{s-1}, and j is influenced by k_m. The length of a path $\{i, k_1, \ldots, k_s, \ldots, k_m, j\}$ is defined as the number of similarity links it contains. The distance between two funds i and j is defined as the length of the shortest path connecting them. The diameter of a network is then defined as the greatest distance

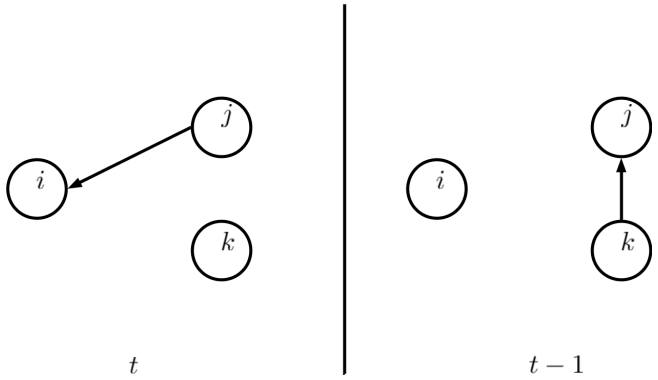


Figure 1.1: Graphical representation of an Intransitive Temporal Triad as described in Definition 1.

between any two pairs of funds. We can now prove the following sufficient condition for identification for model (1.2):

Proposition 1.4.1. *Suppose that the temporal network* $\{W_t\}$ *is non trivial, row-normalized, has no isolated fund and that* $\rho\gamma \neq 0$ *or* $\beta_2\beta_2 \neq 0$.
If there exists a t *such that the network* $W_t^2 W_{t-1}$ *has diameter greater or equal than* 3, *then the model* (1.2) *is identified.*

Proof. See Appendix 1.A.3. □

If we consider the case without fixed effects:

$$\boldsymbol{y}_t = \alpha\boldsymbol{\iota} + \lambda W_t\boldsymbol{y}_t + \rho\boldsymbol{y}_{t-1} + \gamma W_{t-1}\boldsymbol{y}_{t-1} + \beta_1 W_{t-1}\boldsymbol{x}_{t-1} + \beta_2\boldsymbol{x}_{t-1} + \boldsymbol{\mathcal{C}}_t\boldsymbol{\delta}' + \boldsymbol{\epsilon}_t, \qquad (1.5)$$

we can also prove the following:

Proposition 1.4.2. *Suppose that the temporal network* $\{W_t\}$ *is non trivial, row-normalized, has no isolated fund and that* $\rho\gamma \neq 0$ *or* $\beta_2\beta_2 \neq 0$.
If there exists at least one Intransitive Temporal Triad, then model (1.5) *is identified.*

Proof. See Appendix 1.A.1. □

1.4.2 Identification: Correlated Effects

Besides the reflection problem, another identification issue that could hinder the estimation is the presence of correlated effects, i.e., unobserved shocks that may be common to funds that hold similar portfolios. The inclusion of time fixed effects partly helps in reducing this problem but does not eliminate it. Hence, we address this concern in two ways. First, we consider a version of model that includes controls for own and peers portfolio characteristics, such as weighted average portfolio maturity and portfolio credit risk exposure. The weighted average portfolio maturity of fund i at time t, $WAM_{i,t}$, is defined as:

$$WAM_{i,t} = \sum_{k \in \mathcal{K}_{i,t}} \omega_{i,t}^{k} m_{i,t}^{k}, \tag{1.6}$$

where $\mathcal{K}_{i,t}$ is the set of securities in fund i portfolio at time t, $\omega_{i,t}^{k}$ is the weight of security k at time t and $m_{i,t}^{k}$ is its maturity in days. Similarly, the peer weighted average portfolio maturity for fund i at time t, $PeerWAM_{i,t}$, is defined as:

$$PeerWAM_{i,t} = \frac{\sum_{j \in P_{i,t}} WAM_{j,t}}{|P_{i,t}|}. \tag{1.7}$$

In order to measure a fund i credit risk exposure we use, as a proxy, the portfolio weighted average credit default spread, $CDS_{i,t}$. In particular, let $\mathcal{B}_{i,t}$ be the set of all issuers in the portfolio of fund i at time t and, for $b \in \mathcal{B}_{i,t}$, let $s_{b,t}$ be the average credit default spread of issuer b for the month t. Then $CDS_{i,t}$ is defined as:

$$CDS_{i,t} = \sum_{b \in \mathcal{B}_{i,t}} \omega_{i,t}^{b} s_{b,t}, \tag{1.8}$$

where $\omega_{i,t}^{b}$ is the weight at time t of issuer b in the portfolio of fund i. Similarly, the peer portfolio weighted average credit default spread, $PeerCDS_{i,t}$, is defined as:

$$PeerCDS_{i,t} = \frac{\sum_{j \in P_{i,t}} CDS_{j,t}}{|P_{i,t}|}. \tag{1.9}$$

Our second strategy to address this endogeneity problem is the design a quasi-random experiment exploiting the different portfolio exposures of each fund to European banks during summer 2011. The intuition behind this choice is based on the anecdotal evidence that, during the peak of the 2011 European debt, MMMFs with high exposure to European banks suffered large outflow shocks. Identification is achieved by showing that fund flows are significantly correlated with an exogenous characteristic of the peers, their exposure to European banks, that is relevant for peer flow but otherwise random for fund flows. We define a *peer treatment* flow variable, $W_t \boldsymbol{y}_t^{Treatment}$, whose i-th element is given by:

$$\frac{\sum_{j \in P_{i,t}} y_{j,t} \mathbb{1}_{\{Treatment=1\}}}{|P_{i,t}|}, \tag{1.10}$$

where the treatment variable is equal to 1 for the funds that, from May to December 2011, were in the top quartile with the highest exposure to European banks. In Appendix 1.B.1 we then estimate model (1.2) by substituting the peer flow variable $W_t \boldsymbol{y}_t$ with its treatment counterpart.

1.5 Portfolio Similarity Measure and Similarity Temporal Network

The definition of a similarity measure between funds' portfolios is not a trivial task and inherently entails the choice of the dimensions along which similarities are more relevant. For instance, two funds may have portfolios that are similar in terms of instrument types allocation but with completely different borrowers or, on the other hand, they may have portfolios similar in terms of borrowers but with very different maturities. In our application, we measure similarities by aggregating the securities in each portfolio at the borrower level. However, in order to reduce the information loss at the maturity level, we also adjust each borrower exposure for the corresponding maturities. Specifically, let $\mathcal{B}$ be the set of all possible borrowers and let $|\mathcal{B}| = B$. For each fund i and borrower

$b \in \mathcal{B}$, we define the *Weighted Maturity Exposure* of i to b at time t as:

$$\mathrm{WME}_{i,t}^{b} = \sum_{k \in \mathcal{K}_{i,t}^{b}} \omega_{i,t}^{k} m_{i,t}^{k}, \tag{1.11}$$

where $\mathcal{K}_{i,t}^{b}$ is the set of instruments in the portfolio of i at time t whose issuer is b, $\omega_{i,t}^{k}$ is the weight of that instrument and $m_{i,t}^{k}$ its effective maturity. Then, for each fund i we have the B-dimensional vector

$$\mathrm{WME}_{i,t} = \left(\mathrm{WME}_{i,t}^{1}, \mathrm{WME}_{i,t}^{2}, \ldots, \mathrm{WME}_{i,t}^{B}\right). \tag{1.12}$$

of the weighted maturity exposures. If i and j are two funds, we define the *Weighted Maturity Similarity* between them at time t, denoted $\mathrm{WMS}_t(i,j)$, as the cosine similarity between their vectors of weighted maturity exposure, i.e.:

$$\mathrm{WMS}_t(i,j) = \frac{\mathrm{WME}_{i,t} \cdot \mathrm{WME}_{j,t}}{||\mathrm{WME}_{i,t}||\,||\mathrm{WME}_{j,t}||}. \tag{1.13}$$

By definition, $\mathrm{WMS}_t(i,j) = 1$ if and only if i and j are exposed to the same borrowers with the same weighted maturity exposures. Also, $\mathrm{WMS}_t(i,j) = 0$ if and only if i and j do not have any borrower in common at time t. As shown in Figure 1.2, Treasury Funds exhibit the highest similarity, with an average value of 0.94, followed by the Government Agency Funds with an average pairwise similarity of 0.61. Then we have the Prime Funds with an average weighted maturity similarity of 0.31 and Single State Funds with an average similarity of only 0.04. Thus, the similarity within fund categories seems to decrease with the increase in the size of the available investable universe. Given that different fund categories can invest in different types of asset, we also have that the cross similarity between funds of different types is low with an high probability mass at zero, as shown in Table 1.1. Figure 1.3 shows that the evolution over time of the average WMS within fund categories. There is a slight increase over time for the average similarity of Treasury funds, while Prime funds show a small but steady decline in similarity from the second half of 2012. Interestingly, we observe an increase in the average similarity

Table 1.1: Average cross similarity between fund categories where the similarity is measured with the Weighted Maturity Similarity measures defined in (1.13).

	Treasury	Government	Prime	Single State
Treasury	0.94	0.31	0.37	0.003
Government	0.31	0.61	0.25	0.004
Prime	0.37	0.25	0.31	0.005
Single State	0.003	0.004	0.005	0.04

of government funds from the beginning of 2013 up to the first half of 2014.

Given the weighted maturity similarity introduced before, we are now able to construct the Similarity Temporal Network. In order to do so, we restrict our focus only on the funds that are always present throughout the entire period of our analysis. This leaves us with a subset of 363 MMMFs.

At each time t, we say that there is a link from a fund i to a fund j, i.e. j belongs to the Peer Group of i, if and only if j belongs to the 10% of funds with the highest weighted maturity similarity to i. In Appendix 1.B we show the robustness of our results to the choice of this threshold. For each fund i the outgoing links have weights normalized to sum 1:

$$w_{i,j} = \frac{1}{k_i^{out}}, \tag{1.14}$$

where k_i^{out} is out-degree of node i. Hence for each period t, we construct a directed network with adjacency matrix W_t, that is row-normalized by (1.14). The temporal network is then given by the collection of adjacency matrices over time:

$$\mathcal{W} = \{W_1, W_2, \ldots, W_T\}, \tag{1.15}$$

where T is the length of the time series. The in-degree of a fund i can be interpreted as the number of funds that could potentially be affected by the flows of fund i, while its out-degree represents the number of funds whose flows could potential spillover to i. In our analysis, peers are those funds that bear an high degree of similarity in their

holdings according to the weighted maturity similarity.

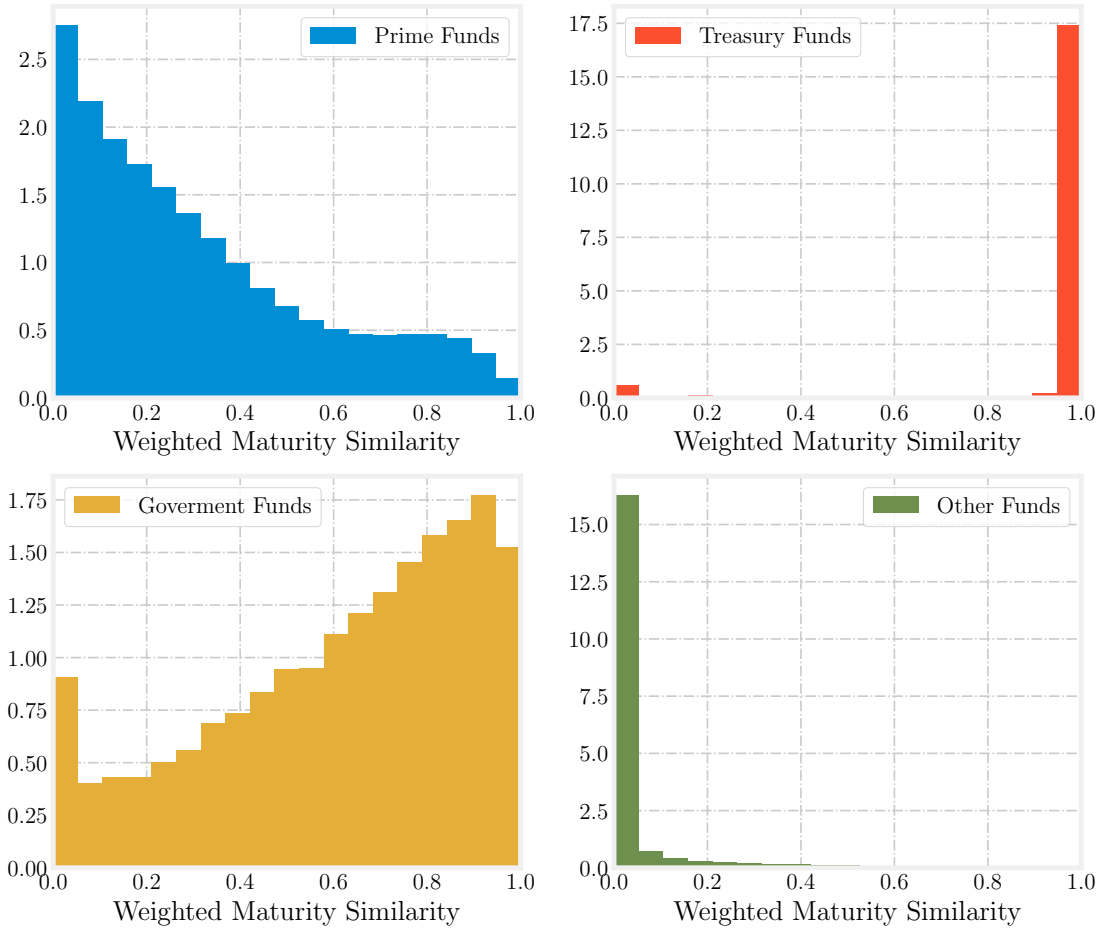


Figure 1.2: Histogram of the Weighted Maturity Similarity within each MMMF category.

1.6 Estimation results

The empirical application uses monthly data[7] from regulatory filings of US Money Market Mutual funds, from January 2011 to September 2015, and as usual in the literature,

[7]See Section 1.3 for more details about the data-set.

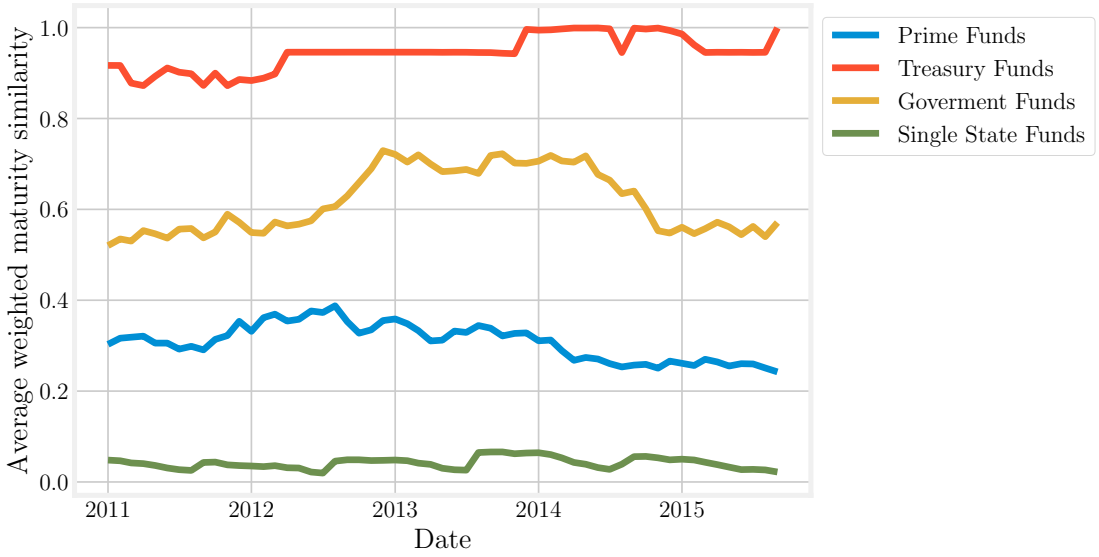


Figure 1.3: Evolution over time of the average Weighted Maturity Similarity measure within each MMMF category.

we estimate the dependent variable $Flow_{i,t}$ as

$$Flow_{i,t} = \frac{TNA_{i,t} - (1 + r_{i,t})TNA_{i,t-1}}{TNA_{i,t-1}}, \tag{1.16}$$

where $r_{i,t}$ is the yield of fund i during the period $(t-1, t)$ and $TNA_{i,t}$ is the total net asset value of fund i at time t. In order to reduce the impact of outliers we trim the data on the dependent variable using the $0.5th$ and $99.5th$ percentiles as cutoff. Moreover, to keep the data-set balanced, for each extreme value, we remove all the observations for that fund. Our main focus is on the universe composed of all funds except Treasury money market funds and the corresponding results are those we will maily describe. However, we also estimate the model and present the results when considering all funds and only Prime funds.

Table 1.2 contains basic statistics about the data. The dependent variable $Flow$ is expressed in percentage terms, the monthly fund $Yield$ is instead expressed in basis points, while the control variables $LogFundSize$ and $LogFamilySize$ are defined as the natural logarithm of the fund or fund family size expressed in millions. Model (1.2) can

Table 1.2: Main statistics of data set from monthly regulatory filings of US Money Market Mutual funds covering the period from January 2011 to September 2015.

	count	mean	std	min	25%	50%	75%	max
$Flow\,(\boldsymbol{y})$	17442	−0.28	5.35	−25.96	−2.51	−0.55	1.64	36.66
$Yield\,(\boldsymbol{x})$	17442	1.44	0.73	0.00	0.92	1.38	1.91	5.98
$WFlow\,(\boldsymbol{Wy})$	17442	−0.21	1.65	−6.70	−1.16	−0.21	0.71	8.04
$WYield\,(\boldsymbol{Wx})$	17442	1.41	0.49	0.38	1.02	1.37	1.75	3.21
$LogFundSize$	17442	7.33	1.84	2.83	5.89	7.15	8.72	11.81
$LogFamilySize$	17442	10.31	2.15	4.09	8.69	10.87	11.92	13.03

be rewritten in stacked form as:

$$\boldsymbol{y} = \boldsymbol{\alpha} \otimes \boldsymbol{\iota} + \boldsymbol{\iota_T} \otimes \boldsymbol{\zeta}_n + \lambda W\boldsymbol{y} + \rho\boldsymbol{y}_{-1} + \gamma W_{-1}\boldsymbol{y}_{-1} + \beta_1 W_{-1}\boldsymbol{x}_{-1} + \beta_2\boldsymbol{x}_{-1} + \boldsymbol{\mathcal{C}\delta} + \boldsymbol{\epsilon}, \quad (1.17)$$

where $\boldsymbol{y}, \boldsymbol{y}_{-1}, \boldsymbol{x}_{-1}$ and $\boldsymbol{\epsilon}$ are TN column vectors, $\boldsymbol{\mathcal{C}}$ is a $TN \times K$ matrix, and W and W_{-1} are block diagonal $TN \times TN$ matrices. Using the two demeaning operators defined in Appendix 1.A:

$$B = F'_{T,T-1} \otimes \boldsymbol{I}_N, \quad (1.18)$$

$$(\boldsymbol{I} - H) = (\boldsymbol{I} - \boldsymbol{I}_T \otimes \frac{1}{N}(\boldsymbol{\iota}_N\boldsymbol{\iota}'_N)), \quad (1.19)$$

we can rewrite the model after removing the fixed effects as:

$$\begin{aligned} B(\boldsymbol{I} - H)\boldsymbol{y} =& \lambda B(\boldsymbol{I} - H)W\boldsymbol{y} + \rho B(\boldsymbol{I} - H)\boldsymbol{y}_{-1} + \gamma B(\boldsymbol{I} - H)W_{-1}\boldsymbol{y}_{-1} + \\ &+ \beta_2 B(\boldsymbol{I} - H)\boldsymbol{x}_{-1} + \beta_1 B(\boldsymbol{I} - H)W_{-1}\boldsymbol{x}_{-1} + B(\boldsymbol{I} - H)\boldsymbol{\mathcal{C}\delta} + \boldsymbol{\nu}^*. \end{aligned} \quad (1.20)$$

As already observed the variable $B(\boldsymbol{I} - H)W\boldsymbol{y}$ is endogenous but, due to the demeaning operation, also $B(\boldsymbol{I} - H)\boldsymbol{y}_{-1}$ and $B(\boldsymbol{I} - H)W_{-1}\boldsymbol{y}_{-1}$ may be correlated with the error term. Hence, we need instruments for $[B(\boldsymbol{I} - H)W\boldsymbol{y},\ B(\boldsymbol{I} - H)\boldsymbol{y}_{-1},\ B(\boldsymbol{I} - H)W_{-1}\boldsymbol{y}_{-1}]$. Let $\boldsymbol{X} = [\boldsymbol{x}_{-1}, W_{-1}\boldsymbol{x}_{-1}, \boldsymbol{\mathcal{C}}]$ be the matrix of all exogenous regressors. It is easy to

verify that $\boldsymbol{I}, W, W_{-1}$, WW_{-1}, W^2 and W^2W_{-1} are linearly independent, hence the parameters are identified and we can use $B(\boldsymbol{I}-H)W\boldsymbol{X}$, $B(\boldsymbol{I}-H)W^2\boldsymbol{X}$,..., as valid instruments for $B(\boldsymbol{I}-H)W\boldsymbol{y}$. As common in the dynamic panel literature, and following Lee and Yu [2014], $(\boldsymbol{I}-H)\boldsymbol{y}_{-1}$, $W(\boldsymbol{I}-H)\boldsymbol{y}_{-1}$, $W^2(\boldsymbol{I}-H)\boldsymbol{y}_{-1}$,..., and $(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}$, $W(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}$, $W^2(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}$,..., can be instruments for $B(\boldsymbol{I}-H)\boldsymbol{y}_{-1}$ and $B(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}$.

Therefore, in order to estimate model (1.20) we use a 2SLS procedure that can be seen a special case of the approach proposed in Lee and Yu [2014].

Let $\boldsymbol{Z} = \big[B(\boldsymbol{I}-H)W\boldsymbol{y},\, B(\boldsymbol{I}-H)\boldsymbol{y}_{-1},\, B(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}, B(\boldsymbol{I}-H)\boldsymbol{X}\big]$, and define the instrument matrix

$$\begin{aligned}\boldsymbol{Q} = \big[&(\boldsymbol{I}-H)\boldsymbol{y}_{-1}, W(\boldsymbol{I}-H)\boldsymbol{y}_{-1}, W^2(\boldsymbol{I}-H)\boldsymbol{y}_{-1}, \ldots, \\ &(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}, W(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}, W^2(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}, \ldots, \qquad (1.21)\\ &B(\boldsymbol{I}-H)\boldsymbol{X}, B(\boldsymbol{I}-H)W\boldsymbol{X}, B(\boldsymbol{I}-H)W^2\boldsymbol{X}, ...\big]. \qquad (1.22)\end{aligned}$$

Thus if we define $\boldsymbol{P}_Q = \boldsymbol{Q}(\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'$ and $\boldsymbol{y}^* = B(\boldsymbol{I}-H)\boldsymbol{y}$, then

$$\widehat{\theta} = (\boldsymbol{Z}'\boldsymbol{P}_Q\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{P}_Q\boldsymbol{y}^*. \qquad (1.23)$$

When estimating the model without fixed effects or with only time fixed effect we use, instead, the Generalized Spatial 2SLS method introduced by Kelejian and Prucha [1998] and further developed by Lee [2003] which yield an asymptotically optimal estimator when the error are i.i.d.. For instance, let consider the model with only time fixed effects

$$\boldsymbol{y} = \boldsymbol{\alpha}\otimes\boldsymbol{\iota} + \lambda W\boldsymbol{y} + \rho\boldsymbol{y}_{-1} + \gamma W_{-1}\boldsymbol{y}_{-1} + \beta_1 W_{-1}\boldsymbol{x}_{-1} + \beta_2\boldsymbol{x}_{-1} + \boldsymbol{\mathcal{C}}\boldsymbol{\delta} + \boldsymbol{\epsilon}, \qquad (1.24)$$

and demeaned representation

$$(\boldsymbol{I}-H)\boldsymbol{y}=\lambda(\boldsymbol{I}-H)W\boldsymbol{y}+\rho(\boldsymbol{I}-H)\boldsymbol{y}_{-1}+\gamma(\boldsymbol{I}-H)W_{-1}\boldsymbol{y}_{-1}+ \\ +\beta_1(\boldsymbol{I}-H)W_{-1}\boldsymbol{x}_{-1}+\beta_2(\boldsymbol{I}-H)\boldsymbol{x}_{-1}+(\boldsymbol{I}-H)\boldsymbol{\mathcal{C}}\boldsymbol{\delta}+\boldsymbol{\nu}. \tag{1.25}$$

In our case the Generalized Spatial 2SLS reduces to a two-step procedure. Let define the matrix of all exogenous regressors $\boldsymbol{X}=\big[\boldsymbol{y}_{-1},W_{-1}\boldsymbol{y}_{-1},\boldsymbol{x}_{-1},W_{-1}\boldsymbol{x}_{-1},\boldsymbol{\mathcal{C}}\big]$ and let $\boldsymbol{Z}=\big[(\boldsymbol{I}-H)W\boldsymbol{y},(\boldsymbol{I}-H)\boldsymbol{X}\big]$. During the first step of the procedure, we estimate a 2SLS using as instruments $\widetilde{\boldsymbol{Q}}=\big[(\boldsymbol{I}-H)\boldsymbol{X},(\boldsymbol{I}-H)W\boldsymbol{X},(\boldsymbol{I}-H)W^2\boldsymbol{X}\big]$. Let $\boldsymbol{P}_{\widetilde{\boldsymbol{Q}}}=\widetilde{\boldsymbol{Q}}(\widetilde{\boldsymbol{Q}}'\widetilde{\boldsymbol{Q}})^{-1}\widetilde{\boldsymbol{Q}}'$, then we get a consistent estimator given by $\widetilde{\boldsymbol{\theta}}=(\boldsymbol{Z}'\boldsymbol{P}_{\widetilde{\boldsymbol{Q}}}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{P}_{\widetilde{\boldsymbol{Q}}}\boldsymbol{y}$.
Using this first step estimate and the reduced form of (1.25) we can compute

$$\mathbb{E}\big[(\boldsymbol{I}-H)W\boldsymbol{y}(\widetilde{\boldsymbol{\theta}})|\boldsymbol{X}\big]=(\boldsymbol{I}-H)W(\boldsymbol{I}-\widetilde{\lambda}W)^{-1}(\widetilde{\rho}\boldsymbol{I}+\widetilde{\gamma}W_{-1})\boldsymbol{y}_{-1}+ \\ +(\boldsymbol{I}-H)W(\boldsymbol{I}-\widetilde{\lambda}W)^{-1}(\widetilde{\beta_2}\boldsymbol{I}+\widetilde{\beta_1}W_{-1})\boldsymbol{x}_{-1}. \tag{1.26}$$

The second step consists in estimating a 2SLS using as instruments

$$\boldsymbol{Q}=\big[\mathbb{E}\big[(\boldsymbol{I}-H)W\boldsymbol{y}(\widetilde{\boldsymbol{\theta}})|\boldsymbol{X}\big],\,(\boldsymbol{I}-H)\boldsymbol{X}\big].$$

Hence, since the model is just-identified we obtain that:

$$\widehat{\boldsymbol{\theta}}=(\boldsymbol{Q}'\boldsymbol{Z})^{-1}\boldsymbol{Q}'\boldsymbol{y}. \tag{1.27}$$

From Table 1.3 we see that the coefficient (λ) of the contemporaneous peer flows is positive, statistically significant, and robust to the introduction of time and fund fixed effects. Moreover, peer effects are also robust to the introduction of control variables for fund own and peer portfolio characteristics. For instance, column (4) shows the results of the estimation when controlling for fund own portfolio weighted maturity $WAM_{i,t-1}$, peer portfolio weighted maturity $PeerWAM_{i,t-1}$, fund and peer portfolio credit risk

Table 1.3: This table contains the main estimation results for the peer effect model (1.2). Panel A focuses on a universe of funds composed of all fund categories except Treasury Money Market funds. Panel B shows instead the results considering all fund categories and only Prime Money Market funds. Each model is estimated with standard controls for fund characteristics, i.e. natural logarithm of fund and fund family size. However models (4), (5), (6) also control for funds and peers portfolio characteristics. In particular, when *Additional controls* indicates WAM, we control for the fund and peers portfolio weighted average maturity, while CDS represents the inclusion, as controls, of the fund and peers portfolio credit risk exposure, computed as weighted average of the credit default spreads of the issuers belonging to the portfolio. Standard errors are always clustered on time and fund. Significance Levels: 0.01 '***', 0.05 '**', 0.1 '*'.

Panel A:

	All Funds except Treasuries			All Funds except Treasuries		
	(1)	(2)	(3)	(4)	(5)	(6)
$W_t \times Flow_t$	0.7565^{***}	0.5454^{***}	0.7597^{***}	0.7275^{***}	0.7654^{***}	0.6833^{***}
	(0.1473)	(0.1703)	(0.2417)	(0.1668)	(0.1962)	(0.1863)
$Flow_{t-1}$	-0.1228^{***}	-0.1230^{***}	-0.1271^{***}	-0.1261^{***}	-0.1272^{***}	-0.1259^{***}
	(0.0205)	(0.0206)	(0.0185)	(0.0185)	(0.0185)	(0.0185)
$W_{t-1} \times Flow_{t-1}$	0.1268^{***}	0.1255^{***}	0.1679^{***}	0.1705^{***}	0.1707^{***}	0.1684^{***}
	(0.0336)	(0.0422)	(0.0404)	(0.0391)	(0.0396)	(0.0397)
$Yield_{t-1}$	-0.1493	-0.1507	0.3409^{**}	0.2705^{*}	0.3204^{**}	0.2885^{*}
	(0.1031)	(0.1041)	(0.1620)	(0.1616)	(0.1626)	(0.1609)
$W_{t-1} \times Yield_{t-1}$	0.1823	0.1771	0.2681	0.3380	0.2721	0.3291
	(0.1350)	(0.1798)	(0.2352)	(0.2360)	(0.2303)	(0.1609)
Controls:	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls:	No	No	No	WAM, CDS	CDS	WAM
Spatial lag IV:	4	4	4	4	4	4
Fixed Effects:	No	Time	Time, Fund	Time, Fund	Time, Fund	Time, Fund
N. Obs.	17,442	17,442	17,442	17,442	17,442	17,442

Panel B:

	All Funds			Only Prime Funds		
	(7)	(8)	(9)	(10)	(11)	(12)
$W_t \times Flow_t$	0.6421^{***}	0.4740^{***}	0.3929^{***}	0.4907^{*}	0.3613^{*}	0.3611^{**}
	(0.0905)	(0.1075)	(0.1413)	(0.2765)	(0.2182)	(0.1828)
$Flow_{t-1}$	-0.1262^{***}	-0.1267^{***}	-0.1340^{***}	-0.0973^{***}	-0.0984^{***}	-0.1113^{***}
	(0.0188)	(0.0188)	(0.0170)	(0.0363)	(0.0370)	(0.0356)
$W_{t-1} \times Flow_{t-1}$	0.1224^{***}	0.0947^{***}	0.1313^{***}	0.1322^{***}	0.1338^{***}	0.1360^{***}
	(0.0347)	(0.0312)	(0.0366)	(0.0353)	(0.0468)	(0.0510)
$Yield_{t-1}$	-0.1999^{**}	-0.1841^{**}	0.2627	0.0063	0.0036	0.2220
	(0.0937)	(0.0932)	(0.1752)	(0.1360)	(0.1394)	(0.2433)
$W_{t-1} \times Yield_{t-1}$	0.2497^{*}	0.1560	0.4215^{**}	0.0481	-0.1024	-0.1797
	(0.1355)	(0.1354)	(0.2122)	(0.1516)	(0.1982)	(0.2777)
Controls:	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls:	No	No	No	No	No	No
Spatial lag IV:	4	4	4	4	4	4
Fixed Effects:	No	Time	Time, Fund	No	Time	Time, Fund
N. Obs.	20,691	20,691	20,691	8,721	8,721	8,721

exposure $CDS_{i,t-1}$ and $PeerCDS_{i,t-1}$. Peer effects remain positive and statistically significant also when considering all fund categories or only Prime funds.
The presence of positive peer effects causes the, so called, spatial multiplier effect. That is, an initial exogenous flow shock may be endogenously amplified by the system resulting in a total higher impact to the industry and potentially also affecting funds that did not initially received the shock but holds similar portfolios, causing selling pressure on funds that hold similar securities that, in turn, may have limited liquidity. The financial stability implications of peer effects will be analyzed in detail in Section 1.7.
We also find a positive and statistically significant relationship with the lagged peer flows, suggesting the presence of temporal flow shocks spillovers from peers which will be studied in more details in Section 1.6.1. In line with the literature on the Flow-Performance relationship, our empirical results show that this relationship is positive and statistically significant in the Money Market Mutual Fund industry.
As shown in Appendix 1.B and Table 1.5, these results are robust to the inclusion of Treasury MMMFs and to different ways to construct the similarity network.

1.6.1 Impact Measures

Due to the presence of spatial interactions, interpretation of spatial models based on parameters point estimates may lead to misleading conclusions. Indeed, in a classical linear regression model, the parameters estimates represent the *(direct) marginal effect* of the exogenous variables on the dependent one. On the contrary, this is not true for spatial models. As suggested in [LeSage and Pace, 2009], a partial derivatives approach to assess the impact of changes in the value of an independent variable should be preferred for spatial models. The reduced form of model (1.17) is given by:

$$\begin{aligned}\boldsymbol{y} =&(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1}+(\boldsymbol{I}-\lambda W)^{-1}(\beta_2 I+\beta_1 W_{-1})\boldsymbol{x}_{-1}+\\ &+(\boldsymbol{I}-\lambda W)^{-1}(\boldsymbol{\mathcal{C}\delta}+\boldsymbol{\alpha}\otimes\boldsymbol{\iota}+\boldsymbol{\iota_T}\otimes\boldsymbol{\zeta}_n)+(\boldsymbol{I}-\lambda W)^{-1}\boldsymbol{\epsilon}. \qquad (1.28)\end{aligned}$$

Hence the matrices of partial derivatives of $\boldsymbol{y}$ with respect to the error term, past flow and yield are given respectively by:

$$\frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{\epsilon}_t} = (\boldsymbol{I} - \lambda W_t)^{-1}, \tag{1.29}$$

$$\frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{y}_{t-1}} = (\boldsymbol{I} - \lambda W_t)^{-1}(\rho \boldsymbol{I} + \gamma W_{t-1}), \tag{1.30}$$

$$\frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{x}_{t-1}} = (\boldsymbol{I} - \lambda W_t)^{-1}(\beta_2 \boldsymbol{I} + \beta_1 W_{t-1}), \tag{1.31}$$

where, for instance, we can interpret the element $\left(\frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{\epsilon}_t}\right)_{ij}$ as the impact on the flow of fund i of a simultaneous unit shock to the flow of fund j. Each one of the previous partial derivatives is a matrix usually referred to as an impact matrix. Let denote by $\boldsymbol{S}_x(W)$, with $x \in \{\boldsymbol{\epsilon}, \boldsymbol{y}_{-1}, \boldsymbol{x}_{-1}\}$, the generic $TN \times TN$ block diagonal impact matrix[8] for model (1.17). Moreover, we define $\boldsymbol{S}_x(W_t)$ as the generic time t impact matrix obtained considering only the time t similarity network W_t. In order to provide insights on the spatial dependence between the flows of MMMFs, we will, for simplicity, focus only on the generic time t impact matrix of simultaneous flow shocks $\boldsymbol{S}_\epsilon(W_t)$. This is an $N \times N$ matrix given by:

$$\boldsymbol{S}_\epsilon(W_t) = \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{\epsilon}_t} = (\boldsymbol{I} - \lambda W_t)^{-1} = \begin{bmatrix} \boldsymbol{S}_\epsilon(W_t)_{11}, & \boldsymbol{S}_\epsilon(W_t)_{12}, & \dots, & \boldsymbol{S}_\epsilon(W_t)_{1N} \\ \boldsymbol{S}_\epsilon(W_t)_{21}, & \boldsymbol{S}_\epsilon(W_t)_{22}, & \dots, & \boldsymbol{S}_\epsilon(W_t)_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{S}_\epsilon(W_t)_{N1}, & \boldsymbol{S}_\epsilon(W_t)_{N2}, & \dots, & \boldsymbol{S}_\epsilon(W_t)_{NN} \end{bmatrix}. \tag{1.32}$$

The diagonal elements of (1.32) measure the direct effect of a shock ϵ_i on the flow of the same crossectional unit i, while the off-diagonal elements measure the indirect effect of a shock ϵ_i on the flow of funds other than i. For instance, $\boldsymbol{S}_\epsilon(W_t)_{ij}$ measures the indirect effect that a shock to the flow of fund j has on the contemporaneous flow of fund i. Moreover, the sum of the off-diagonal elements of column i measures the total impact

[8]Here $\boldsymbol{S}_x(W)$ is a block diagonal matrix that represent the impact matrix for the overall stacked observations. Since we are considering a temporal network, a similar matrix could be defined for each time period t.

that a shock to fund i has on the flows of all the other funds. Instead, the sum of the off-diagonal elements in row i measures the impact that the shocks to all other funds have on the flow of fund i, or in other words, the vulnerability of fund i to peers shocks. Furthermore, using the series expansion of $(\boldsymbol{I} - \lambda W_t)^{-1}$, the impact matrix $\boldsymbol{S}_\epsilon(W_t)$ can be decomposed as:

$$\boldsymbol{S}_\epsilon(W_t) = (\boldsymbol{I} - \lambda W_t)^{-1} = \boldsymbol{I} + \lambda W_t + \lambda^2 W_t^2 + \lambda^3 W_t^3 \dots . \tag{1.33}$$

Since in the identity matrix the diagonal elements are all one and the off-diagonal elements are all zero, the first term in (1.33) can be interpreted as the first order direct effects of $\boldsymbol{\epsilon}$ on $\boldsymbol{y}$. By construction, the diagonal elements of W_t are all zero, hence, this term can be interpreted as a first order indirect effect of $\boldsymbol{\epsilon}$ on $\boldsymbol{y}$, i.e. the impact on $\boldsymbol{y}$ of shocks to the first-degree neighbors. The diagonal elements of W_t^2 are not necessarily zero, and represent the second order direct effect of $\boldsymbol{\epsilon}$ on $\boldsymbol{y}$, i.e. the impact from a length 2 feedback loop. In this case a length 2 feedback loop is a situation in which a shock ϵ_i has an indirect impact on the flow y_j of a neighbor which, in turn, has an indirect impact back on y_i (see Figure 1.4). It is important to note that, due to the presence of these feedback loops, a one unit change in ϵ_i may have a direct impact on y_i greater than 1. The off-diagonal elements of W_t represent the first order indirect effects of neighbors, i.e. the impact on y_i of a flow shock ϵ_j to a first degree neighbor j. Similarly, the off-diagonal elements of the second term W_t^2 in (1.33) can be interpreted as a second order indirect effect, i.e. the impact from the second degree neighbors. An example of second order indirect impact of j on i is shown in Figure 1.5, where j has an indirect impact on a neighbor k that has, in turn, and indirect impact on i.

Since the impacts change for each observation, a series of summary scalar impact measures for $\boldsymbol{S}_x(W)$ have been proposed in the literature [Elhorst, 2014], [Arbia et al., 2019]. The first one is called *Average Direct Impact*, and measures the average total impact on y_i of a unit change of the variable x_i for $i = 1, \dots, TN$, and is defined as the average of

Figure 1.4: Simultaneous flow shock second order direct impact.

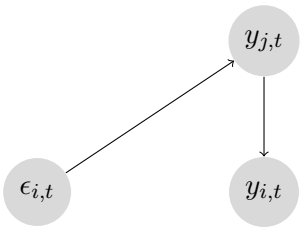
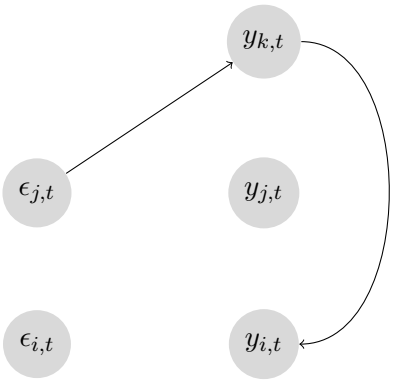


Figure 1.5: Simultaneous flow shock second order indirect impact.



all the diagonal elements in $\boldsymbol{S}_x(W)$:

$$\textit{Average Direct Impact} = \bar{M}(x)_{direct} = \frac{1}{TN}\sum_{i=1}^{TN} S_x(W)_{ii}. \tag{1.34}$$

The second one is a global measure called *Average Total Impact* and is defined as the average of all the row or column sums in $\boldsymbol{S}_x(W)$:

$$\textit{Total Direct Impact} = \bar{M}(x)_{total} = \frac{1}{TN}\sum_{i=1}^{TN}\sum_{j=1}^{TN} S_x(W)_{ij}. \tag{1.35}$$

The difference between the *Average Total Impact* and the *Average Direct Impact* is another measure called *Average Indirect Impact*. It can be computed as the average of the row or column sum of the off-diagonal elements in $\boldsymbol{S}_x(W)$, and can be interpreted as a measure of Spatial Spillovers:

$$\textit{Average Indirect Impact} = \bar{M}(x)_{indirect} = \bar{M}(x)_{total} - \bar{M}(x)_{direct}. \tag{1.36}$$

Panel A, B and C of Table 1.4 contain the estimated impact measures corresponding to columns (3), (9) and (12) of Table 1.3 respectively. The significance levels are computed by simulating each impact measure using the estimated parameters and their clustered covariance matrix as suggested in [LeSage and Pace, 2009].

It is worth noticing that the average direct impact of a flow shock ϵ is greater than 1 because the *Average Direct Impact* takes also into account any feedback loop where the shock to i has a contemporaneous indirect impact on fund j that has an indirect impact on .... that has an indirect impact on i. Using (1.33) we can partition the impact measures into the contribution of first, second and higher degree neighbors, which are shown, for the case of all fund categories except Treasuries, in Panel A of Table 1.4 and Figure 1.6.

Peer Flow Effects

We have defined Peer Flow Effects as the contemporaneous dependence between the flows of similar funds. From Panel A of Table 1.4 we see that the average indirect impact of a flow shock is positive but not statistically significant. However, if we decompose the measure as a function of the peers order we have that the average indirect impacts of a flow shock from first and second order peers are positive and statistically significant, supporting the evidence of a contemporaneous dependence between the flows of similar funds. This dependence may allow the transmission of outflow shocks between peers that, in turn, may result in widespread run like events. If wee look at Figure 1.6, we see that the magnitude of peer flow effects is decreasing with the order of the neighbor degree. Thus, the impact from first and second order neighbors, which have a higher degree of similarity, is higher than the impact from higher order neighbors.

The estimated indirect impact measure of first degree neighbors is 0.7 which means that for a fund i, on average, a one percent flow shock to the top 10% most similar funds may result in 0.7% indirect flow shock to i, even though i did not originally received a direct shock.

Another way to see the impact of peer flow effects is considering the case in which every fund in the network receives a one percent flow shock. Without spatial interaction, the average total impact on each fund would be equal to a one percent shock. However when peer effects are present, if we only consider the contributions from the statistically significant lags, the average total impact on each fund is equal to about a 2.3% percent shock. This increase is due to the spatial multiplier caused by the spatial interactions.

Flow Spillovers

We have defined flow spillovers as the dependence between current flows and lagged peer flows. From Panel A of Table 1.4 we see a negative and statistically significant direct impact of lagged flows. From (1.33) we can rewrite the lagged flow impact matrix as

$$\frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{y}_{t-1}} = (\boldsymbol{I} - \lambda W_t)^{-1}(\rho \boldsymbol{I} + \gamma W_{t-1}), \tag{1.37}$$

$$= (\boldsymbol{I} + \lambda W_t + \lambda^2 W_t^2 + \lambda^3 W_t^3 \dots)(\rho \boldsymbol{I} + \gamma W_{t-1}). \tag{1.38}$$

Hence, it is easy to see that, for instance, the Lag 0 impact matrix is given by:

$$\rho \boldsymbol{I}, \tag{1.39}$$

while the Lag 1 impact matrix is:

$$\rho\lambda W_t + \gamma W_{t-1}, \tag{1.40}$$

and, similarly, the Lag 2 impact matrix is:

$$\rho\lambda^2 W_t^2 + \lambda\gamma W_t W_{t-1}. \tag{1.41}$$

The off-diagonal elements of the first term in (1.40) represent the indirect impact of the time t first degree neighbors, while the off-diagonal elements of the second term represent

the indirect impact of the time $t-1$ first degree neighbors. What (1.40) implies is that the flow at time $t-1$ of a fund j, may have a first order indirect impact on the flow at time t of a fund i both if j was a first degree neighbor of i at time $t-1$, but also if j is a first degree neighbor of i at time t.

It is important to highlight that the Lag 2 impact matrix (1.41) captures the impact of both, spatial and temporal, second order degree neighbors. On the one hand, a fund j is a second order spatial neighbor of i if, at time t, j is a neighbor of a fund k, which is also a neighbor, at time t, of fund i. On the other hand, a fund j is a second order temporal neighbor of i if, at time $t-1$, j is a neighbor of a fund k which, at time t, is a neighbor of i.

We find a positive and weakly significant indirect impact of lagged peer flows on current flows from first and second degree neighbors, supporting the evidence of temporal spillovers between the flows of similar funds. Hence, a 1% change in peer flows at time $t-1$ of a fund i, may result on average in a 0.12% change in the time t flows of fund i. Moreover, as shown in Figure 1.6, the magnitude of indirect impact for lagged flows is rapidly decreasing with the order of the neighbor degree.

Flow Performance Relationship

The yield impact matrix (1.31) can be interpreted as the flow-performance relationship implied by our spatial model. The literature on the flow performance relationship studies the dependence between the flows of a fund and its own past performance. In our model, this dependence is summarized by average direct impact measure for the past yield in Tables 1.4. Consistently with the empirical evidence, this impact measure is positive and statistically significant. From (1.33) we have that:

$$\frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{x}_{t-1}} = (\boldsymbol{I} - \lambda W_t)^{-1}(\beta_2 \boldsymbol{I} + \beta_1 W_{t-1}), \tag{1.42}$$

$$= (\boldsymbol{I} + \lambda W_t + \lambda^2 W_t^2 + \lambda^3 W_t^3 \dots)(\beta_2 \boldsymbol{I} + \beta_1 W_{t-1}). \tag{1.43}$$

Hence the Lag 0 impact matrix is given by:

$$\beta_2 \boldsymbol{I}, \tag{1.44}$$

while the Lag 1 and 2 impact matrices are, respectively:

$$\beta_2 \lambda W_t + \beta_1 W_{t-1}, \tag{1.45}$$

and

$$\beta_2 \lambda^2 W_t^2 + \beta_1 \lambda W_t W_{t-1}. \tag{1.46}$$

It is important to highlight that the Lag 1 impact matrix (1.45), accounts for the past yield impact of both current and $t-1$ first degree neighbors. Table 1.4 suggest the existence of a positive and significant dependence between the flow a fund and the past performance of its peers. In particular, we find this relationship to be statistically significant for first and second degree neighbors and decreasing with the peer spatial lag.

1.7 Resilience to widespread flow shocks

When a fund receives an extreme outflow, in the absence of redemption gates, the fund is forced to sell an equivalent amount of assets to generate liquidity and meet the redemption demand. If the extreme outflow is greater than its percentage holding in daily liquid assets, the fund manager may also need to sell part of its holdings in less liquid assets that may also be present in the portfolio of other Money Market Funds. If a large proportion of funds is forced to - simultaneously - sell holdings in less liquid assets, then fund managers may incur in losses due to the fire-sale mechanism and the assets finite liquidity.

It is therefore important to asses the magnitude of the shock that the MMMFs industry is able to withstand without triggering potential fire-sales spillovers. The answer is trivial if there exist no spatial dependence between fund flows. If each fund is required to maintain a 10% holding in daily liquid assets, then the MMMFs industry should be

able to withstand a system wide 10% flow shock.
However, this is not the case when Peer Flow Effects are presents. For instance, from Table 1.4 we know that a flow shock to a given fund may have indirect impacts on other funds resulting in a total impact greater than the initial shock. Moreover, the total impact of a flow shock is not homogeneous among funds and depends on the fund's position within the network and its connections.

As described in Section 1.5, the in-degree k_i^{in} of a fund i can be interpreted as number of funds that could, potentially, be indirectly affected by a flow shock to i. If N is the total number of funds in the network, the *in-degree centrality* of fund i is defined as

$$C_i^n = \frac{k_i^{in}}{N}. \tag{1.47}$$

Fixing a time instant t and using the estimate of the spatial dependence parameter $\hat{\lambda} = 0.7597$, we can measure the total impact of flow shock s to a fund i as

$$\sum_{i,j} \left[(\boldsymbol{I} - \hat{\lambda} W_t)^{-1} \boldsymbol{\epsilon}_s^i \right]_{ij}, \tag{1.48}$$

where $\boldsymbol{\epsilon}_s^i$ is an $N \times 1$ flow shock vector such that $(\boldsymbol{\epsilon}_s^i)_j = 0$ for $j \neq i$ and $(\boldsymbol{\epsilon}_s^i)_i = s$. Equation (1.48) considers the impact of neighbors for every possible lag, however, as we have seen in Table 1.4, the impact of peers flow shock is statistically significant only up to the second order lag. Thus, in order to study the resiliency of the money market industry and focus only on the statistically significant relationships, we consider in this Section the decomposition of the impact matrix up to the first two lags:

$$\sum_{i,j} \left[(\boldsymbol{I} + \hat{\lambda} W_t + \hat{\lambda}^2 W_t^2) \boldsymbol{\epsilon}_s^i \right]_{ij}. \tag{1.49}$$

Figure 1.7 reports, for a sample of months, the relationship between the total impact of a 10% flow shock to a fund (y axis) and its in-degree centrality (x axis). We see that shocks to funds with higher in-degree centrality tend to have a higher total impact,

consistently across the different dates.

In order to study the resilience of the MMMFs industry to widespread flow shocks we analyze how the percentage of funds that receive a total outflow greater than the percentage of daily liquid assets (assumed to be 10%), i.e. the percentage of funds forced to sell less liquid assets, changes as function of the percentage of funds shocked and the magnitude of the flow shock. Without spatial interactions, if the magnitude of the outflow shock is less than 10% then we expect the percentage of funds forced to sell less liquid assets to be equal to 0% independently of the percentage of funds initially shocked, while, if the magnitude of the outflow shock is greater or equal than 10%, we expect the percentage of funds forced to sell less liquid assets to increase linearly with the percentage of funds initially shocked.

Let us define f_{in} as the percentage of funds initially receiving a flow shock and f_{out} as the percentage of funds forced to sell less liquid assets. Using the September 2015 snapshot of the similarity temporal network defined in Section 1.5, for each $f_{in} \in (0\%, 100\%)$ and for each outflow shock s in $[5\%, 7\%, 10\%, 20\%, 30\%, 40\%, 50\%]$, we randomly select 1000 times $N \cdot f_{in}$ unique funds and we simultaneously shock them while leaving all the other funds shock equal to zero. After the shock, for each pair (f_{in}, s), we compute f_{out} as the average percentage of funds resulting in a total outflow that is grater than 10%. The results are shown in the left panel of Figure 1.8. If the flow shock magnitude is $s = 5\%$, no fund is forced to sell less liquid assets. However, focusing on the case $s = 7\%$ we see that, due to the spatial interaction, if the percentage of funds contemporaneously receiving the shock is high enough, then f_{out} starts becoming greater than zero and grows rapidly even if the initial flow shock is below the level of daily liquid assets. Fixing the percentage of funds shocked f_{in}, we also see that f_{out} is increasing in the shock magnitude s meaning that, if the flow shock is strong enough, then fewer funds need to be shocked in order for the industry to observe a high percentage of funds forced to sell less liquid assets. For instance, if 50% of the funds receive an outflow shock of 40%, more than 90% of the funds may suffer a total outflow of 10% or more. The black dashed line represents,

for comparison, the relationship between f_{in} and f_{out} for any flow shock greater than 10% assuming instead no spatial interactions.
Furthermore, the right panel of Figure 1.8 shows how f_{out} changes as function of s when the 20% of funds shocked have a high in-degree centrality (blue line) or low in-degree centrality (red line). If the funds shocked can have an impact only on few funds, then f_{out} stays almost constant with respect to s. On the other hand, if the funds receiving the initial shock have the potential to influence a high number of funds, then f_{out} rapidly grows as the shock magnitude s increases.

It is not possible to identify, a priori, the number of funds simultaneously selling less liquid assets that will likely trigger fire sales spillovers, because it may depend on the market conditions and the asset liquidity. However, for our analysis we can define a threshold τ, corresponding to the number of funds simultaneously selling less liquid assets, i.e. simultaneously receiving a flow shock greater than the percentage of daily liquid assets, to characterize the resilience the MMF industry. The two dimensions of a run-like event are the magnitude of the flow shock and the percentage of funds shocked. For each combinations of the two, if the resulting percentage of funds selling less liquid assets is greater than τ, then we say that we are in the fire-sales spillover or vulnerability region. Figure 1.9 shows for various values of τ the different fire-sales regions in grey, and its lower boundary as a solid black line. The red dashed line shows the lower bound for the fire-sales region assuming no Peer Effects. The first observation stemming from Figure 1.9 is that, independently of the threshold τ, accounting for the presence of Peer Effects shows a larger vulnerability region than predicted by the case without spatial interaction. If $\tau = 50\%$, not accounting for Peer Effects shows that the risk of fire sales spillovers exists only if 50% of more of the industry receives a shock greater than 10%. However, the presence of spatial interactions highlights that the fire sales risk may be present also if a smaller percentage of funds receives a strong enough shock. For instance, if 20% of the funds receive a flow shock greater than 80%. Obviously, the greater τ the smaller is the fire sales spillover region.

1.8 Conclusion

Using a novel data-set of portfolio-level holdings of US Money Market Mutual Funds we find evidence of positive and statistically significant endogenous peer flow effects. Their economic implication is also significant. A fund may indirectly suffer a 0.75% flow shock as a result of a 1% shocks to its most similar funds. These indirect impacts are decreasing with the network distance between two funds and statistically significant up to second order neighbors. Moreover, we show that the endogenous amplification of a flow shock caused by peer effects has a substantial implications for the vulnerability of the MMMFs industry to run-like behaviours. The total impact of a flow shock is heterogeneous and depends on the centrality of the fund shocked. More importantly, peer effects create a non linear vulnerability region suggesting that flow shocks below the percentage of daily liquid assets can still generate fire-sale spillovers if the percentage of funds shocks is high enough.

Appendix 1.A Proofs: Identification

1.A.1 No Fixed Effects

Without loss of generality, and for ease of explanation, we will not consider the control variables $\boldsymbol{\mathcal{C}}_t\boldsymbol{\delta}'$ in this section. Let first start by rewriting the structural model without any fixed effect as:

$$\boldsymbol{y} = \alpha\boldsymbol{\iota} + \lambda W\boldsymbol{y} + \rho\boldsymbol{y}_{-1} + \gamma W_{-1}\boldsymbol{y}_{-1} + \beta_1 W_{-1}\boldsymbol{x}_{-1} + \beta_2\boldsymbol{x}_{-1} + \boldsymbol{\epsilon}, \tag{1.50}$$

where $\boldsymbol{y}$, $\boldsymbol{y}_{-1}$, $\boldsymbol{x}_{-1}$ and $\boldsymbol{\epsilon}$ are TN column vectors given by:

$$\begin{aligned}
\boldsymbol{y} &= (\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_T)', \\
\boldsymbol{y}_{-1} &= (\boldsymbol{y}_0, \boldsymbol{y}_1, \dots, \boldsymbol{y}_{T-1})', \\
\boldsymbol{x}_{-1} &= (\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{T-1})', \\
\boldsymbol{\epsilon} &= (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_T)'.
\end{aligned}$$

Moreover, W and W_{-1} are $TN \times TN$ block diagonal matrices given by $diag(W_1, \dots, W_T)$ and $diag(W_0, \dots, W_{T-1})$, respectively. The reduced form of (1.50) can then be written as:

$$\begin{aligned}
\boldsymbol{y} = \alpha(\boldsymbol{I} - \lambda W)^{-1}\boldsymbol{\iota} &+ (\boldsymbol{I} - \lambda W)^{-1}(\rho\boldsymbol{I} + \gamma W_{-1})\boldsymbol{y}_{-1} + \\
&+ (\boldsymbol{I} - \lambda W)^{-1}(\beta_2\boldsymbol{I} + \beta_1 W_{-1})\boldsymbol{x}_{-1} + \\
&+ (\boldsymbol{I} - \lambda W)^{-1}\boldsymbol{\epsilon}.
\end{aligned} \tag{1.51}$$

Let assume that W and W_{-1} are distinct, row normalized and do not contain any zero row. In order to prove Proposition 1.4.2, we first prove the following Lemma:

Lemma 1.A.1. *Suppose that the previous mild conditions are satisfied and that $\rho\gamma \neq 0$ or $\beta_2\beta_2 \neq 0$. Model* (1.50) *is identified if and only if $\exists\, t$ such that the matrices I, W_t, W_{t-1} and W_tW_{t-1} are linearly independent.*

Proof. Let assume that $(\alpha, \lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\alpha', \lambda', \rho', \gamma', \beta_1', \beta_2')$ are two set of structural parameters that lead to the same reduced form model. Then we have

$$\alpha(\boldsymbol{I} - \lambda W)^{-1}\boldsymbol{\iota} = \alpha'(\boldsymbol{I} - \lambda' W)^{-1}\boldsymbol{\iota} \tag{1.52}$$

$$(\boldsymbol{I} - \lambda W)^{-1}(\rho \boldsymbol{I} + \gamma W_{-1}) = (\boldsymbol{I} - \lambda' W)^{-1}(\rho' \boldsymbol{I} + \gamma' W_{-1}) \tag{1.53}$$

$$(\boldsymbol{I} - \lambda W)^{-1}(\beta_2 \boldsymbol{I} + \beta_1 W_{-1}) = (\boldsymbol{I} - \lambda' W)^{-1}(\beta_2' \boldsymbol{I} + \beta_1' W_{-1}). \tag{1.54}$$

Since $\forall\, \lambda$ we have that[9] $W(I - \lambda W)^{-1} = (I - \lambda W)^{-1} W$, if we multiply by $(I - \lambda' W)(I - \lambda W)$ equalities (1.53) and (1.54) , we get

$$(\rho - \rho')I + (\lambda\rho' - \lambda'\rho)W + (\gamma - \gamma')W_{-1} + (\lambda\gamma' - \lambda'\gamma)WW_{-1} = 0, \tag{1.55}$$

and

$$(\beta_2 - \beta_2')I + (\lambda\beta_2' - \lambda'\beta_2)W + (\beta_1 - \beta_1')W_{-1} + (\lambda\beta_1' - \lambda'\beta_1)WW_{-1} = 0. \tag{1.56}$$

If I, W, W_{-1} and WW_{-1} are linearly independent, then from (1.55) and (1.56) the following two systems of equations must be satisfied:

$$\begin{cases} \rho - \rho' = 0 \\ \lambda\rho' - \lambda'\rho = 0 \\ \gamma - \gamma' = 0 \\ \lambda\gamma' - \lambda'\gamma = 0 \end{cases} \tag{1.57}$$

$$\begin{cases} \beta_2 - \beta_2' = 0 \\ \lambda\beta_2' - \lambda'\beta_2 = 0 \\ \beta_1 - \beta_1' = 0 \\ \lambda\beta_1' - \lambda'\beta_1 = 0 \end{cases} \tag{1.58}$$

[9] This can be easily seen if we consider the series expansion $(I - \lambda W)^{-1} = \sum_{k=0}^{\infty} \lambda^k W^k$.

Solving (1.57) and (1.58) yields then $\alpha = \alpha'$, $\lambda = \lambda'$, $\rho = \rho'$, $\gamma = \gamma'$, $\beta_1 = \beta_1'$ and $\beta_2 = \beta_2'$, which means that Model (1.50) is identified.

Let now suppose that I, W, W_{-1} and WW_{-1} are linearly dependent, i.e. $\exists\, \theta_1, \theta_2, \theta_3$, with at least one different from zero, such that:

$$WW_{-1} = \theta_1 I + \theta_2 W + \theta_3 W_{-1}. \tag{1.59}$$

Since, by assumption, W and W_{-1} are row-normalized, we have that $W\boldsymbol{\iota} = W_{-1}\boldsymbol{\iota} = \boldsymbol{\iota}$, which implies $\theta_1 + \theta_2 + \theta_3 = 1$. By substituting (1.59) into (1.53), we obtain:

$$(\rho - \rho' + \theta_1(\lambda\gamma' - \lambda'\gamma))I + (\lambda\rho' - \lambda'\rho + \theta_2(\lambda\gamma' - \lambda'\gamma))W + (\gamma - \gamma' + \theta_3(\lambda\gamma' - \lambda'\gamma))W_{-1} = 0. \tag{1.60}$$

Since, by assumption, W and W_{-1} are different and both have zeros on the main diagonal (because no fund is connected to itself), I, W and W_{-1} are linearly independent. Therefore the coefficients of (1.60) satisfy:

$$\begin{cases} \rho - \rho' + \theta_1(\lambda\gamma' - \lambda'\gamma) = 0 \\ \lambda\rho' - \lambda'\rho + \theta_2(\lambda\gamma' - \lambda'\gamma) = 0 \\ \gamma - \gamma' + \theta_3(\lambda\gamma' - \lambda'\gamma) = 0 \end{cases} \tag{1.61}$$

If we sum the first two equations of (1.61) we obtain:

$$\begin{aligned} & \rho - \rho' + \theta_1(\lambda\gamma' - \lambda'\gamma) + \lambda\rho' - \lambda'\rho + \theta_2(\lambda\gamma' - \lambda'\gamma) = 0, \\ \Longrightarrow\ & \rho - \rho' + \lambda\rho' - \lambda'\rho + (1 - \theta_3)(\lambda\gamma' - \lambda'\gamma) = 0, \\ \Longrightarrow\ & \rho - \rho' + \lambda\rho' - \lambda'\rho + \lambda\gamma' - \lambda'\gamma - \theta_3(\lambda\gamma' - \lambda'\gamma) = 0, \\ \Longrightarrow\ & \rho(1 - \lambda') - \rho'(1 - \lambda) + \lambda\gamma' - \lambda'\gamma - \theta_3(\lambda\gamma' - \lambda'\gamma) = 0, \\ \Longrightarrow\ & \rho(1 - \lambda') - \rho'(1 - \lambda) + \lambda\gamma' - \lambda'\gamma + \gamma - \gamma' - \gamma + \gamma' - \theta_3(\lambda\gamma' - \lambda'\gamma) = 0, \\ \Longrightarrow\ & \rho(1 - \lambda') - \rho'(1 - \lambda) + \gamma(1 - \lambda') - \gamma'(1 - \lambda) - \gamma + \gamma' - \theta_3(\lambda\gamma' - \lambda'\gamma) = 0, \\ \Longrightarrow\ & (\rho + \gamma)(1 - \lambda') - (\rho' + \gamma')(1 - \lambda) - \gamma + \gamma' - \theta_3(\lambda\gamma' - \lambda'\gamma) = 0. \end{aligned}$$

Since $\gamma - \gamma' + \theta_3(\lambda\gamma' - \lambda'\gamma) = 0$, from the last equation we also have that

$$(\rho + \gamma)(1 - \lambda') - (\rho' + \gamma')(1 - \lambda) = 0, \tag{1.62}$$

which implies that the third equation of (1.61) can be obtained as the sum of the first two. Hence, in order for (ρ, λ, γ) and $(\rho', \lambda', \gamma')$ to lead to the same reduced form, only two equations need to be satisfied. Therefore, Model (1.50) is not identified.
We have proven, so far, that Model (1.50) is identified if and only if I, W, W_{-1} and WW_{-1} are linearly independent. The lemma is then obtained by observing that W, W_{-1} and WW_{-1} are block diagonal. Thus I, W, W_{-1} and WW_{-1} are linearly independent if and only if $\exists\, t$ such that the matrices I, W_t, W_{t-1} and W_tW_{t-1} are linearly independent. □

The result and proof of Lemma 1.A.1 is similar to Proposition 1 in [Bramoullé et al., 2009], with the main difference being that in our case endogenous and exogenous peer effects are not transmitted through the same network but rather, through two consecutive time instant of a more general temporal network.

Proposition 1.A.1. *Suppose that the previous mild conditions are satisfied and that* $\rho\gamma \neq 0$ *or* $\beta_2\beta_2 \neq 0$. *If there exists at least one Intransitive Temporal Triad, then model* (1.50) *is identified.*

Proof. Let now assume that there exist at least one Intransitive Temporal Tried. This means that $\exists\, t$ and i, j, k such that i is influenced by j at time t, j is influenced by k at time $t-1$, but i is not influenced by k at time t of $t-1$. In other words we have that $(W_t)_{ik} = (W_{t-1})_{ik} = 0$, $(W_t)_{ij} \neq 0$ and $(W_{t-1})_{jk} \neq 0$. This implies that $(W_tW_{t-1})_{ik} = (W_t)_{i,} \cdot (W_{t-1})_{,k} \geq (W_t)_{i,j}(W_{t-1})_{j,k} > 0$. Hence, I, W_t, W_{t-1} and W_tW_{t-1} are linearly independent, which implies by Lemma 1.A.1, that Model (1.50) is identified.

□

1.A.2 With time fixed effects

Without loss of generality, and for ease of explanation, we will not consider the control variables $\boldsymbol{\mathcal{C}}_t\boldsymbol{\delta}'$ in this section. Let first start by rewriting the structural model with time fixed effects as:

$$\boldsymbol{y} = \boldsymbol{\alpha} \otimes \boldsymbol{\iota} + \lambda W\boldsymbol{y} + \rho\boldsymbol{y}_{-1} + \gamma W_{-1}\boldsymbol{y}_{-1} + \beta_1 W_{-1}\boldsymbol{x}_{-1} + \beta_2\boldsymbol{x}_{-1} + \boldsymbol{\epsilon}, \tag{1.63}$$

where $\boldsymbol{\alpha}$ is the $T \times 1$ vector of time fixed effects, $\boldsymbol{y}$, $\boldsymbol{y}_{-1}$, $\boldsymbol{x}_{-1}$ and $\boldsymbol{\epsilon}$ are TN column vectors given by:

$$\begin{aligned}
\boldsymbol{y} &= (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_T)', \\
\boldsymbol{y}_{-1} &= (\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{T-1})', \\
\boldsymbol{x}_{-1} &= (\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{T-1})', \\
\boldsymbol{\epsilon} &= (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \ldots, \boldsymbol{\epsilon}_T)'.
\end{aligned}$$

Local Demeaning

Similarly to [Bramoullé et al., 2009], we first eliminate the Time Fixed Effects using a demeaning operator. In this section we first focus on the local demeaning operator, i.e. for each variable we subtract from each fund the average taken over its neighbors in the network:

$$\begin{aligned}
(\boldsymbol{I} - W)\boldsymbol{y} = (\boldsymbol{I} - W)\lambda W\boldsymbol{y} + \rho(\boldsymbol{I} - W)\boldsymbol{y}_{-1} + \gamma(\boldsymbol{I} - W)W_{-1}\boldsymbol{y}_{-1} + \beta_1(\boldsymbol{I} - W)W_{-1}\boldsymbol{x}_{-1} \\
+ \beta_2(\boldsymbol{I} - W)\boldsymbol{x}_{-1} + (\boldsymbol{I} - W)\boldsymbol{\epsilon}.
\end{aligned} \tag{1.64}$$

The demeaned structural model (1.64) has now no time fixed effect, and in this context the demeaned reduced form becomes:

$$\begin{aligned}(\boldsymbol{I}-W)\boldsymbol{y} =&(\boldsymbol{I}-W)(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1}+\\ &+(\boldsymbol{I}-W)(\boldsymbol{I}-\lambda W)^{-1}(\beta_2\boldsymbol{I}+\beta_1 W_{-1})\boldsymbol{x}_{-1}+\\ &+(\boldsymbol{I}-W)(\boldsymbol{I}-\lambda W)^{-1}\boldsymbol{\epsilon}. \end{aligned} \tag{1.65}$$

As in the previous Section, we assume that W and W_{-1} are distinct, row normalized and do not contain any zero row. Hence, in order to prove Proposition 1.4.1, we first prove the following Lemma:

Lemma 1.A.2. *Suppose that the previous mild conditions are satisfied and that* $\rho\gamma \neq 0$ *or* $\beta_2\beta_2 \neq 0$. *Model* (1.64) *is identified if and only if* $\exists t$ *such that the matrices* I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 *and* $W_t^2W_{t-1}$ *are linearly independent.*

Proof. Let assume that $(\alpha, \lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\alpha', \lambda', \rho', \gamma', \beta_1', \beta_2')$ are two set of structural parameters that lead to the same reduced form model (1.65). Then we have

$$(\boldsymbol{I}-W)(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1}) = (\boldsymbol{I}-W)(\boldsymbol{I}-\lambda' W)^{-1}(\rho'\boldsymbol{I}+\gamma' W_{-1}), \tag{1.66}$$

$$(\boldsymbol{I}-W)(\boldsymbol{I}-\lambda W)^{-1}(\beta_2\boldsymbol{I}+\beta_1 W_{-1}) = (\boldsymbol{I}-W)(\boldsymbol{I}-\lambda' W)^{-1}(\beta_2'\boldsymbol{I}+\beta_1' W_{-1}), \tag{1.67}$$

or equivalently, since $W(I-\lambda W)^{-1} = (I-\lambda W)^{-1}W$ for every λ,

$$(\boldsymbol{I}-\lambda W)^{-1}(\boldsymbol{I}-W)(\rho\boldsymbol{I}+\gamma W_{-1}) = (\boldsymbol{I}-\lambda' W)^{-1}(\boldsymbol{I}-W)(\rho'\boldsymbol{I}+\gamma' W_{-1}), \tag{1.68}$$

$$(\boldsymbol{I}-\lambda W)^{-1}(\boldsymbol{I}-W)(\beta_2\boldsymbol{I}+\beta_1 W_{-1}) = (\boldsymbol{I}-\lambda' W)^{-1}(\boldsymbol{I}-W)(\beta_2'\boldsymbol{I}+\beta_1' W_{-1}). \tag{1.69}$$

If we multiply both members of the previous two equalities by $(I-\lambda' W)(I-\lambda W)$, we then get

$$\begin{aligned}(\rho-\rho')\boldsymbol{I}+(\gamma-\gamma')W_{-1}+(\rho'-\rho+\lambda\rho'-\lambda'\rho)W+(\gamma'-\gamma+\lambda\gamma'-\lambda'\gamma)WW_{-1}+\\ +(\lambda'\rho-\lambda\rho')W^2+(\lambda'\gamma-\lambda\gamma')W^2W_{-1}=0,\end{aligned} \tag{1.70}$$

and

$$(\beta_2-\beta_2')\boldsymbol{I}+(\beta_1'-\beta_1)W_{-1}+(\beta_2-\beta_2'+\lambda\beta_2'-\lambda'\beta_2)W+(\beta_1-\beta_1'+\lambda'\beta_1-\lambda\beta_1')WW_{-1}+$$
$$+(\lambda'\beta_2-\lambda\beta_2')W^2+(\lambda\beta_1'-\lambda'\beta_1)W^2W_{-1}=0. \tag{1.71}$$

If I, W, W_{-1}, WW_{-1}, W^2 and W^2W_{-1} are linearly independent, then from (1.70) and (1.71) the following two systems of equations must be satisfied:

$$\begin{cases} \rho-\rho'=0 \\ \gamma-\gamma'=0 \\ \rho-\rho'+\lambda\rho'-\lambda'\rho=0 \\ \gamma-\gamma'+\lambda'\gamma-\lambda\gamma'=0 \\ \lambda'\rho-\lambda\rho'=0 \\ \lambda\gamma'-\lambda'\gamma=0 \end{cases} \tag{1.72}$$

$$\begin{cases} \beta_2-\beta_2'=0 \\ \beta_1-\beta_1'=0 \\ \beta_2-\beta_2'+\lambda\beta_2'-\lambda'\beta_2=0 \\ \beta_1-\beta_1'+\lambda'\beta_1-\lambda\beta_1'=0 \\ \lambda'\beta_2-\lambda\beta_2'=0 \\ \lambda\beta_1'-\lambda'\beta_1=0 \end{cases} \tag{1.73}$$

Since by assumption $\rho\gamma\neq 0$ or $\beta_2\beta_2\neq 0$, from (1.72) and (1.73) we have that the model is identified.

Suppose now that I, W, W_{-1}, WW_{-1}, W^2 and W^2W_{-1} are linearly dependent. If

$$WW_{-1}=\theta_1 I+\theta_2 W_{-1}+\theta_3 W, \tag{1.74}$$

then $\theta_1+\theta_2+\theta_3=1$, and by substitution in (1.70), $(\alpha,\lambda,\rho,\gamma,\beta_1,\beta_2)$ and $(\alpha',\lambda',\rho',\gamma',\beta_1',\beta_2')$

lead to the same reduced form model if and only if

$$\begin{cases} \rho - \rho' + \theta_1(\gamma' - \gamma + \lambda\gamma' - \lambda'\gamma) = 0 \\ \gamma - \gamma' + \theta_2(\gamma' - \gamma + \lambda\gamma' - \lambda'\gamma) = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\gamma' - \gamma + \lambda\gamma' - \lambda'\gamma) = 0 \\ \lambda'\rho - \lambda\rho' = 0 \\ \lambda'\gamma - \lambda\gamma' = 0 \end{cases} \tag{1.75}$$

which is implies

$$\begin{cases} \rho - \rho' + \lambda'\rho - \lambda\rho' + \theta_1(\gamma - \gamma' + \lambda'\gamma - \lambda\gamma') = 0 \\ \gamma' - \gamma + \lambda\gamma' - \lambda'\gamma + \theta_2(\gamma - \gamma' + \lambda'\gamma - \lambda\gamma') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\gamma - \gamma' + \lambda'\gamma - \lambda\gamma') = 0 \end{cases} \tag{1.76}$$

From (1.76) it is easy to see that the third equation can be obtained as the sum of the first two. Hence, only two equations need to be satisfied for $(\alpha, \lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\alpha', \lambda', \rho', \gamma', \beta_1', \beta_2')$ to lead to the same reduced form model. Therefore the model is not identified. If instead

$$W^2 = \theta_1 I + \theta_2 W_{-1} + \theta_3 W + \theta_4 W W_{-1}, \tag{1.77}$$

then $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 1$. The two set of parameters lead to the same reduced form model if and only if

$$\begin{cases} \rho - \rho' + \theta_1(\lambda'\rho - \lambda\rho') = 0 \\ \gamma - \gamma' + \theta_2(\lambda'\rho - \lambda\rho') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\lambda'\rho - \lambda\rho') = 0 \\ \gamma' - \gamma + \lambda\gamma' - \lambda'\gamma + \theta_4(\lambda'\rho - \lambda\rho') = 0 \\ \lambda'\gamma - \lambda\gamma' = 0 \end{cases} \tag{1.78}$$

Since the last equation of (1.78) can be obtained as the sum of the first four, the following system has the same solution set:

$$\begin{cases} \rho - \rho' + \theta_1(\lambda'\rho - \lambda\rho') = 0 \\ \gamma - \gamma' + \lambda'\gamma - \lambda\gamma' + \theta_2(\lambda'\rho - \lambda\rho') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\lambda'\rho - \lambda\rho') = 0 \\ \gamma' - \gamma + \lambda\gamma' - \lambda'\gamma + \theta_4(\lambda'\rho - \lambda\rho') = 0 \end{cases} \tag{1.79}$$

It is easy to see that in (1.79) the last equation can be obtained as the sum of the first three. Hence, without altering the solution set, we can consider the following system of equations:

$$\begin{cases} \rho - \rho' + \theta_1(\lambda'\rho - \lambda\rho') = 0 \\ (\theta_2 + \theta_3)(\lambda'\rho - \lambda\rho') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\lambda'\rho - \lambda\rho') = 0 \end{cases} \tag{1.80}$$

Again, since the third equation can be obtained as the sum of the first two, only two equation need to be satisfied for $(\alpha, \lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\alpha', \lambda', \rho', \gamma', \beta_1', \beta_2')$ to lead to the same reduced form, and the model is not identified. If instead

$$W^2W = \theta_1 I + \theta_2 W_{-1} + \theta_3 W + \theta_4 W W_{-1} + \theta_5 W^2, \tag{1.81}$$

with $\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5 = 1$. By substitution into (1.70), the two set of parameters lead to the same reduced form model if and only if

$$\begin{cases} \rho - \rho' + \theta_1(\lambda'\gamma - \lambda\gamma') = 0 \\ \gamma - \gamma' + \theta_2(\lambda'\gamma - \lambda\gamma') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\lambda'\gamma - \lambda\gamma') = 0 \\ \gamma' - \gamma + \lambda\gamma' - \lambda'\gamma + \theta_4(\lambda'\gamma - \lambda\gamma') = 0 \\ \lambda'\rho - \lambda\rho' + \theta_5(\lambda'\gamma - \lambda\gamma') = 0 \end{cases} \tag{1.82}$$

Since the last equation of (1.82) can be obtained as the sum of the first four, an equivalent

system with the same solution set is given by

$$\begin{cases} \rho - \rho' + \lambda'\rho - \lambda\rho' + (\theta_1 + \theta_5)(\lambda'\gamma - \lambda\gamma') = 0 \\ \gamma - \gamma' + \theta_2(\lambda'\gamma - \lambda\gamma') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\lambda'\gamma - \lambda\gamma') = 0 \\ \gamma' - \gamma + \lambda\gamma' - \lambda'\gamma + \theta_4(\lambda'\gamma - \lambda\gamma') = 0 \end{cases} \tag{1.83}$$

Again, we have that the last equation can be obtained as the sum of the first three, which implies

$$\begin{cases} \rho - \rho' + \lambda'\rho - \lambda\rho' + (\theta_1 + \theta_5)(\lambda'\gamma - \lambda\gamma') = 0 \\ \lambda\gamma' - \lambda'\gamma + (\theta_2 + \theta_4)(\lambda'\gamma - \lambda\gamma') = 0 \\ \rho' - \rho + \lambda\rho' - \lambda'\rho + \theta_3(\lambda'\gamma - \lambda\gamma') = 0 \end{cases} \tag{1.84}$$

From (1.84) it is easy to see that only two equations need to be satisfied for $(\alpha, \lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\alpha', \lambda', \rho', \gamma', \beta_1', \beta_2')$ to lead to the same reduced form, and the model is not identified. Since I, W, W_{-1}, WW_{-1}, W^2 and W^2W_{-1} are all block-diagonal matrices, we have that model (1.64) is identified if and only if $\exists\, t$ such that the matrices I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 and $W_t^2W_{t-1}$ are linearly independent. □

We can now prove the following proposition.

Proposition 1.A.2. *Suppose that the usual mild conditions are satisfied and that* $\rho\gamma \neq 0$ *or* $\beta_2\beta_2 \neq 0$. *If there exist a* t *such that the network* $W_t^2W_{t-1}$ *has diameter greater or equal than* 3, *then model* (1.64) *is identified.*

Proof. Let's assume that $\exists\ t$ such that the network $W_t^2W_{t-1}$ has diameter greater or equal than 3. This implies that we can find two funds i and j in $W_t^2W_{t-1}$ such that $(W_t^2W_{t-1})_{ij} > 0$, and $(W_t^2)_{ij} = (W_tW_{t-1})_{ij} = (W_t)_{ij} = (W_{t-1})_{ij} = 0$. Hence, I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 and $W_t^2W_{t-1}$ are linearly independent and model (1.64) is identified. □

Global Demeaning

As standard in the panel literature and similarly to [Lee and Yu, 2010], we can also eliminate the time fixed effects using a global demeaning operator. Hence, let define the demeaning operator J_n as $J_n = I_n - \frac{1}{n}\boldsymbol{\iota}_n\boldsymbol{\iota}_n' = I_n - H_n$. Since W_t is row-normalized, we have that $J_nW_t = J_nW_t(J_n + \frac{1}{n}\boldsymbol{\iota}_n\boldsymbol{\iota}_n') = J_nW_tJ_n$. Therefore, the demeaned structural model has no time fixed effects and can be written as:

$$J_n\boldsymbol{y}_t = \lambda J_nW_t\boldsymbol{y}_t + \rho J_n\boldsymbol{y}_{t-1} + \gamma J_nW_{t-1}\boldsymbol{y}_{t-1} + \beta_1 J_nW_{t-1}\boldsymbol{x}_{t-1} + \beta_2 J_n\boldsymbol{x}_{t-1} + J_n\boldsymbol{\epsilon}_t, \quad (1.85)$$

or equivalently

$$\begin{aligned} J_n\boldsymbol{y}_t = \lambda(J_nW_t)(J_n\boldsymbol{y}_t) + \rho(J_n\boldsymbol{y}_{t-1}) + \gamma(J_nW_{t-1})(J_n\boldsymbol{y}_{t-1}) \\ + \beta_1(J_nW_{t-1})(J_n\boldsymbol{x}_{t-1}) + \beta_2(J_n\boldsymbol{x}_{t-1}) + J_n\boldsymbol{\epsilon}_t. \end{aligned} \quad (1.86)$$

It is worth noticing that (1.86) is a spatial dynamic model without fixed effects similar to (1.50) but with demeaned variables. However, its identification cannot be proved by using Proposition 1.A.2 because the matrix J_nW_t is not row-normalized in general. Starting with the stacked structural model with time fixed effects (1.63) and considering the block diagonal global demeaning operator $\boldsymbol{I} - H$, we have the following form for the demeaned structural model

$$(\boldsymbol{I}-H)\boldsymbol{y} = (\boldsymbol{I}-H)\lambda W\boldsymbol{y}+(\boldsymbol{I}-H)(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1}+(\boldsymbol{I}-H)(\beta_2\boldsymbol{I}+\beta_1 W_{-1})\boldsymbol{x}_{-1}+(\boldsymbol{I}-H)\boldsymbol{\epsilon}, \quad (1.87)$$

and the following demeaned reduced form model

$$\begin{aligned} (\boldsymbol{I} - H)\boldsymbol{y} =& (\boldsymbol{I} - H)(\boldsymbol{I} - \lambda W)^{-1}(\rho\boldsymbol{I} + \gamma W_{-1})\boldsymbol{y}_{-1} + \\ &+(\boldsymbol{I} - H)(\boldsymbol{I} - \lambda W)^{-1}(\beta_2\boldsymbol{I} + \beta_1 W_{-1})\boldsymbol{x}_{-1} + \\ &+(\boldsymbol{I} - H)(\boldsymbol{I} - \lambda W)^{-1}\boldsymbol{\epsilon}. \end{aligned} \quad (1.88)$$

Lemma 1.A.3. *Suppose that W and W_{-1} are distinct, row normalized, do not contain zero rows and that I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 and $W_t^2W_{t-1}$ are linearly independent. If $\theta_1 I + \theta_2 W_{t-1} + \theta_3 W_t + \theta_4 W_t W_{t-1}$ has identical rows, then $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$.*

Proof. Let suppose that $\theta_1 I+\theta_2 W_{-1}+\theta_3 W+\theta_4 WW_{-1}$ has identical rows. Since $W\boldsymbol{\iota} = \boldsymbol{\iota}$, left multiplying $\theta_1 I+\theta_2 W_{-1}+\theta_3 W+\theta_4 WW_{-1}$ by W does not change the matrix. Hence

$$\theta_1 I + \theta_2 W_{-1} + \theta_3 W + \theta_4 WW_{-1} = \theta_1 W + \theta_2 WW_{-1} + \theta_3 W^2 + \theta_4 W^2 W_{-1}, \tag{1.89}$$

which implies

$$\theta_1 I + \theta_2 W_{-1} + (\theta_3 - \theta_1) W + (\theta_4 - \theta_2) WW_{-1} + \theta_3 W^2 + \theta_4 W^2 W_{-1} = 0. \tag{1.90}$$

Since by assumption I, W, W_{-1}, WW_{-1}, W^2 and W^2W_{t-1} are linearly independent, then we have that $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$. □

Lemma 1.A.4. *Suppose that the usual mild conditions are satisfied and that $\rho\gamma \neq 0$ or $\beta_2\beta_2 \neq 0$. If $\exists\, t$ such that I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 and $W_t^2W_{t-1}$ are linearly independent, then model* (1.88) *is identified.*

Proof. Let assume that $(\lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\lambda', \rho', \gamma', \beta_1', \beta_2')$ are two set of structural parameters that lead to the same reduced form model (1.88) and let define $\boldsymbol{\eta} = (\lambda, \rho, \gamma)$ and $\boldsymbol{\beta} = (\lambda, \beta_1, \beta_2)$. Hence, we have that $\forall\, \boldsymbol{y}_{t-1}$ and $\forall\, \boldsymbol{x}_{t-1}$

$$(\boldsymbol{I}-\lambda W_t)^{-1}(\rho \boldsymbol{I}+\gamma W_{t-1})\boldsymbol{y}_{t-1}-\varphi(\boldsymbol{y}_{t-1}, \boldsymbol{\eta})\boldsymbol{\iota} = (\boldsymbol{I}-\lambda' W_t)^{-1}(\rho' \boldsymbol{I}+\gamma' W_{t-1})\boldsymbol{y}_{t-1}-\varphi(\boldsymbol{y}_{t-1}, \boldsymbol{\eta}')\boldsymbol{\iota}, \tag{1.91}$$

and

$$(\boldsymbol{I}-\lambda W_t)^{-1}(\beta_2 \boldsymbol{I}+\beta_1 W_{t-1})\boldsymbol{x}_{t-1}-\varphi(\boldsymbol{x}_{t-1}, \boldsymbol{\beta})\boldsymbol{\iota} = (\boldsymbol{I}-\lambda' W_t)^{-1}(\beta_2' \boldsymbol{I}+\beta_1' W_{t-1})\boldsymbol{x}_{t-1}-\varphi(\boldsymbol{x}_{t-1}, \boldsymbol{\beta}')\boldsymbol{\iota}, \tag{1.92}$$

where

$$\varphi(\boldsymbol{y}_{t-1}, \boldsymbol{\eta}) = \frac{1}{N}\boldsymbol{\iota}'(\boldsymbol{I} - \lambda W_t)^{-1}(\rho \boldsymbol{I} + \gamma W_{t-1})\boldsymbol{y}_{t-1}, \tag{1.93}$$

$$\varphi(\boldsymbol{x}_{t-1}, \boldsymbol{\beta}) = \frac{1}{N}\boldsymbol{\iota}'(\boldsymbol{I} - \lambda W_t)^{-1}(\beta_2 \boldsymbol{I} + \beta_1 W_{t-1})\boldsymbol{x}_{t-1}. \tag{1.94}$$

Multiplying (1.91) by $(\boldsymbol{I} - \lambda' W_t)(\boldsymbol{I} - \lambda W_t)$ we obtain

$$\begin{aligned}[(\rho - \rho')\boldsymbol{I} + (\gamma - \gamma')W_{t-1} + (\lambda\rho' - \lambda'\rho)W_t + (\lambda\gamma' - \lambda'\gamma)W_t W_{t-1}]\boldsymbol{y}_{t-1} = \\ = (1 - \lambda')(1 - \lambda)[\varphi(\boldsymbol{y}_{t-1}, \boldsymbol{\eta}) - \varphi(\boldsymbol{y}_{t-1}, \boldsymbol{\eta}')]\boldsymbol{\iota}.\end{aligned} \tag{1.95}$$

Since the right-hand side of equation (1.95) is a constant vector and since (1.95) must be satisfied for every $\boldsymbol{y}_{t-1}$, we have that the matrix $(\rho - \rho')\boldsymbol{I} + (\gamma - \gamma')W_{t-1} + (\lambda\rho' - \lambda'\rho)W_t + (\lambda\gamma' - \lambda'\gamma)W_t W_{t-1}$ has equal rows. From Lemma 1.A.3, if I, W_t, W_{t-1}, $W_t W_{t-1}$, W_t^2 and $W_t^2 W_{t-1}$ are linearly independent, then

$$\begin{cases} \rho - \rho' = 0 \\ \lambda\rho' - \lambda'\rho = 0 \\ \gamma - \gamma' = 0 \\ \lambda\gamma' - \lambda'\gamma = 0 \end{cases} \tag{1.96}$$

Since, by assumption, $\rho\gamma \neq 0$, system (1.96) implies that $\lambda = \lambda'$, $\rho = \rho'$, $\gamma = \gamma'$. Similarly we can prove that, if I, W_t, W_{t-1}, $W_t W_{t-1}$, W_t^2 and $W_t^2 W_{t-1}$ are linearly independent, then also $\beta_1 = \beta_1'$ and $\beta_2 = \beta_2'$. Hence, model (1.88) is identified. □

1.A.3 With fund and time fixed effects

If we have both time and fund fixed effects, the structural equation of the model can be written as:

$$\boldsymbol{y}_t = \alpha_t \boldsymbol{\iota} + \boldsymbol{\zeta}_n + \lambda W_t \boldsymbol{y}_t + \rho \boldsymbol{y}_{t-1} + \gamma W_{t-1}\boldsymbol{y}_{t-1} + \beta_1 W_{t-1}\boldsymbol{x}_{t-1} + \beta_2 \boldsymbol{x}_{t-1} + \boldsymbol{\epsilon}_t, \tag{1.97}$$

where $\boldsymbol{\zeta}_n$ is the vector of fund fixed effects, which is equal to

$$\boldsymbol{\zeta}_n = (f_1, f_2, \ldots, f_n)'. \tag{1.98}$$

It is convenient to rewrite model (1.97) in a stacked structural form as:

$$\boldsymbol{y} = \boldsymbol{\alpha} \otimes \boldsymbol{\iota} + \boldsymbol{\iota_T} \otimes \boldsymbol{\zeta}_n + \lambda W \boldsymbol{y} + \rho \boldsymbol{y}_{-1} + \gamma W_{-1} \boldsymbol{y}_{-1} + \beta_1 W_{-1} \boldsymbol{x}_{-1} + \beta_2 \boldsymbol{x}_{-1} + \boldsymbol{\epsilon}, \tag{1.99}$$

with $\boldsymbol{\iota_T}$ being the $T \times 1$ column vector of ones. Time fixed effects $\boldsymbol{\alpha} \otimes \boldsymbol{\iota}$ can be removed as in Appendix 1.A.2 using the global demeaning operator $\boldsymbol{I} - H$, i.e.

$$\begin{aligned}(\boldsymbol{I} - H)\boldsymbol{y} = (\boldsymbol{I} - H)\lambda W \boldsymbol{y} + (\boldsymbol{I} - H)(\rho \boldsymbol{I} + \gamma W_{-1})\boldsymbol{y}_{-1} \\ + (\boldsymbol{I} - H)(\beta_2 \boldsymbol{I} + \beta_1 W_{-1})\boldsymbol{x}_{-1} + (\boldsymbol{I} - H)(\boldsymbol{\iota_T} \otimes \boldsymbol{\zeta}_n) + \boldsymbol{\nu},\end{aligned} \tag{1.100}$$

where $(\boldsymbol{I} - H)(\boldsymbol{\iota_T} \otimes \boldsymbol{\zeta}_n)$ could be interpreted as transformed fund fixed effects. In order to remove the fund fixed effects, we use the *Forward Orthogonal Deviation* (FOD) operator which has been widely employed in the dynamic panel literature [Arellano and Bover, 1995] and, recently, also applied to spatial dynamic panel models with fixed effects [Lee and Yu, 2014]. In our case, the FOD operator is defined as:

$$B = F'_{T,T-1} \otimes \boldsymbol{I}_N, \tag{1.101}$$

where $F_{T,T-1}$ is $T \times T-1$ matrix given by

$$F_{T,T-1} = diag\left[\frac{T-1}{T}, \quad \ldots, \quad \frac{1}{2}\right]^{\frac{1}{2}} \times \tag{1.102}$$

$$\begin{bmatrix} 1 & -(T-1)^{-1} & -(T-1)^{-1} & \ldots & -(T-1)^{-1} & -(T-1)^{-1} & -(T-1)^{-1} \\ 0 & 1 & -(T-2)^{-1} & \ldots & -(T-2)^{-1} & -(T-2)^{-1} & -(T-2)^{-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -1 \end{bmatrix}' . \tag{1.103}$$

The intuition behind the operator B is that $\forall\, i$ and $\forall\, t = 1, \ldots, T-1$, the observation y_{it} is transformed into $y^*_{it} = \sqrt{\frac{T-t}{T-t+1}} \cdot \left[y_{it} - \frac{1}{T-t}\sum_{h=t+1}^{T} y_{ih}\right]$, hence y^*_{it} depends only on current and future values, but not on past ones. The advantage of using the operator B is that it does not introduce autocorrelation in the error term, but requires the loss of the last temporal observation for each cross-sectional unit. After the application of B, the demeaned structural equation becomes:

$$B(\boldsymbol{I}-H)\boldsymbol{y} = B(\boldsymbol{I}-H)\lambda W\boldsymbol{y} + B(\boldsymbol{I}-H)(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1} + B(\boldsymbol{I}-H)(\beta_2\boldsymbol{I}+\beta_1 W_{-1})\boldsymbol{x}_{-1} + \boldsymbol{\nu}^*, \tag{1.104}$$

Similarly, the demeaned reduced form model without fixed effects is given by:

$$\begin{aligned} B(\boldsymbol{I}-H)\boldsymbol{y} =& B(\boldsymbol{I}-H)(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1} + \\ &+ B(\boldsymbol{I}-H)(\boldsymbol{I}-\lambda W)^{-1}(\beta_2\boldsymbol{I}+\beta_1 W_{-1})\boldsymbol{x}_{-1} + \\ &+ B(\boldsymbol{I}-H)(\boldsymbol{I}-\lambda W)^{-1}(\boldsymbol{\iota_T}\otimes\boldsymbol{\zeta}_n) + \\ &+ B(\boldsymbol{I}-H)(\boldsymbol{I}-\lambda W)^{-1}\boldsymbol{\epsilon}. \end{aligned} \tag{1.105}$$

We observe that, while the fund fixed effects are in general still present in the reduced form model, our focus remains on the identification of the structural parameters $(\lambda, \rho, \gamma, \beta_1, \beta_2)$.

Lemma 1.A.5. *Suppose that the usual mild conditions are satisfied and that* $\rho\gamma \neq 0$ *or* $\beta_2\beta_2 \neq 0$. *If* $\exists t$ *such that* I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 *and* $W_t^2W_{t-1}$ *are linearly independent, then model* (1.104) *is identified.*

Proof. Let define $\boldsymbol{\eta} = (\lambda, \rho, \gamma)$ and $\boldsymbol{\beta} = (\lambda, \beta_1, \beta_2)$. The demeaned reduced form (1.105) can be rewritten as:

$$\begin{aligned} B(\boldsymbol{I} - H)\boldsymbol{y} =& B(\boldsymbol{I} - \lambda W)^{-1}(\rho\boldsymbol{I} + \gamma W_{-1})\boldsymbol{y}_{-1} - B\boldsymbol{\varphi}(\boldsymbol{y}_{-1}, \boldsymbol{\eta}) + \\ &+ B(\boldsymbol{I} - \lambda W)^{-1}(\beta_2\boldsymbol{I} + \beta_1 W_{-1})\boldsymbol{x}_{-1} - B\boldsymbol{\varphi}(\boldsymbol{x}_{-1}, \boldsymbol{\eta}) + \\ &+ B(\boldsymbol{I} - H)(\boldsymbol{I} - \lambda W)^{-1}(\boldsymbol{\iota_T} \otimes \boldsymbol{\zeta}_n) + \\ &+ B(\boldsymbol{I} - H)(\boldsymbol{I} - \lambda W)^{-1}\boldsymbol{\epsilon}, \end{aligned} \tag{1.106}$$

where

$$\boldsymbol{\varphi}(\boldsymbol{y}_{-1}, \boldsymbol{\eta}) = H(\boldsymbol{I} - \lambda W)^{-1}(\rho\boldsymbol{I} + \gamma W_{-1})\boldsymbol{y}_{-1} \tag{1.107}$$

$$\boldsymbol{\varphi}(\boldsymbol{x}_{-1}, \boldsymbol{\beta}) = H(\boldsymbol{I} - \lambda W)^{-1}(\beta_2\boldsymbol{I} + \beta_1 W_{-1})\boldsymbol{x}_{-1}. \tag{1.108}$$

Let $(\lambda, \rho, \gamma, \beta_1, \beta_2)$ and $(\lambda', \rho', \gamma', \beta_1', \beta_2')$ be two set of structural parameters that lead to the same reduced form model (1.106). Hence, we have that $\forall\, \boldsymbol{y}_{-1}$ and $\forall\, \boldsymbol{x}_{-1}$:

$$B(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1}-B\boldsymbol{\varphi}(\boldsymbol{y}_{-1}, \boldsymbol{\eta}) = B(\boldsymbol{I}-\lambda' W)^{-1}(\rho'\boldsymbol{I}+\gamma' W_{-1})\boldsymbol{y}_{-1}-B\boldsymbol{\varphi}(\boldsymbol{y}_{-1}, \boldsymbol{\eta'}), \tag{1.109}$$

and

$$B(\boldsymbol{I}-\lambda W)^{-1}(\beta_2\boldsymbol{I}+\beta_1 W_{-1})\boldsymbol{x}_{-1}-B\boldsymbol{\varphi}(\boldsymbol{x}_{-1}, \boldsymbol{\beta}) = B(\boldsymbol{I}-\lambda' W)^{-1}(\beta_2'\boldsymbol{I}+\beta_1' W_{-1})\boldsymbol{x}_{-1}-B\boldsymbol{\varphi}(\boldsymbol{x}_{-1}, \boldsymbol{\beta'}). \tag{1.110}$$

From (1.109) we have that:

$$B\big[(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1}-(\boldsymbol{I}-\lambda' W)^{-1}(\rho'\boldsymbol{I}+\gamma' W_{-1})\boldsymbol{y}_{-1}+\boldsymbol{\varphi}(\boldsymbol{y}_{-1}, \boldsymbol{\eta'})-\boldsymbol{\varphi}(\boldsymbol{y}_{-1}, \boldsymbol{\eta})\big] = \boldsymbol{0}. \tag{1.111}$$

The matrix B is right invertible with null space of dimension 1 that is spanned by the

unit vector, therefore $\forall\, \boldsymbol{y}_{-1}\ \exists\, k_y$ such that:

$$(\boldsymbol{I}-\lambda W)^{-1}(\rho\boldsymbol{I}+\gamma W_{-1})\boldsymbol{y}_{-1}-(\boldsymbol{I}-\lambda' W)^{-1}(\rho'\boldsymbol{I}+\gamma' W_{-1})\boldsymbol{y}_{-1}+\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta'})-\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta})=k_y\boldsymbol{\iota}. \tag{1.112}$$

Similarly, from (1.110) we have that

$$(\boldsymbol{I}-\lambda W)^{-1}(\beta_2\boldsymbol{I}+\beta_1 W_{-1})\boldsymbol{x}_{-1}-(\boldsymbol{I}-\lambda' W)^{-1}(\beta_2'\boldsymbol{I}+\beta_1' W_{-1})\boldsymbol{x}_{-1}+\boldsymbol{\varphi}(\boldsymbol{x}_{-1},\boldsymbol{\beta'})-\boldsymbol{\varphi}(\boldsymbol{x}_{-1},\boldsymbol{\beta})=k_x\boldsymbol{\iota}. \tag{1.113}$$

Multiplying (1.112) by $(\boldsymbol{I}-\lambda' W)(\boldsymbol{I}-\lambda W)$ we obtain that $\forall \boldsymbol{y}_{-1}$:

$$\begin{aligned}\big[(\rho-\rho')\boldsymbol{I}+(\gamma-\gamma')W_{-1}+(\lambda\rho'-\lambda'\rho)W+(\lambda\gamma'-\lambda'\gamma)WW_{-1}\big]\boldsymbol{y}_{-1}=\\ (\boldsymbol{I}-\lambda' W)(\boldsymbol{I}-\lambda W)\big[\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta})-\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta'})\big]+k_y(1-\lambda)(1-\lambda')\boldsymbol{\iota}.\end{aligned} \tag{1.114}$$

The matrix inside the square brackets on the left-hand side of (1.114) is a block-diagonal matrix, $k_y(1-\lambda)(1-\lambda')\boldsymbol{\iota}$ is a constant vector, while $\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta})-\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta'})$ is a stacked vector that can be rewritten as

$$\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta})-\boldsymbol{\varphi}(\boldsymbol{y}_{-1},\boldsymbol{\eta'})= \tag{1.115}$$

$$\Big(\varphi(\boldsymbol{y}_0,\boldsymbol{\eta})-\varphi(\boldsymbol{y}_0,\boldsymbol{\eta'}),\ldots,\varphi(\boldsymbol{y}_{t-1},\boldsymbol{\eta})-\varphi(\boldsymbol{y}_{t-1},\boldsymbol{\eta'}),\ldots,\varphi(\boldsymbol{y}_{T-1},\boldsymbol{\eta})-\varphi(\boldsymbol{y}_{T-1},\boldsymbol{\eta'})\Big)'. \tag{1.116}$$

It is easy to see that if we focus on a generic time instant $t-1$, equation (1.114) implies that

$$(\rho-\rho')\boldsymbol{I}+(\gamma-\gamma')W_{t-1}+(\lambda\rho'-\lambda'\rho)W_t+(\lambda\gamma'-\lambda'\gamma)W_tW_{t-1}$$

has identical rows. Hence, by Lemma 1.A.3, we can prove that if I, W_t, W_{t-1}, W_tW_{t-1}, W_t^2 and $W_t^2W_{t-1}$ are linearly independent, model (1.105) is identified. □

Appendix 1.B Robustness Analysis

In this appendix we study the robustness of peer flow effects to different methodologies and similarity measures for the construction of the similarity temporal network. In particular, focusing on all funds except Treasury MMMFs, as in Section 1.6, in the first 4 cases we fix the similarity measure to be the weighted maturity similarity WMS and consider different ways to define the network, while in the last 4 cases we fix network construction methodology to be the same as the one described in Section 1.5 and consider different similarity measures. Hence, we consider a total of eight cases as described below:

Case 1: Here the network construction methodology is similar to the one described in Section 1.5, but outgoing links from a fund i are created only with the top 5% of most similar funds according with the WMS measure.

Case 2: Here the network construction methodology is similar to the one described in Section 1.5, but outgoing links from a fund i are created only with the top 15% of most similar funds according with the WMS measure.

Case 3: Here the similarity temporal network is constructed by linking together all the funds belonging to the same fund category.

Case 4: Here the similarity temporal network is constructed by linking together all the funds with a WMS measure different from zero, weighting each link by the similarity itself.

Case 5: Here we measure the similarity between fund portfolios using the cosine similarity. In particular, let $\mathcal{B}$ be the set of all possible borrowers and let $|\mathcal{B}| = B$. For each fund i, time period t and borrower $b \in \mathcal{B}$, $w_{i,t}^b$ is the weight of issuer b in the portfolio of fund i at time t. Hence, the portfolio of fund i at time t can be represented by the B-dimensional vector

$$w_{i,t} = \left(w_{i,t}^1, w_{i,t}^2, \ldots, w_{i,t}^B\right). \tag{1.117}$$

Thus, if i and j are two funds, the cosine similarity between i and j at time t is given by:

$$\text{CosineSimilarity}_t(i,j) = \frac{w_{i,t} \cdot w_{j,t}}{||w_{i,t}||\,||w_{j,t}||}. \tag{1.118}$$

The similarity temporal network is then constructed creating outgoing links from a fund i only with the top 10% of most similar funds.

Case 6: Here we measure the similarity between fund portfolios using a measure, that we call *Weighted Life Similarity* that is similar to the WMS but uses the weighted life instead of weighted maturity as a proxy for interest rate risk. In particular, let $\mathcal{B}$ be the set of all possible borrowers and let $|\mathcal{B}| = B$. For each fund i and borrower $b \in \mathcal{B}$, we define the *Weighted Life Exposure* of i to b at time t as:

$$\text{WLE}^b_{f,t} = \sum_{k=1}^{k_b} w_k l_k, \tag{1.119}$$

where k_b is the number of instruments in the portfolio of i whose issuer is b, w_k is the weight of that instrument and l_k its effective life. Then, for each fund i we have the B-dimensional vector

$$\text{WLE}_{i,t} = \Big(\text{WLE}^1_{i,t}, \text{WLE}^2_{i,t}, \ldots, \text{WLE}^B_{i,t}\Big). \tag{1.120}$$

of the weighted life exposures. If i and j are two funds, we define the *Weighted Life Similarity* between i and j at time t, denoted $\text{WLS}_t(i,j)$, as the cosine similarity between their vectors of weighted life exposures, i.e.:

$$\text{WLS}_t(i,j) = \frac{\text{WLE}_{i,t} \cdot \text{WLE}_{j,t}}{||\text{WLE}_{i,t}||\,||\text{WLE}_{j,t}||}. \tag{1.121}$$

The similarity temporal network is then constructed creating outgoing links from a fund i only with the top 10% of most similar funds.

Case 7: Here we measure the similarity between fund portfolios using the weighted maturity similarity that is computed not only on the borrower level, but on the borrower-

asset class level. In particular, let $\mathcal{B}$ be the set of all possible borrowers and $\mathcal{A}$ be the set of all possible instrument asset classes, with $|\mathcal{B}| = B$ and $|\mathcal{A}| = A$. For each fund i, borrower $b \in \mathcal{B}$ and asset class $a \in \mathcal{A}$, we define the *Weighted Maturity Exposure* of i to b and a at time t as:

$$\mathrm{WME}_{i,t}^{a,b} = \sum_{k=1}^{k_{a,b}} w_k m_k, \tag{1.122}$$

where $k_{a,b}$ is the number of instruments in the portfolio of i whose issuer is b that belong to the asset class a, w_k is the weight of that instrument and m_k its effective maturity. Then, for each fund i we have the $A \times B$-dimensional vector

$$\mathrm{WME}_{i,t}^{A,B} = \left(\mathrm{WLE}_{i,t}^{1,1}, \mathrm{WLE}_{i,t}^{2,1}, \ldots, \mathrm{WLE}_{i,t}^{A,1}, \ldots, \mathrm{WLE}_{i,t}^{A,B}\right). \tag{1.123}$$

of the weighted maturity exposures. If i and j are two funds, we define the *Weighted Maturity Similarity* between i and j at time t, denoted $\mathrm{WMS}_t^{A,B}(i,j)$, as the cosine similarity between their vectors of weighted maturity exposures, i.e.:

$$\mathrm{WMS}_t^{A,B}(i,j) = \frac{\mathrm{WME}_{i,t}^{A,B} \cdot \mathrm{WME}_{j,t}^{A,B}}{||\mathrm{WME}_{i,t}^{A,B}||\,||\mathrm{WME}_{j,t}^{A,B}||}. \tag{1.124}$$

The similarity temporal network is then constructed creating outgoing links from a fund i only with the top 10% of most similar funds.

Case 8: Here we measure the similarity between fund portfolios using the weighted maturity similarity that is computed not on the borrower level, but only on the asset class level. In particular, let $\mathcal{A}$ be the set of all possible instrument asset classes, with $|\mathcal{A}| = A$. For each fund i and asset class $a \in \mathcal{A}$, we define the *Weighted Maturity Exposure* of i to a at time t as:

$$\mathrm{WME}_{i,t}^{a} = \sum_{k=1}^{k_a} w_k m_k, \tag{1.125}$$

where k_a is the number of instruments in the portfolio of i that belong to the asset class a, w_k is the weight of that instrument and m_k its effective maturity. Then, for each fund i we have the A-dimensional vector

$$\mathrm{WME}_{i,t}^{A} = \Big(\mathrm{WLE}_{i,t}^{1}, \mathrm{WLE}_{i,t}^{2}, \ldots, \mathrm{WLE}_{i,t}^{A}\Big). \tag{1.126}$$

of the weighted maturity exposures. If i and j are two funds, we define the *Weighted Maturity Similarity* between i and j at time t, denoted $\mathrm{WMS}_{t}^{A}(i,j)$, as the cosine similarity between their vectors of weighted maturity exposures, i.e.:

$$\mathrm{WMS}_{t}^{A}(i,j) = \frac{\mathrm{WME}_{i,t}^{A} \cdot \mathrm{WME}_{j,t}^{A}}{||\mathrm{WME}_{i,t}^{A}||\,||\mathrm{WME}_{j,t}^{A}||}. \tag{1.127}$$

The similarity temporal network is then constructed creating outgoing links from a fund i only with the top 10% of most similar funds.

From Panel A of Table 1.5 we see that the estimate of peer effects remains positive and statistically significant if we change the threshold for the construction of the similarity temporal network from the top 10% to the top 5% or top 15% of most similar funds. Moreover, the analysis of Case 3 and 4 shows that this result is not specific of the way we have used to define the spatial dependence. Indeed, if we construct the network by simply linking together funds belonging to the same category, we still find a positive and statistically significant spatial correlation coefficient, and the same is true if the network is constructed by linking all the funds with a similarity different from zero and each link is weighted by the similarity itself.

Moreover, from Panel B of Table 1.5 peer effects also appear to be robust to the choice of the particular similarity measure. For instance, peer effects remain positive and statistically significant when the similarity measure does not uses any proxy for interest rate risk (Case 5) or uses a different proxy such as instrument effective life (Case 6). The same is true when the weighted maturity similarity is computed at joint level issuer-asset class (Case 7) or only at the asset class level (Case 8).

The estimated impact measures and their partitioning as function of the neighbor order are shown in Tables 1.6 and 1.7 for the cases using different network construction methodologies and different similarity measures, respectively. In each case we find that the indirect impact of a flow shock to first and second degree neighbors is positive and statistically significant, except for Case 7 where only first degree indirect impact are statistically significant.

1.B.1 Quasi-Random Experiment

One identification challenge for Peer Effects, besides the reflection problem, may be due to the presence of unobserved common shocks that affects the flows of similar funds. In order to tackle this problem we design a quasi-random experiment exploiting the funds different exposure to European banks during summer 2011. The success of the experiment is deeply connected to the choice of the treatment variable and its ability to provide a source of exogenous variation, not driven by unobserved common factors, that is relevant for fund peer flows but otherwise random for fund own flows.

The intuition behind our experiment choice is based on the evidence that, during the peak of the European Sovereign Debt Crisis, money market funds with high exposure to European Banks suffered larger outflows relative to funds with low exposure [Chernenko and Sunderam, 2014]. Hence, our treatment variable is constructed as the interaction of two dummy variables. The first is a dummy that is equal to one for the funds with European exposure in the top quartile, while the second one is a variable that identifies the crisis period, i.e. is equal to one if the date is between May and September 2011. Hence the model estimated in this quasi-random experiment is given by:

$$\boldsymbol{y}_t = \alpha_t \boldsymbol{\iota} + \boldsymbol{\zeta}_n + \lambda W_t \boldsymbol{y}_t^{Treatment} + \rho \boldsymbol{y}_{t-1} + \gamma W_{t-1} \boldsymbol{y}_{t-1} + \beta_1 W_{t-1} \boldsymbol{x}_{t-1} + \beta_2 \boldsymbol{x}_{t-1} + \boldsymbol{\mathcal{C}}_t \boldsymbol{\delta}' + \boldsymbol{\epsilon}_t, \tag{1.128}$$

where the i-th component of the peer treatment flow variable is equal to

$$\frac{\sum_{j \in P_{i,t}} y_{j,t} \mathbb{1}_{\{Treatment=1\}}}{|P_{i,t}|}. \tag{1.129}$$

A graphical analysis of our experiment is shown in Figure 1.10. The plot in the bottom panel shows the average flow over time for funds with relative high and low European exposure. Outside the crisis period, we observe a substantial similarity between the flows, which is also confirmed by looking at the flow distributions conditional on the dummy variable for European exposure (first plot in the top row). However, during the crisis period, we observe a sharp divergence in behavior, which justifies the definition of the treatment variable as an interaction, since the overall distribution of flows during the crisis period is qualitatively similar to the flow distribution outside the crisis (second plot in the top row). Thus, unobserved common factors affecting the flows of similar funds over time or in the cross section are differenced out by the treatment variable.

The strength and power of any results from this experiment is also connected to treatment group variation between a fund and its peers. Given the definition of the treatment variable, a fund belonging to one treatment group is likely to have peers, defined as the top 10% of most similar funds, also belonging to the same treatment group. If indeed there was no variation between the treatment group of a fund and that of its peers, it could be difficult to translate any evidence of peer effects outside of the experiment boundaries, even if theoretically justified. Fortunately, this is not the case here. The third plot in the first row of Figure 1.10 shows the distribution of the average peer treatment status for the two treatment groups. We see that funds with low European exposure have on average 20% of their peers belonging to the treatment group 1, with cases where this percentage can be as high as 60%. Hence, any evidence of peer effects would account for dependencies both within and between treatment groups.

From Table 1.8 we find a positive and strongly significant relationship between own flows and the treated version of peer flows, which is consistent with our main empirical results presented in Section 1.6 and also suggests that Peer Effects do not seems to be driven by unobserved common factors. The result is also robust to the inclusion of controls for own and peer European exposure.

Appendix 1.C Counterfactual Experiment

In order to better understand and interpret the empirical results in Section 1.6 we provide here a simulated counterfactual experiment study in which the model is estimated on data where the dependent variable $Flows_{i,t}$ ($y_{i,t}$) is simulated using different data generating processes. In particular, the objective of this section is to understand how the estimate of the spatial dependence coefficient λ and its statistical significance changes in cases when the assumed similarity temporal network does not reflect the real spatial dependence structure. For instance, we want to understand how the model coefficient estimate behaves when there is, actually, no spatial dependence between the flow of funds or when the real spatial dependence structure is different from the similarity temporal network used for the model estimation.
The data generating process for $\boldsymbol{y}_t$ is defined as:

$$\boldsymbol{y}_t = \alpha_t \boldsymbol{\iota} + \boldsymbol{\zeta}_n + \lambda G \boldsymbol{y}_t + \boldsymbol{\epsilon}_t, \tag{1.130}$$

where α_t and $\boldsymbol{\zeta}_n$ are, respectively, the time and fund fixed effects, G is the row-normalized adjacency matrix corresponding to the spatial structure, λ is the spatial dependence parameter and $\boldsymbol{\epsilon}_t$ is the vector of error terms distributed as $N(0, \sigma \boldsymbol{I})$. For the time and fund fixed effects we use the estimated effects from the data, moreover, we assume that $\lambda = 0.5$ and σ is equal to the empirical flow standard deviation. The simulation of (1.130), is then performed using the corresponding reduced form

$$\boldsymbol{y}_t = (\boldsymbol{I} - \lambda G)^{-1}(\alpha_t \boldsymbol{\iota} + \boldsymbol{\zeta}_n + \boldsymbol{\epsilon}_t), \tag{1.131}$$

In order to simulate the data generating process for the flow variable $\boldsymbol{y}_t$ we need, however, to define the spatial dependence structure. We assume that the network guiding the dependence between the funds belongs to the Erdös-Rényi family of random graphs $G(n, p)$ [Erdös et al., 1959], where n represents the number of nodes in the graph - in our case the number of funds - and p represents the probability of link formation. In

the $G(n,p)$ model, a graph is constructed by starting with n nodes and then edges are added independently from each other with probability p. The value of p is also equal to the expected density of the network. Hence, for each simulation, we use an adjacency matrix G that is randomly drawn from $G(n,p)$ and row-normalized. After simulating (1.131) for $t = 1, ..., T$, we estimate the model (1.2) using the similarity temporal network defined in Section 1.5 and the empirical data for the exogenous variables. We perform 1000 simulations for different values of the link probability p. The case in which $p = 0$ represent the situation where there exists no spatial dependence between the funds flow. For each value of p Table 1.9 reports the average estimated λ, the average standard error of the estimate and its average p-value. The simulation study shows that when there is no spatial dependence or when the real spatial structure is different from the one used for the estimation, on average, the spatial dependence parameter is not statistically significant. This results does not change by changing the expected density of the network.

Appendix 1.D Temporal Network Description

Figure 1.11 gives a static representation of the similarity temporal network, where only the links that are present at least 70% of the time are shown. We see that there is high connectivity within funds category and 4 clusters can be clearly identified.

It is interesting to look at the in-degree distribution of the network since the in-degree of a fund i represents the number of funds that could potentially be affected by a flow shock to i. Figure 1.12 shows the in-degree distribution of all funds (top left corner) aggregated over all the time periods. We see that there are funds with a low in-degree, but also funds that could potentially have an impact on the flows of more than 100 other MMMFs. This means that the in-degree is not uniform among funds and that the impact, on the rest of the industry, of a shock to a fund vastly depends on which is the fund that has suffered the shock. The other panels of Figure 1.12 show the in-degree distribution for the different categories of MMMFs. The Prime MMMFs in-degree distribution exhibits a positive skew, while the in-degree distribution of Treasury funds show a negative skew. In a temporal network setting it is natural to study how links between nodes change over time. In particular, it is interesting to analyze the persistence of the links in the network. A link between two funds at time t is said to be persistent if there is a non-negligible probability that they will be connected at time $t + \Delta t$. In order to measure *link persistence* (see Nicosia et al. [2013] and Tang et al. [2010]), let consider the unweighted temporal network $\{A_t\}_{t=0,\dots,T}$, with:

$$(A_t)_{ij} = \begin{cases} 1 & \text{if} \quad (W_t)_{ij} \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.132}$$

If we consider a fund i and two consecutive time instants t and $t+1$, the topological overlap of the neighborhood of i in t and $t+1$ is defined as:

$$C_i(t, t+1) = \frac{\sum_j (A_t)_{ij}(A_{t+1})_{ij}}{\sqrt{[\sum_j (A_t)_{ij}][\sum_j (A_{t+1})_{ij}]}}. \tag{1.133}$$

Taking the average of (1.133) over all possible consecutive time intervals $[t, t+1]$, we obtain a measure of the link persistence of fund i:

$$C_i = \frac{1}{T}\sum_{t=0}^{T-1} C_i(t, t+1). \tag{1.134}$$

C_i is equal to 1 if and only if fund i has the same links in all the possible time intervals $[t, t+1]$, and is equal to 0 if no link is ever observed in two consecutive time instants. An overall measure of link persistence for the temporal network $\{W_t\}_{t=0,\dots,T}$, called temporal correlation coefficient, can be obtained by averaging (1.134) over all the funds:

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{1.135}$$

The similarity temporal network defined in (1.15) has an high link persistence measure $C = 0.7883$ suggesting that the degree of similarity between funds tends to be stable over time. Figure 1.13 shows the link persistence distribution for all the funds in the network (top left panel) and for the different types of money market funds. If we focus on Prime funds, it is worth noticing that, if on average we observe a link persistence of 0.71, there are funds exhibiting a temporal correlation coefficient as high as 0.96.

Table 1.4: This table contains the estimated impact measures and their partitioning for the first 5 lags. Each panel is dedicated to one of the three fund universes considered in Table 1.3 and the impact measures are computed using the corresponding estimated parameters with both fund and time fixed effects. Significance Levels: 0.01 '***', 0.05 '**', 0.1 '*'.

Panel A: All Funds except Treasuries						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1270^{***}		0.3408^{**}	
Lag 1		0.7597^{***}		0.0713^{*}		0.5270^{***}
Lag 2	0.0108^{**}	0.5663^{**}	0.0008^{*}	0.0534^{*}	0.0072^{**}	0.3932^{**}
Lag 3	0.0054	0.4330	0.0004	0.0407	0.0037	0.3932
Lag 4	0.0035	0.3295	0.0003	0.0309	0.0024	0.2286
Total impacts	1.0285^{***}	3.1336	-0.1247^{***}	0.2947	0.3601^{**}	2.1743

Panel B: All Funds						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1340^{***}		0.2627^{*}	
Lag 1		0.3929^{***}		0.0786^{*}		0.5247^{***}
Lag 2	0.0022^{*}	0.1521	0.0004^{*}	0.0304^{*}	0.002^{*}	0.2033^{*}
Lag 3	0.0005	0.1521	0.0001	0.0120	0.0007	0.0802
Lag 4	0.0002	0.0236	$3.9 \cdot 10^{-5}$	0.0047	0.0002	0.0315
Total impacts	1.0031^{***}	0.6441^{*}	-0.1334^{***}	0.1289^{*}	0.2667	0.8603^{*}

Panel C: Only Prime Funds						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1113^{***}		0.2219	
Lag 1		0.3611^{*}		0.0957		-0.09948
Lag 2	0.0042	0.1261	0.0009	0.0336	-0.0009	-0.0350
Lag 3	0.0009	0.0461	0.0002	0.0122	-0.0002	-0.0127
Lag 4	0.0002	0.0167	$6.7 \cdot 10^{-5}$	0.0044	-0.0001	-0.0046
Total impacts	1.0056^{***}	0.5595	-0.1100^{***}	0.1486	0.2207	-0.1545

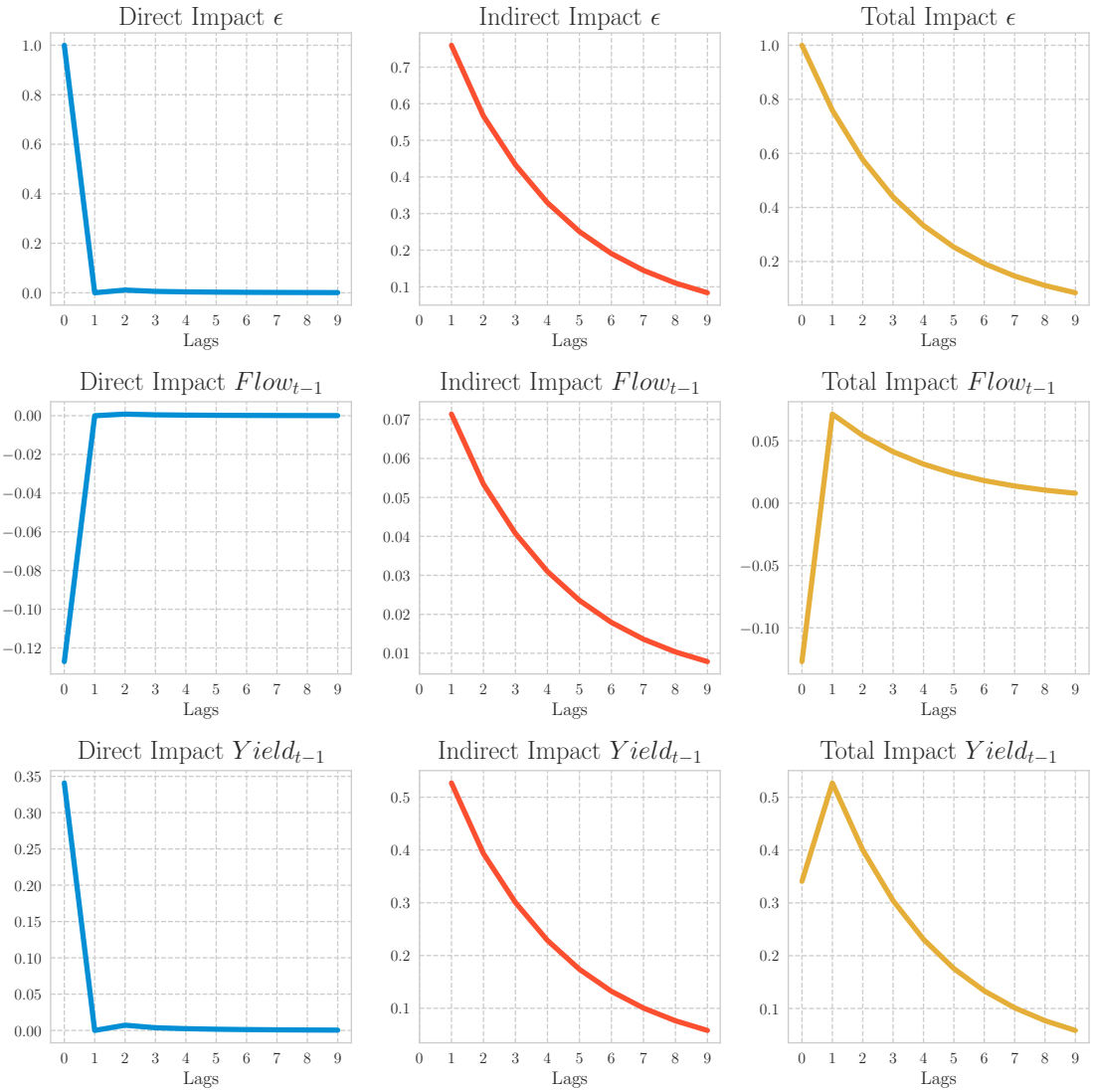


Figure 1.6: This Figure shows the first 10 lag partitioning of the impact measures for model (3) in Table 1.3. The first row contains the impact measure plots for a contemporaneous flow shock ϵ, while the second and third row contains the impact measures for past flow and past yield, respectively.

Figure 1.7: This Figure shows the scatter plot of a fund in-degree centrality and the total impact of a 10% flow shock to that fund according to the impact matrix (1.46). Each panel refers to the network snapshot of a single date.

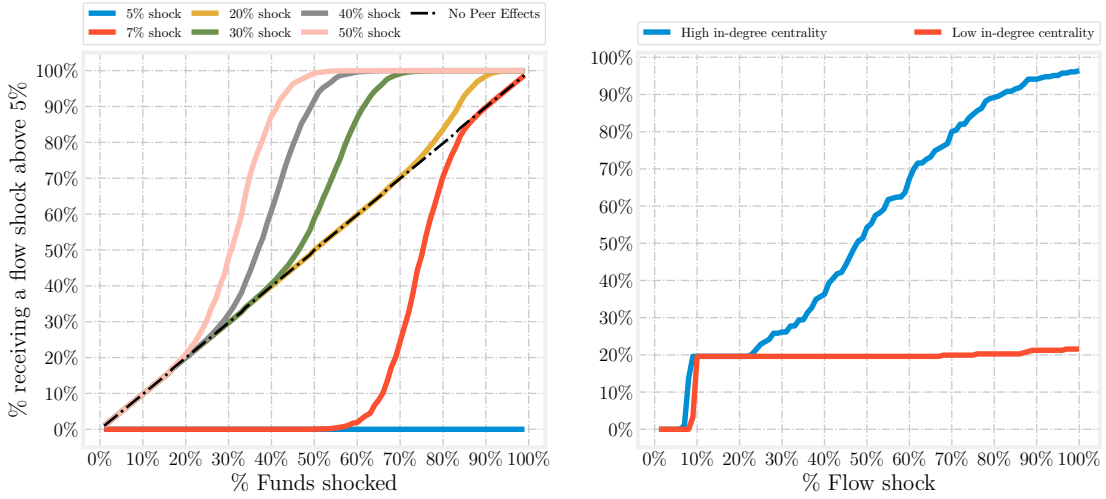


Figure 1.8: This Figure describes the implications of model (3) in Table 1.3 for the resilience of the Money Market Mutual Fund Industry. Using the September 2015 snapshot of the similarity temporal network defined in Section 1.5, the left panel shows how the percentage of funds forced to sell less liquid assets changes as a function of the percentage of fund shocked under different assumptions on the magnitude of the flow shock. The black dashed line shows the relationship under the assumption of no peer effects. Instead, the right panel shows how the percentage of funds forced to sell less liquid assets changes with respect to the magnitude of a flow shock when the funds shocked are the 20% of funds with the highest (blue line) and lowest (red line) in-degree centrality.

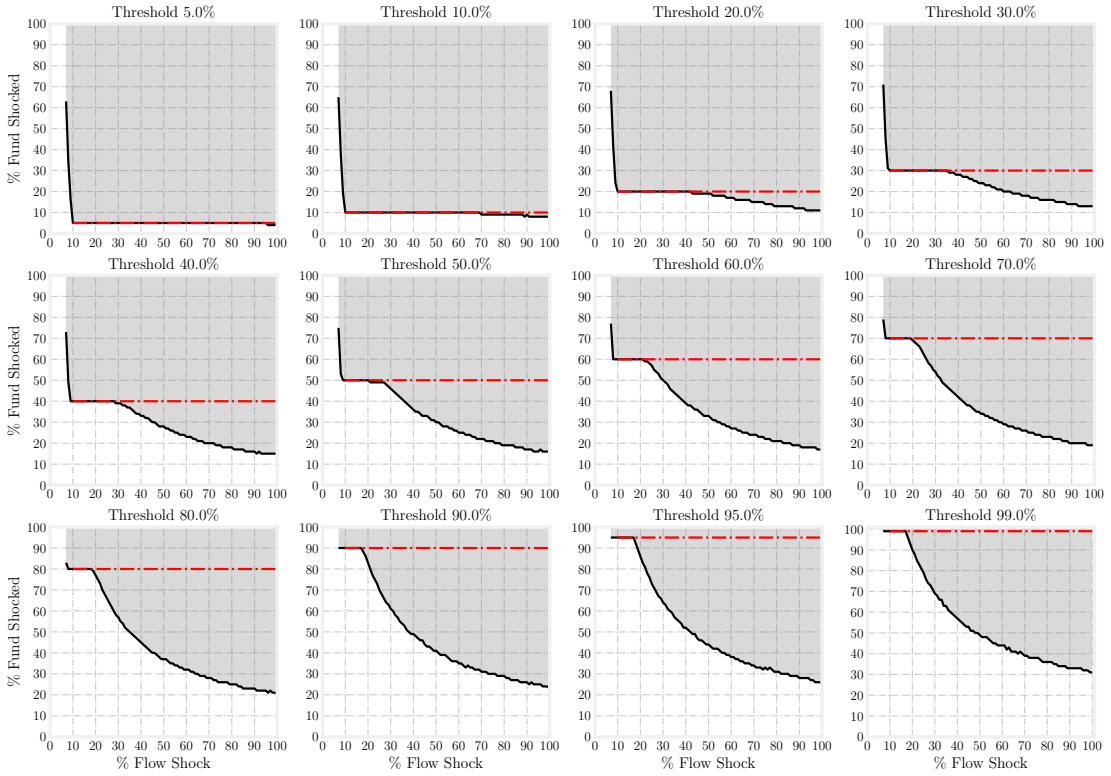


Figure 1.9: This Figure shows the resilience regions of Money Market Mutual Fund Industry, as defined in Section 1.7, for the September 2015 snapshot of the similarity temporal network. Each panel considers a different threshold, from 5% to 99%, that represents the percentage of funds that need to be simultaneously receiving a flow shock greater than the percentage of daily liquid assets in order for fire-sales spillovers to be triggered. Each plot uses as axes the two dimensions of a run like event: the x-axis is the magnitude of the flow shock and the y-axis is the percentage of funds shocked. The grey area, delimited by the tick black line, denotes the (vulnerability) region on the plane where the threshold τ is reached or surpassed. This region is computed under the assumption of peer effects using the impact matrix (1.49). The dashed red line is, instead, the the boundary of the vulnerability region assuming no peer effects.

Table 1.5: This Table contains the robustness results for model (1.2) with respect to the definition of the spatial dependence between the funds. Panel A contains the estimated parameters in four cases where the spatial dependence id defined using the *Weighted Maturity Similarity* and four different methodologies for the construction of the similarity network. Instead, the four cases considered in Panel B use the same network construction methodology as in the main empirical results of Section 1.6 but four different similarity measures. Standard errors are clustered with respect to time and fund. Significance Levels: 0.01 '***', 0.05 '**', 0.1 '*'.

Panel A: Different network				
	$Flow_t$			
	Case 1	Case 2	Case 3	Case 4
$W_t \times Flow_t$	0.8077***	0.7402***	0.8952***	0.6639***
	(0.2599)	(0.1987)	(0.2505)	(0.1800)
$Flow_{t-1}$	−0.1272***	−0.1271***	−0.1291***	−0.1271***
	(0.0184)	(0.0181)	(0.0188)	(0.0186)
$W_{t-1} \times Flow_{t-1}$	0.2112***	0.1264***	0.1703***	0.1314***
	(0.0504)	(0.0248)	(0.0285)	(0.0377)
$Yield_{t-1}$	0.3193**	0.3820**	0.3836**	0.3791**
	(0.1590)	(0.1700)	(0.1914)	(0.1701)
$W_{t-1} \times Yield_{t-1}$	0.3936	0.1385	0.0980	0.1172
	(0.2557)	(0.2203)	(0.2315)	(0.2597)
Controls:	Yes	Yes	Yes	Yes
Additional controls:	No	No	No	No
Spatial lag IV:	4	4	4	4
Fixed Effects:	Fund, Time	Fund, Time	Fund, Time	Fund, Time
N. Obs.	17,442	17,442	17,442	17,442
Panel B: Different similarity measure				
	$Flow_t$			
	Case 5	Case 6	Case 7	Case 8
$W_t \times Flow_t$	0.9451***	0.7441***	0.6317**	0.9305***
	(0.2477)	(0.2353)	(0.3049)	(0.2965)
$Flow_{t-1}$	−0.1293***	−0.1266***	−0.1272***	−0.1304***
	(0.0181)	(0.0185)	(0.0185)	(0.0177)
$W_{t-1} \times Flow_{t-1}$	0.1289***	0.1260***	0.1857***	0.1587***
	(0.0397)	(0.0344)	(0.0386)	(0.0244)
$Yield_{t-1}$	0.1615	0.2891*	0.3588**	0.4548***
	(0.1499)	(0.1601)	(0.1633)	(0.1592)
$W_{t-1} \times Yield_{t-1}$	0.4750**	0.2940	0.2382	−0.0464
	(0.2290)	(0.2328)	(0.2465)	(0.2081)
Controls:	Yes	Yes	Yes	Yes
Additional controls:	No	No	No	No
Spatial lag IV:	2	2	2	2
Fixed Effects:	Fund, Time	Fund, Time	Fund, Time	Fund, Time
N. Obs.	17,442	17,442	17,442	17,442

Table 1.6: This Table contains the estimated impact measures and their first 5 lags partitioning for the four robustness cases presented in Panel A of Table 1.5. Significance Levels: 0.01 '***', 0.05 '**', 0.1 '*'.

Case 1						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000***		−0.1272***		0.3193*	
Lag 1		0.8076***		0.1083**		0.6514***
Lag 2	0.0087**	0.6435**	0.0010**	0.0865**	0.0068*	0.5193**
Lag 3	0.0049	0.5219	0.0006*	0.0700*	0.0039*	0.4210
Lag 4	0.00344	0.4221	0.0004	0.0566	0.0027	0.3405
Total impacts	1.0275***	4.1723	−0.1238***	0.5601**	0.3410**	3.3658

Case 2						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000***		−0.1270***		0.3819**	
Lag 1		0.7401***		0.0322		0.4212**
Lag 2	0.0182**	0.5296**	0.0004	0.0234	0.0100*	0.3017*
Lag 3	0.0080	0.3974	0.0003	0.0174	0.0044	0.2263
Lag 4	0.0051	0.2950	0.0001	0.0129	0.0028	0.1679
Total impacts	1.0424***	2.8063	−0.1258***	0.1229	0.4055**	1.5977

Case 3						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000***		−0.1291***		0.3835**	
Lag 1		0.8951***		0.0547**		0.4413***
Lag 2	0.0106**	0.7906**	0.0006***	0.0483***	0.0052**	0.3898**
Lag 3	0.0093*	0.7079*	0.0005***	0.0483***	0.0046*	0.3490*
Lag 4	0.0083	0.6337	0.0005***	0.0387***	0.0041	0.3124
Total impacts	1.1001***	8.4380	−0.1230***	0.5158***	0.4329***	4.1605

Case 4						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000***		−0.1270***		0.3791**	
Lag 1		0.66391***		0.0470		0.3689*
Lag 2	0.0142**	0.4265**	0.0008	0.0303	0.0077	0.2371
Lag 3	0.0030	0.2896	0.0002	0.0205	0.0016	0.1609
Lag 4	0.0027	0.1915	0.0001	0.0135	0.0015	0.1064
Total impacts	1.0228***	1.95285	−0.1256***	0.1385	0.3916**	1.0852

Table 1.7: This Table contains the estimated impact measures and their first 5 lags partitioning for the four robustness cases presented in Panel B of Table 1.5. Significance Levels: 0.01 '***', 0.05 '**', 0.1 '*'.

Case 5						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1292^{***}		0.1614	
Lag 1		0.9450^{***}		0.0067		0.6275^{***}
Lag 2	0.0169^{***}	0.8762^{***}	$-9.2 \cdot 10^{-5}$	0.0064	0.0104^{***}	0.5826^{***}
Lag 3	0.0100^{**}	0.8340^{**}	$2.3 \cdot 10^{-6}$	0.0060	0.0064^{**}	0.5541^{**}
Lag 4	0.0080^{*}	0.7896^{*}	$1.4 \cdot 10^{-5}$	0.0056	0.0052^{*}	0.5245^{*}
Total impacts	1.1195^{***}	17.090^{*}	-0.1289^{***}	0.1223	0.2390	11.3505^{*}
Case 6						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1266^{***}		0.2890^{**}	
Lag 1		0.7440^{***}		0.0318		0.5091^{**}
Lag 2	0.0109^{**}	0.5427^{**}	0.0003	0.0233	0.0071^{*}	0.3716^{*}
Lag 3	0.0109	0.4066	0.0002	0.0174	0.0035	0.2783
Lag 4	0.0109	0.3031	0.0001	0.0129	0.0022	0.2074
Total impacts	1.0273^{***}	2.8799	-0.1256^{***}	0.1233	0.3072^{**}	1.9711
Case 7						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1271^{***}		0.3587^{**}	
Lag 1		0.6317^{**}		0.1053^{**}		0.4648^{**}
Lag 2	0.0073	0.3917	0.0010^{*}	0.0655^{*}	0.0052	0.2884
Lag 3	0.0030	0.2490	0.0005	0.0415	0.0022	0.1832
Lag 4	0.0016	0.1575	0.0003	0.0263	0.0012	0.1159
Total impacts	1.0144^{***}	1.7009	-0.1249^{***}	0.2839^{***}	0.3690^{***}	1.2518
Case 8						
	ϵ_t		$Flow_{t-1}$		$Yield_{t-1}$	
	ADI	*AII*	*ADI*	*AII*	*ADI*	*AII*
Lag 0	1.000^{***}		-0.1304^{***}		0.4548^{***}	
Lag 1		0.9304^{***}		0.0372^{*}		0.3768
Lag 2	0.0195^{**}	0.8462^{**}	0.0004	0.0343^{**}	0.0080	0.3426
Lag 3	0.0130^{*}	0.7924^{*}	0.0003	0.0319^{**}	0.0053	0.3209
Lag 4	0.0105	0.7390	0.0003	0.0297^{*}	0.0042	0.2992
Total impacts	1.1448^{***}	13.2368	-0.1256^{***}	0.5313^{**}	0.5138^{**}	5.3609

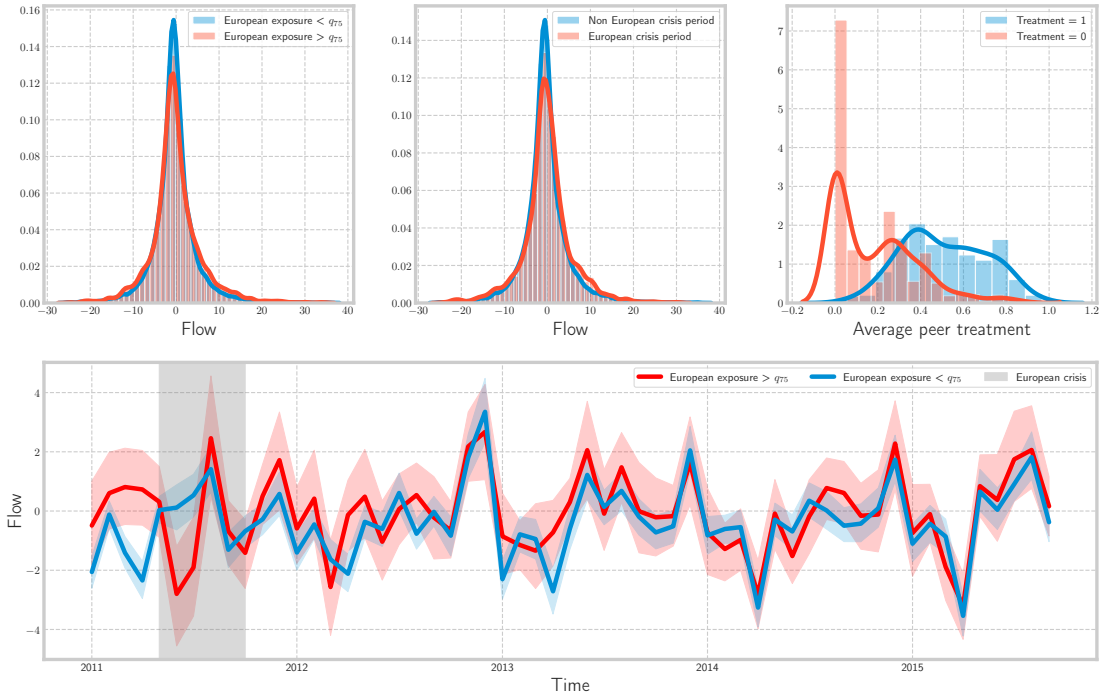


Figure 1.10: This Figure is a graphical analysis of the quasi-random experiment described in Section 1.B.1. The first plot in the top row shows the flow distributions conditional on the value of the dummy variable for european exposure. The second plot shows the flow distributions conditional on the value of the dummy variable for the crisis period. Instead, the last plot in the top row shows the distribution of the average peer treatment status for the two treatment groups. Finally, the bottom plot shows the average flow over time for the two treatment groups.

Table 1.8: This Table contains the estimation results for the quasi-random experiment expressed by model (1.128). The reported standard errors are the HAC standard errors proposed by Driscoll and Kraay [1998] and the significance levels are as follows: 0.01 '***', 0.05 '**', 0.1 '*'.

	$Flow_t$			
$W_t \times Flow_t^{Treatment}$	0.8752^{***} (0.0992)	0.8018^{***} (0.0914)	0.8301^{***} (0.0899)	0.8025^{***} (0.1112)
$Flow_{t-1}$	-0.1221^{***} (0.0222)	-0.1171^{***} (0.0199)	-0.1175^{***} (0.0199)	-0.1175^{***} (0.0199)
$W_{t-1} \times Flow_{t-1}$	0.1421^{***} 0.0477)	0.1628^{***} (0.0451)	0.1610^{***} (0.0441)	0.1631^{***} (0.0445)
$Yield_{t-1}$	-0.1508^{**} (0.0685)	0.2988^{*} (0.1564)	0.3485^{**} (0.1423)	0.3371^{**} (0.1453)
$W_{t-1} \times Yield_{t-1}$	0.1731 (0.1967)	0.2532 (0.2012)	0.2818 (0.1979)	0.2678 (0.2039)
Controls:	Yes	Yes	Yes	Yes
Additional Controls:	No	No	European exposure	European exposure, Peer European exposure
Fixed Effects:	Time	Fund, Time	Fund, Time	Fund, Time
N. Obs.	17,442	17,442	17,442	17,442

Table 1.9: This Table contains the results of the counterfactual experiment described in Section 1.C. Using the data generating process (1.131), where the spatial dependence parameter $\lambda = 0.5$ and the adjacency matrix G is random draw from the Erdös-Rényi family of random graphs $G(n, p)$, we perform 1000 simulations for each value of the linking probability p and estimate the peer effect model (1.2) using the similarity temporal network defined in Section 1.5. Here we report the average estimated spatial parameter, the average standard error and the percentage of the simulations where the p-value of $\overline{\lambda}$ is less than 1%.

	$\overline{\lambda}$	$\overline{S.E.}$	% p-value < 0.01
$p = 0$	0.4746	0.4514	4.4%
$p = 0.1$	0.4980	0.4540	3.1%
$p = 0.3$	0.4382	0.4466	2.4%
$p = 0.5$	0.4291	0.4529	2.7%

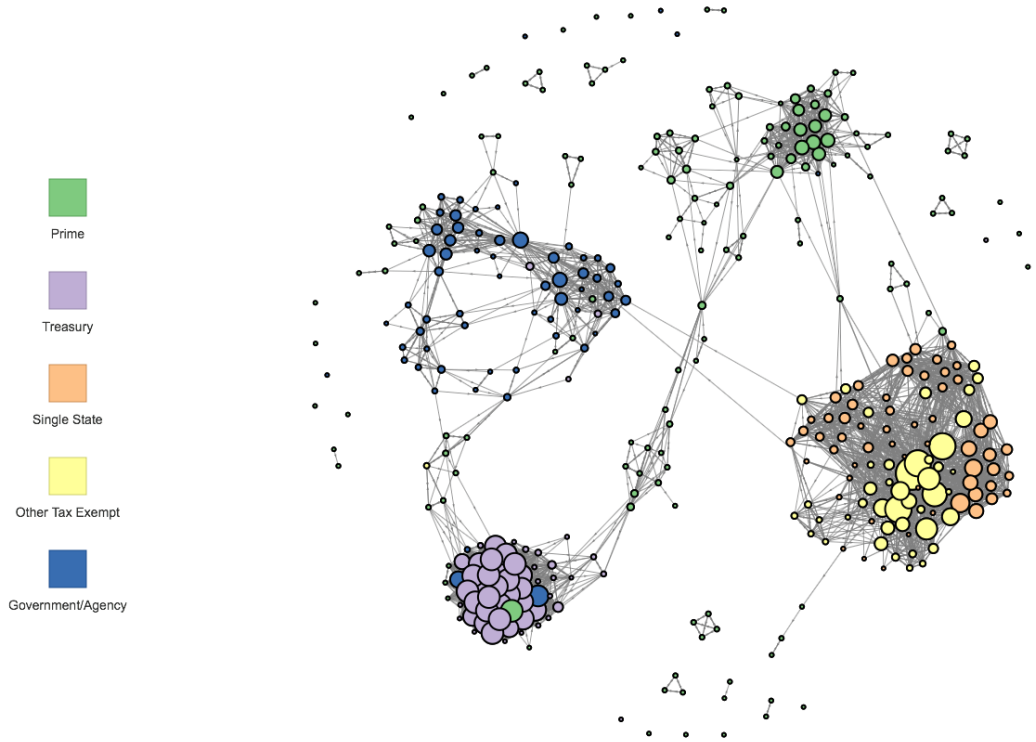


Figure 1.11: Aggregate representation of the similarity temporal network where only the links that are present at least 70% of the time are shown.

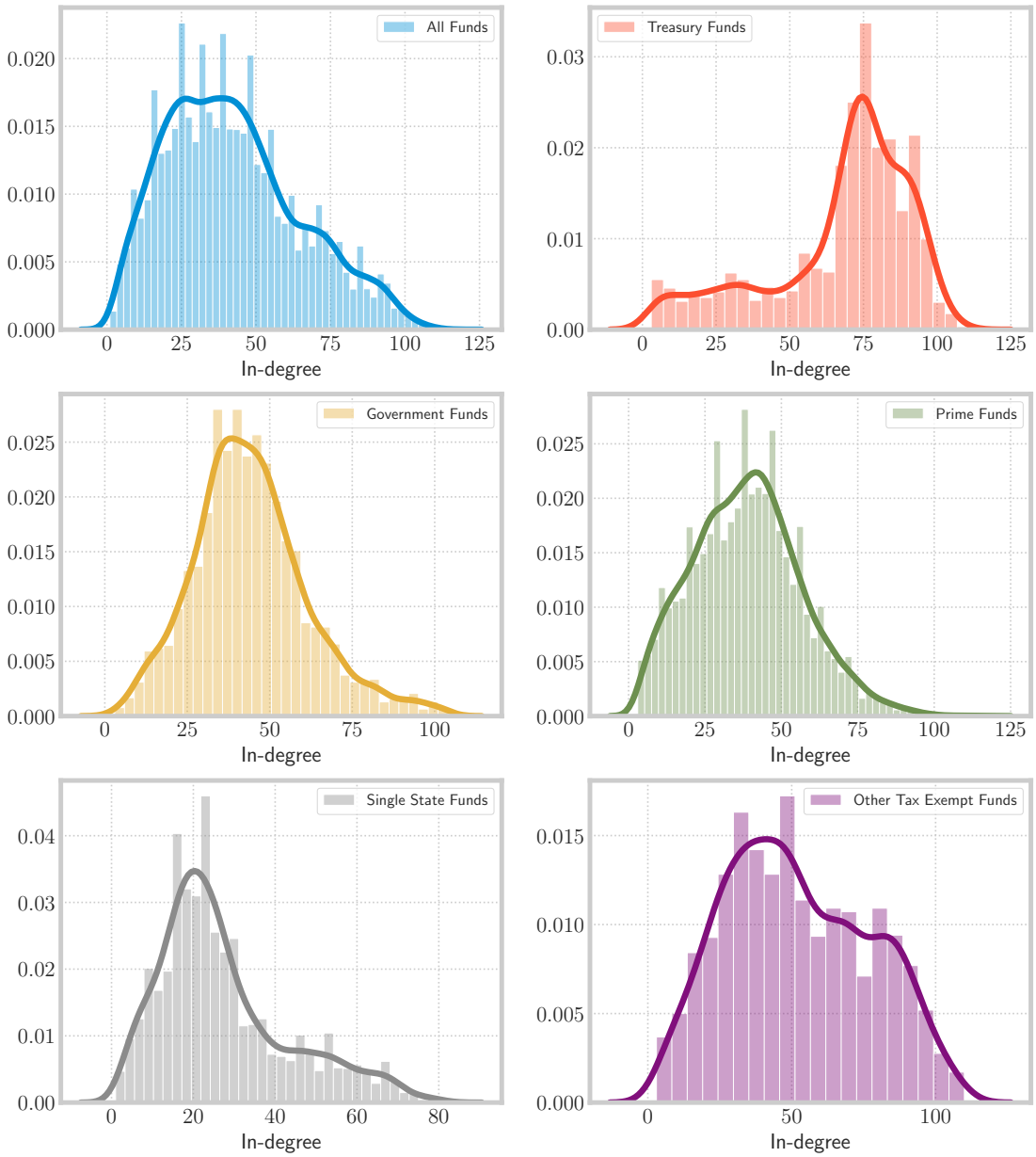


Figure 1.12: In-degree distribution for the aggregate of all funds (top left panel) and for each fund category (from top right panel to bottom right).

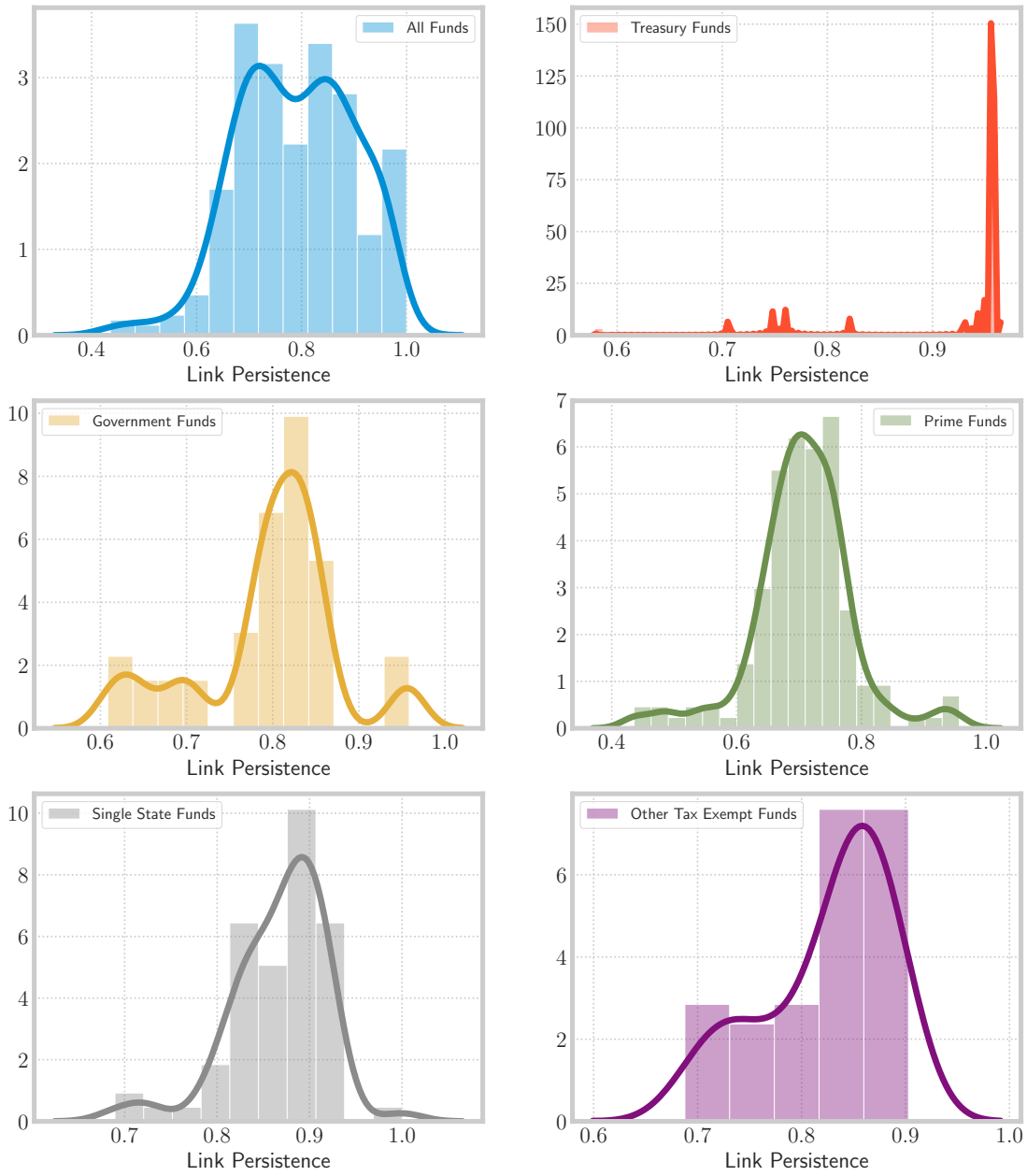


Figure 1.13: Distribution of the link persistence measure defined in (1.134) for the aggregate of all funds (top left panel) and for each fund category (from top right panel to bottom right).

Part II

Longevity risk modeling

Chapter 2

Annuity portfolios: risk margin, longevity and interest rate hedging[1]

2.1 Introduction

Assessing and hedging the longevity risk, i.e. the risk of unexpected changes in the survivorship of individuals, has become increasingly important for insurance companies and pension funds in the last decades. The unprecedented and unforeseen increase in the file expectancy of individuals and, in particular, of policyholders, annuitants and pension fund members, is posing serious threats in terms of solvency to insurance portfolios and pension funds. The need to manage longevity risk generated the potential for the creation of a new market. Longevity bonds were first proposed by the academic literature [Blake and Burrows, 2001] at the beginning of the 2000s, as instruments to hedge against the fluctuations of mortality likelihoods in a portfolio of policies. Interestingly, the first market transactions followed few years later. However, the first attempts to transfer longevity risk to capital markets were unsuccessful, mainly because of the

[1]The material for this chapter is taken from De Rosa, Luciano, and Regis [2016].

lack of standardization and high charges [Cairns et al., 2008]. It has been only in the last few years that the volume of longevity transfers boomed. Before this peak, many researchers studied the optimal design of these transfers [Dowd et al., 2006] and their effectiveness [Ngai and Sherris, 2011], paving the way to the actual start of the market. As of today, the number and the size of the transactions concerning the transfer of longevity risk exposures from insurers and pension plans to reinsurance or specialist is growing fast. Trades completed in 2015 only accounted for nearly 40 billion euros of notional. More and more often, longevity transfers take the form of longevity swaps, that are standardized, indexed-based solutions. These deals call for a cheaper coverage for longevity sellers on the one hand, but forces them to deal with basis risk on the other. Basis risk arises because of the non-customized nature of the transfer. Technically, it is caused by the difference between the dynamics of the mortality pattern of the seller's portfolio and the dynamics of the reference population that serves as an index in the transaction.

We propose dynamic hedging strategies for annuity portfolios, that use standardized longevity contracts, and we evaluate their effectiveness. We consider the impact of basis risk and of the strategy rebalancing frequency, as in De Rosa et al. [2017], but we account also for the presence of interest rate risk, due to the stochastic nature of the short rate. The inclusion of this source of risk in our analysis is relevant, for many reasons. First, the need to evaluate insurance liabilities at fair value requires careful consideration of the impact of interest rate risk in the management of the whole risk profile of annuity providers. We highlight that the mathematical reserve, which is the quantity that needs to be hedged, is affected by the stochastic nature of the discount factors, as well as by the randomness in mortality rates. Second, considering uncertain interest rates allows us to clarify that the management of the assets and liabilities of an insurer or annuity provider is intertwined, because the exposure to the same source of risk (interest rate variability) affects both their investments and their obligations. Third, our work sheds light on the importance of assessing how coverage of longevity through indexed instruments, such as synthetic longevity bonds, may vary the exposure to interest rate risk.

Previous actuarial literature mostly focused on longevity risk hedging strategies, when interest rates are deterministic. Only a few papers stressed the importance of jointly considering longevity and interest rate risk coverage when interest rates are uncertain. Stevens et al. [2011] evaluated the effect of investment risk on the effectiveness of natural hedging of mortality and financial risk. Tsai et al. [2011] considered a static annuity hedging strategy to cover the aggregate risk deriving from both mortality and interest rate stochasticity when the short rate follows a [Cox et al., 1985a] model. They find that hedging with a longevity bond improves the effectiveness of the hedge, which is highly sensitive to the parameters of the interest rate process. We extend their analysis, by considering longevity basis risk as well. We model interest rates following the well-known Vasicek (1977) mode, that allows us to include the possibility of negative rates. We consider Greek-based dynamic hedges as in Luciano et al. [2012b] and Cairns [2013], for instance.

To evaluate the effectiveness of our strategies, we track the hedging errors and we analyze their moments. Moreover, in order to compare the relative impact of the different risk sources (longevity risk and interest rate risk), we compute the risk margins of different hedging strategies when longevity basis risk is present and interest rate is stochastic. The risk margin is defined as the 99.5% Value-at-Risk of the hedged portfolio at a future time horizon $t = 30$ years. We consider such a long-time horizon, having in mind that the liabilities of an insurer can be very volatile in very distant future. We evaluate different longevity hedging strategies when the interest rate risk is fully hedged, partially hedged or left completely unhedged.

Our calibrated example, based on UK data, shows that the impact of interest rate risk on the moment of the hedging error of the strategies is substantial. For instance, the standard deviation of the hedging error of the Delta-Gamma strategy doubles when interest rate risk is present and left unhedged. However, if we measure the effectiveness of each strategy in terms of the required risk margin, the picture changes significantly. The risk margins required for an unhedged annuity contract are comparable both with or without interest rate risk. Interestingly, we find that, when interest rate risk is present,

it accounts only for 10% of the total risk margin. The remaining 90% of the risk margin is, instead, required only by the uncertainty about the future mortality rates.

2.2 Longevity and interest rate risk modeling

We consider a standard filtered probability space $(\Omega, \mathcal{H}, \mathbb{Q})$ satisfying the usual assumptions and a filtration on this space, $\mathcal{H}_t$, that collects all the information available from the financial market we will describe below and all the information concerning the longevity dynamics relevant for the insurance products entering the portfolio or hedging strategies. The measure $\mathbb{Q}$ is the so-called risk-neutral measure. In order to keep the notation simple, we assume that there is no risk premium in the longevity market or, equivalently, that the dynamics of the mortality intensities we will describe hereafter are the same under such measure and the historical one. The interest rate dynamics under the risk-neutral measure will instead incorporate a risk premium, that we will estimate.

2.2.1 Longevity risk model

We assume that the time to death τ of an individual is modeled by a doubly stochastic process, namely a Poisson process with stochastic intensity. Such mortality intensity is homogeneous for individuals belonging to the same cohort and population. The dynamics of the mortality of a given cohort x in a population is described by a square-root process of the type:

$$d\lambda_x(t) = (a_x + b_x\lambda_x(t))dt + \sigma_x\sqrt{\lambda_x}dW_x(t), \tag{2.1}$$

with $a_x > 0, b_x > 0, \sigma_x > 0, \lambda_x(0) = \lambda_0 \in \mathbb{R}^{++}$. The assumption $b_x > 0$ ensures that the process has no mean reversion. The individual ages over time, and the drift simply ensures that the expected change in the intensity is affine and increasing in the intensity itself. This model, in particular, is akin to the well-known Gompertz mortality law in the traditional actuarial literature, because the mortality intensity of an individual is expected to increase exponentially with age.

If the initial point λ_0 is strictly positive, we can impose a restriction on the coefficients

that guarantees that the mortality intensity $\lambda_x(t)$ is strictly positive for every t, almost surely. This condition reads:

$$a_x \geq \frac{\sigma_x^2}{2}. \tag{2.2}$$

It is convenient, for many purposes, especially when considering hedging solutions, having a model that describes the joint mortality dynamics of two populations: the population of insureds of a company, *Portfolio population*, and the population underlying the payoffs of longevity hedging instruments *Reference population*. When the two populations differ, *basis risk* needs to be accounted for when hedging instruments are used by the company. Hence, we assume that the intensity of cohort x in the Reference population follows SDE (2.1).

Basis risk

We assume that the individuals in the insurance portfolio represent a subsample of this Reference population. The mortality intensity of cohort x belonging to the Portfolio population is

$$\lambda_x^{pp} = \delta_x \lambda_x(t) + (1 - \delta_x)\lambda'_x(t), \tag{2.3}$$

with

$$d\lambda'_x(t) = (a' + b'\lambda'_x(t))dt + \sigma'\sqrt{\lambda'_x(t)}dW'_x(t), \tag{2.4}$$

where W_x and $W_x^{'}$ are two independent standard Brownian Motions, $a' > 0$, $\sigma' > 0$, $b' \in \mathbb{R}$ and $0 \leq \delta_x \leq 1$.

The intensity of the insurer's Portfolio population λ_x^{pp} is a convex combination of the Reference population's intensity λ_x and an idiosyncratic component $\lambda_x^{'}$ orthogonal to λ_x. As a consequence, applying Ito's Lemma, it is easy to show that the dynamics of λ_x^{pp} follow a two-factor CIR process. The idiosyncratic component $\lambda_x^{'}$ is specific to the Portfolio population and cannot be hedged using instruments written on the reference population. The parameter δ_x measures the dependence between the two mortality intensities and $1 - \delta_x$ can be easily taken as a measure of basis risk.

It is important to justify properly our interpretation of δ_x as a measure of comovement between the intensity of the two populations. Assuming $0 \leq u \leq t$, in particular, the conditional correlation between $\lambda_x(t)$ and $\lambda_x^{pp}(t)$ is

$$Corr_u\big[\lambda_x^{pp}(t), \lambda_x(t)\big] = \delta_x \sqrt{\frac{Var_u(\lambda_x(t))}{Var_u(\lambda_x^{pp}(t))}}, \tag{2.5}$$

where

$$Var_u\big[\lambda_x(t)\big] = \frac{a\sigma^2}{2b^2}\big(e^{b(t-u)} - 1\big)^2 + \frac{\sigma^2}{b}e^{b(t-u)}\big(e^{b(t-u)} - 1\big)\lambda_x(u), \tag{2.6}$$

$$Var_u\big[\lambda_x^{'}(t)\big] = \frac{a^{'}(\sigma^{'})^2}{2(b^{'})^2}\big(e^{b^{'}(t-u)} - 1\big)^2 + \frac{(\sigma^{'})^2}{b^{'}}e^{b^{'}(t-u)}\big(e^{b^{'}(t-u)} - 1\big)\lambda_x^{'}(u), \tag{2.7}$$

$$Var_u\big[\lambda_x^{pp}(t)\big] = \delta_x^2 Var_u\big[\lambda_x(t)\big] + (1-\delta_x)^2 Var_u\big[\lambda_x^{'}(t)\big]. \tag{2.8}$$

Indeed, when $\delta_x = 0$ the two intensities have zero correlation, while $\delta_x = 1$ implies perfect positive correlation. Having δ_x positive ensures that λ^{pp} is strictly positive. $Corr_u\big[\lambda_x^{pp}(t), \lambda_x(t)\big]$ stays between 0 and 1. Though this may seem restrictive, this assumption is justified by the intuition that when a shock hits the Reference population, increasing for example its mortality intensity, the sub-population is affected similarly, but with a different sensitivity, while divergence between the two intensities is entirely captured by the idiosyncratic risk factor $\lambda^{'}$.

Survival probabilities

In this framework, it is easy to derive the conditional survival probability from t to T is the probability that the individual will be alive at T, given that he or she is alive at t. Indeed,

$$S(t,T) = \mathbb{Q}\left(\tau \geqslant T \mid \tau > t\right) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \lambda_x(s)ds\right) \mid \mathcal{F}_t\right]. \tag{2.9}$$

Given our model 2.1, the expectation for the Reference Population becomes:

$$S(t,T) = A(t,T)e^{-B(t,T)\lambda(t)}, \tag{2.10}$$

where $A(t,T)$ and $B(t,T)$ solve an appropriate system of Riccati equations, being

$$A(t,T) = A(t,T;a,b,\sigma) = \left(\frac{2\gamma e^{\frac{1}{2}(\gamma-b)(T-t)}}{(\gamma-b)\left(e^{\gamma(T-t)}-1\right)+2\gamma}\right)^{\frac{2a}{\sigma^2}}, \tag{2.11}$$

$$B(t,T) = B(t,T;a,b,\sigma) = \frac{2\left(e^{\gamma(T-t)}-1\right)}{(\gamma-b)\left(e^{\gamma(T-t)}-1\right)+2\gamma}, \tag{2.12}$$

where $\gamma = \sqrt{b^2+2\sigma^2}$. As shown in Fung et al. [2014], the above specification guarantees also that the limit of the survival probability, when T diverges, is zero.
The survival probabilities of the Portfolio population can be written as functions of the common and idiosyncratic intensities as follows:

$$S^{pp}(t,T) = \tilde{S}(t,T)\tilde{S}'(t,T) \tag{2.13}$$

$$= \tilde{A}(t,T)\tilde{A}'(t,T)e^{-\tilde{B}(t,T)\delta_x\lambda_x(t)-\tilde{B}'(t,T)(1-\delta_x)\lambda'_x(t)}, \tag{2.14}$$

where

$$\begin{aligned}
\tilde{A}(t,T) &= A(t,T;\alpha,\beta,\eta),\\
\tilde{B}(t,T) &= B(t,T;\alpha,\beta,\eta),\\
\tilde{A}'(t,T) &= A(t,T;\alpha',\beta',\eta'),\\
\tilde{B}'(t,T) &= B(t,T;\alpha',\beta',\eta'),\\
\tilde{\gamma} &= \sqrt{\beta^2+2\eta^2},\\
\tilde{\gamma}' &= \sqrt{(\beta')^2+2(\eta')^2},
\end{aligned}$$

and

$$\begin{aligned}
\alpha &= \delta_x a, \\
\beta &= b, \\
\eta^2 &= \delta_x \sigma^2, \\
\alpha' &= (1-\delta_x) a', \\
\beta' &= b', \\
(\eta')^2 &= (1-\delta_x)(\sigma')^2.
\end{aligned}$$

2.2.2 Interest rate model

We assume that the spot interest rate - or interest rate intensity - follows, under the risk-neutral measure, the well-known Vasicek [1977] process, of the type:

$$dr(t) = \bar{a}(\bar{b} - r(t))dt + \bar{\sigma} dW_r(t), \tag{2.15}$$

with $\bar{a} > 0, \bar{b} > 0, \bar{\sigma} > 0, r(0) = r_0 \in \mathbb{R}^{++}$, where the Wiener process W_r is independent of W_x and W'_x: longevity and interest rate risks are independent. Given the above assumptions, the process is mean reverting to a long-run value, $\bar{b}$, with speed $\bar{a}$. This short rate can display negative paths with positive probability. The discount factor or bond price at time t, for maturity T, associated to our process 2.15, is

$$D(t,T) = \mathbb{E}\left[\exp\left(-\int_t^T r(u)du\right)|\mathcal{F}_t\right] = \bar{A}(t,T)e^{-\bar{B}(t,T)r(t)},$$

where $\bar{A}(t,T)$ and $\bar{B}(t,T)$ are solutions to the Riccati ODEs,

$$\bar{A}(t,T) = exp\left[\left(\bar{b} - \frac{\bar{\sigma}^2}{2\bar{a}^2}\right)\left(\bar{B}(t,T) - T + t\right) - \frac{\bar{\sigma}^2}{4a}\bar{B}(t,T)^2\right], \tag{2.16}$$

$$\bar{B}(t,T) = \frac{1 - e^{-\bar{a}(T-t)}}{\bar{a}}. \tag{2.17}$$

2.2.3 Risk factors

Following a technique described in Jarrow and Turnbull [1994] and Luciano et al. [2012a], which exploits the definitions of "forward" interest rate and mortality intensity respectively, we obtain a rewriting of the zero-coupon bond prices and survival probabilities in terms of risk factors that are very easy to interpret:

$$S(t,T) = e^{-X(t,T)I(t)+Y(t,T)}, \tag{2.18}$$

where

$$\begin{aligned} X(t,T) &= B(t,T), \\ Y(t,T) &= lnA(t,T) - B(t,T)\left[-\frac{\partial lnA(t,T)}{\partial T}\Big|_{(0,t)} + \lambda_x(0)\frac{\partial B(t,T)}{\partial T}\Big|_{(0,t)}\right], \end{aligned}$$

and

$$D(t,T) = e^{-\bar{X}(t,T)J(t)+\bar{Y}(t,T)}, \tag{2.19}$$

where

$$\begin{aligned} \bar{X}(t,T) &= \bar{B}(t,T), \\ \bar{Y}(t,T) &= ln\bar{A}(t,T) - \bar{B}(t,T)\left[-\frac{\partial ln\bar{A}(t,T)}{\partial T}\Big|_{(0,t)} + r(0)\frac{\partial \bar{B}(t,T)}{\partial T}\Big|_{(0,t)}\right], \end{aligned}$$

and where the two risk factors $I(t)$ and $J(t)$ are defined as

$$\begin{aligned} I(t) &= \lambda_x(t) - f(0,t), \\ J(t) &= r(t) - F(0,t). \end{aligned}$$

Here, $f(0,t)$ denotes the forward mortality intensity and $F(0,t)$ the forward interest rate. These risk factors are the difference between the actual realizations of the intensity and short rate at time t and their best forecasts at time 0. They can intuitively be interpreted as forecast errors and they will be the quantities whose exposures we will

hedge. In our setting, it is very important to remark that even the survival probabilities of the Portfolio population can be rewritten as functions of the longevity risk factor $I(t)$, since

$$S^{pp}(t,T) = e^{-X^{pp}(t,T)\delta_x I(t) - X^{'}(t,T)(1-\delta_x)\lambda^{'}_x(t) + Y^{pp}(t,T)},$$

where

$$X^{pp}(t,T) = \tilde{B}(t,T), \tag{2.20}$$

$$X^{'}(t,T) = \tilde{B}^{'}(t,T), \tag{2.21}$$

$$Y^{pp}(t,T) = ln\tilde{A}(t,T) + ln\tilde{A}^{'}(t,T) - \tilde{B}(t,T)f_x(0,t). \tag{2.22}$$

2.3 Hedging an annuity portfolio

2.3.1 Annuity fair value and sensitivities to the risk factors

Our goal is to show how to construct hedging strategies for an annuity contract, that, abstracting from idiosyncratic risk and policyholders' heterogeneity, represents the portfolio of an annuity provider. The annuity has maturity T and annual installments R paid at year-end – written on an individual aged x at time 0. This individual belongs to the portfolio population whose mortality we previously defined and that follows the mortality intensity described by equation (2.3). The fair-value of the annuity is the quantity the annuity provider needs to hedge. This fair value, that is equal to the mathematical reserve of the policy, is computed as

$$N^{pp}(t,T) = R\sum_{u=1}^{T-t} D(t,t+u)S^{pp}(t,t+u), \tag{2.23}$$

that can be equivalently written as

$$N^{pp}(t,T) = \tag{2.24}$$

$$= R\sum_{u=1}^{T-t} e^{-\bar{X}(t,t+u)J(t)+\bar{Y}(t,t+u)} \cdot e^{-X^{pp}(t,t+u)\delta_x I(t) - X^{'}(t,t+u)(1-\delta_x)\lambda^{'}_x(t) + Y^{pp}(t,t+u)}. \tag{2.25}$$

This value is sensitive to changes in the interest rates and in the dynamics of the mortality of the portfolio population. The marginal effect on the value of the reserve caused by any unexpected change in the risk factors can approximated as:

$$dN^{pp} = \frac{\partial N^{pp}}{\partial t}dt + \frac{\partial N^{pp}}{\partial I}dI + \frac{1}{2}\frac{\partial^2 N^{pp}}{\partial I^2}(dI)^2 + \frac{\partial N^{pp}}{\partial \lambda'}d\lambda' + \frac{1}{2}\frac{\partial^2 N^{pp}}{\partial (\lambda')^2}(d\lambda')^2 + \\ + \frac{\partial N^{pp}}{\partial J}dJ + \frac{1}{2}\frac{\partial^2 N^{pp}}{\partial J^2}(dJ)^2, \tag{2.26}$$

where

$$\begin{aligned}
\frac{\partial N^{pp}}{\partial I} &= R\sum_{u=1}^{T-t} D(t,t+u)\Delta^M_{pp}(t,t+u),\\
\frac{\partial^2 N^{pp}}{\partial I^2} &= R\sum_{u=1}^{T-t} D(t,t+u)\Gamma^M_{pp}(t,t+u),\\
\frac{\partial N^{pp}}{\partial \lambda'} &= R\sum_{u=1}^{T-t} D(t,t+u)\Delta'_{pp}(t,t+u),\\
\frac{\partial^2 N^{pp}}{\partial (\lambda')^2} &= R\sum_{u=1}^{T-t} D(t,t+u)\Gamma'_{pp}(t,t+u),\\
\frac{\partial N^{pp}}{\partial J} &= R\sum_{u=1}^{T-t} \Delta^F(t,t+u)S^{pp}(t,t+u),\\
\frac{\partial^2 N^{pp}}{\partial J^2} &= R\sum_{u=1}^{T-t} \Gamma^F(t,t+u)S^{pp}(t,t+u),
\end{aligned}$$

with

$$\Delta^M_{pp}(t,T) := \frac{\partial S^{pp}(t,T)}{\partial I} = -X^{pp}(t,T)\delta_x S^{pp}(t,T) \leq 0, \tag{2.27}$$

$$\Gamma^M_{pp}(t,T) := \frac{\partial^2 S^{pp}(t,T)}{\partial I^2} = \left(X^{pp}(t,T)\delta_x\right)^2 S^{pp}(t,T) \geq 0, \tag{2.28}$$

$$\Delta'_{pp}(t,T) := \frac{\partial S^{pp}(t,T)}{\partial \lambda'} = -X'(t,T)(1-\delta_x) S^{pp}(t,T) \leq 0, \tag{2.29}$$

$$\Gamma'_{pp}(t,T) := \frac{\partial^2 S^{pp}(t,T)}{\partial (\lambda')^2} = \left(X'(t,T)(1-\delta_x)\right)^2 S^{pp}(t,T) \geq 0, \tag{2.30}$$

$$\Delta^F(t,T) := \frac{\partial D(t,T)}{\partial J} = -\bar{X}(t,T) D(t,T) \leq 0, \tag{2.31}$$

$$\Gamma^F(t,T) := \frac{\partial^2 D(t,T)}{\partial J^2} = \bar{X}(t,T)^2 D(t,T) \geq 0. \tag{2.32}$$

Δ^M_{pp} and Δ^F are negative because, as one would expect, the value of the annuity is decreasing in both risk factors. The second order sensitivities are instead positive and the higher I or J, the higher is the sensitivity to the changes in the risk factors. Equation (2.27) highlights that the sensitivity of the annuity with respect to the longevity risk factor I is directly proportional to the parameter δ_x. This is intuitive, given the interpretation of δ_x as the degree of co-movement of the portfolio and of the reference population.

2.3.2 Dynamic hedging strategies using longevity bonds

In this subsection, we propose a dynamic hedging strategy that covers the fair value of the reserve of the annuity described above using longevity bonds. The longevity bonds we consider are instruments written on the survivorship of a Reference population, that follows the dynamics described by equation (2.1). The payoff at maturity T of such an instrument is

$$\exp\left(-\int_t^T \lambda_x(s)ds\right).$$

The value of the longevity bond, as a consequence, is

$$\begin{aligned} M(t) &= D(t,T)S(t,T), \\ &= e^{-\bar{X}(t,T)J(t)+\bar{Y}(t,T)} \cdot e^{-X(t,T)I(t)+Y(t,T)}, \end{aligned} \tag{2.33}$$

while its dynamics can be written as

$$dM = \frac{\partial M}{\partial t}dt + \frac{\partial M}{\partial I}dI + \frac{1}{2}\frac{\partial^2 M}{\partial I^2}(dI)^2 + \frac{\partial M}{\partial J}dJ + \frac{1}{2}\frac{\partial^2 M}{\partial J^2}(dJ)^2, \tag{2.34}$$

where

$$\begin{aligned} \frac{\partial M}{\partial I} &= D(t,T)\Delta^M(t,T), \\ \frac{\partial^2 M}{\partial I^2} &= D(t,T)\Gamma^M(t,T), \\ \frac{\partial M}{\partial J} &= \Delta^F(t,T)S(t,T), \\ \frac{\partial^2 M}{\partial J^2} &= \Gamma^F(t,T)S(t,T), \end{aligned}$$

and $\Delta^M(t,T) = -X(t,T)S(t,T), \Gamma^M(t,T) = X(t,T)^2 S(t,T)$. $\Delta^F(t,T), \Gamma^F(t,T)$ are given by (2.31), (2.32), respectively. A perfect hedge of longevity risk cannot be achieved, unless $\delta_x = 1$, even trading continuously in the hedging instrument. This happens because changes in the idiosyncratic component λ' influence N^{pp}, but do not affect M, as one can conclude by comparing (2.26) and (2.34). Indeed, the value of the hedging portfolio will not perfectly replicate the value of the insurance liabilities, i.e. the overall hedging error will differ from zero. Still, we can disentangle the exposures to the source of risk that can be covered, i.e. $I(t)$, from the unhedgeable risk, i.e. $\lambda_x^{'}(t)$. We consider hedging strategies based on the Greeks we computed above. At a given point in time t, we aim to build an hedging portfolio that nullifies the overall exposure to the risk factors. These strategies (see De Rosa et al. [2017]) are defined according to which Greeks are offset by taking appropriate positions in the hedging instruments. Delta (Gamma) strategies nullify first- (second-) order changes in the longevity and/or financial risk

factors, while Theta strategies nullify first-order time changes effects. Strategies can also be constructed to be self-financing. Computing the hedging strategy, at time t, amounts to solving an appropriate system of equations. For the sake of simplicity, we assume that there exists a number K of hedging instruments (longevity bonds written on generation x of the Reference population or zero coupon bonds, with different maturities $T_i, i = 1, ...K$, to which we refer as H_i), sufficient to guarantee a unique solution to the system of equations:

$$\begin{cases} -\frac{\partial N^{pp}(t)}{\partial I}dI + \sum_{i=1}^{K} n_i \frac{\partial H_i(t)}{\partial I}dI = 0, \\ -\frac{\partial^2 N^{pp}(t)}{\partial I^2}(dI)^2 + \sum_{i=1}^{K} n_i \frac{\partial^2 H_i(t)}{\partial I^2}(dI)^2 = 0, \\ -\frac{\partial N^{pp}(t)}{\partial J}dJ + \sum_{i=1}^{K} n_i \frac{\partial H_i(t)}{\partial J}dJ = 0, \\ -\frac{\partial^2 N^{pp}(t)}{\partial J^2}(dJ)^2 + \sum_{i=1}^{K} n_i \frac{\partial^2 H_i(t)}{\partial J^2}(dJ)^2 = 0, \\ -\frac{\partial N^{pp}(t)}{\partial t}dt + \sum_{i=1}^{K} n_i \frac{\partial H_i(t)}{\partial t}dt = 0, \\ -N^{pp}(t) + \sum_{i=1}^{K} n_i H_i(t) = 0. \end{cases} \tag{2.35}$$

When the system above is considered in its entirety, the strategy is the Delta-Gamma-Theta hedging strategy of longevity and interest rate risk, and $K \geq 6$ in order to ensure uniqueness of the solution. The first two equations nullify the Delta and Gamma exposure to the longevity risk factor, respectively, while the third and fourth nullify the Delta and Gamma exposure to the financial risk factor, respectively. The fifth equation nullifies the Theta, i.e. the deterministic change of the value of the portfolio in time. The sixth equation guarantees that the portfolio is self-financing. Considering only a part of the equations in (2.35) leads to different strategies, named after the Greeks that are nullified. The strategy is applied at any rebalancing date, using the same set of instruments. Right before the rebalancing date t, we evaluate the portfolio. This value

is the profit or loss of the hedge. We finance this quantity through a bank account, that accrues or charges a fixed rate r_0. From this bank account the payments related to the annuity contract are also taken. As usual, we define the hedging error of the dynamic strategy as the absolute value of the bank account. At any t, then, the system of equations is solved and the hedge, i.e. the positions in the hedging instruments, are computed.

2.4 Numerical analysis: Longevity and Interest Rate risk effects on Risk Margins

In this section, we consider an insurer who has sold an annuity contract to a Scottish male born in 1946, who was aged 64 on 31/12/2010 (i.e. $x = 65$). We compare the risk margins to be set apart in order to meet the obligations at a future date equal to $t = 30$ years with a confidence level of 99.5%, when both longevity and interest rate risk are present, and different dynamic hedging strategies are performed. We consider a realistic situation in which longevity risk cannot be fully hedged through dynamic hedging, due to the presence of Basis Risk. Moreover, we assume that the insurer computes risk margins consistently with the Solvency II regulation, that is, as the 99.5% confidence level Value-at-Risk of her portfolio at a future date. Despite the fact that solvency capital requirements are computed with a time horizon of 1 year, we use a time horizon 30 years, consistently with a long-term view of the risks of an annuity provider.

2.4.1 Calibration

The fit of the mortality model follows the procedure described by De Rosa et al. [2017]. We jointly calibrate the parameters of our Reference and Portfolio mortality models to the generations of UK and Scottish males born in 1946, who were aged 64 on 31/12/2010 (i.e. $x = 65$), using the data provided by the Human Mortality Database. Under the constraint given by condition (2.2), we fix 01/01/1991 as the observation point (individuals have all reached age 44) and we fit the observed survival probabilities $S(0,t), S^{pp}(0,t)$

Table 10: Reference and Portfolio population joint calibration results.

a	b	σ	δ_x	$a^{'}$	$b^{'}$	$\sigma^{'}$
$3.3357 \cdot 10^{-5}$	0.0727	0.0082	0.9897	0.0077	0.0155	$4.4463 \cdot 10^{-08}$

Table 11: Vasicek calibrated parameters.

$\bar{a}$	$\bar{b}$	$\bar{\sigma}$	$r(0)$
0.233821	0.030637	0.009400	0.007600

with t=1,...20. We fit our models minimizing the Rooted Mean Squared Error (RMSE) between the model-implied and the observed survival probabilities. The calibration error is 0.00015, and the values of the calibrated parameters are shown in Table 10.

Consistently, the interest rate model is calibrated to the UK Government Bond market[2]. In particular, the diffusion parameter σ_r of the Vasicek model is derived from a monthly time series of short rates (source: OECD) going from January 1978 to December 2010, using maximum likelihood estimation. The other parameters ($\bar{a}$ and $\bar{b}$) are then estimated fitting the model implied term structure to the observed one given UK government bonds at 31st of December 2010, by minimizing the rooted mean square error. Table 11 reports the values of the calibrated parameters. As one would expect from the Vasicek model, and to represent the current situation of interest rate markets, we allow for the possibility of negative interest rates. Figure 14 shows the simulated percentiles, from 0.01 to 0.99, of the short rate r_t.

2.4.2 Delta-Gamma hedging with Basis and Interest Rate Risk

In this section, we analyze the effectiveness of a Delta-Gamma hedging strategy for longevity risk, when interest rate is stochastic. We compare two different scenarios. In the first, the financial risk factor is left unhedged. In the second, it is Delta hedged. We compare the results with the ones described in De Rosa et al. [2017], where interest rate

[2]Parameters are taken from Jevtić and Regis [2015].

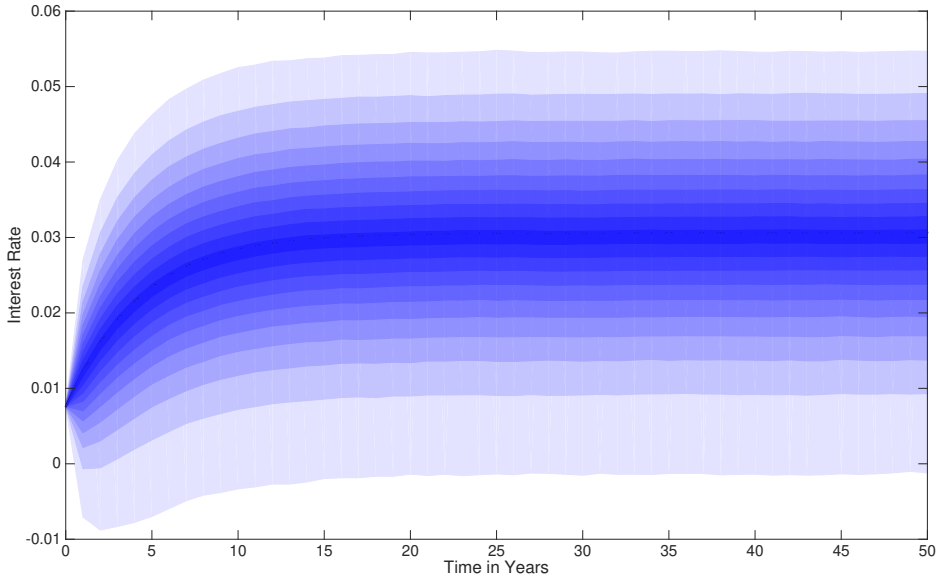


Figure 14: Percentiles of the Interest Rate from 1% to 99%.

risk was fully hedged, assuming a constant short rate $r_0 = 2\%$. We report this last case here not only for completeness, but also because it provides a benchmark to understand the effect of interest rate risk in longevity hedges.

Delta-Gamma hedging of longevity risk is performed at three different rebalancing frequencies (3, 6 and 12 months), using three longevity bonds with rolling maturities of 10, 15, 20 years. From system (2.35), we understand that three longevity bonds are needed to build a self-financing strategy. Being the strategy self-financing, we track the hedging error simply by storing, in a Bank Account, the difference between the value of the hedging portfolio and value of the annuity at each rebalancing date.

We perform Delta-hedging of interest rate risk as well, by adding a fourth longevity bond (with rolling matuirty of 5 years) to the hedging portfolio, rather than using a zero-coupon bond. By doing so, the composition of the hedging portfolio at each rebalancing date can be determined by solving the following system of 4 equations in 4

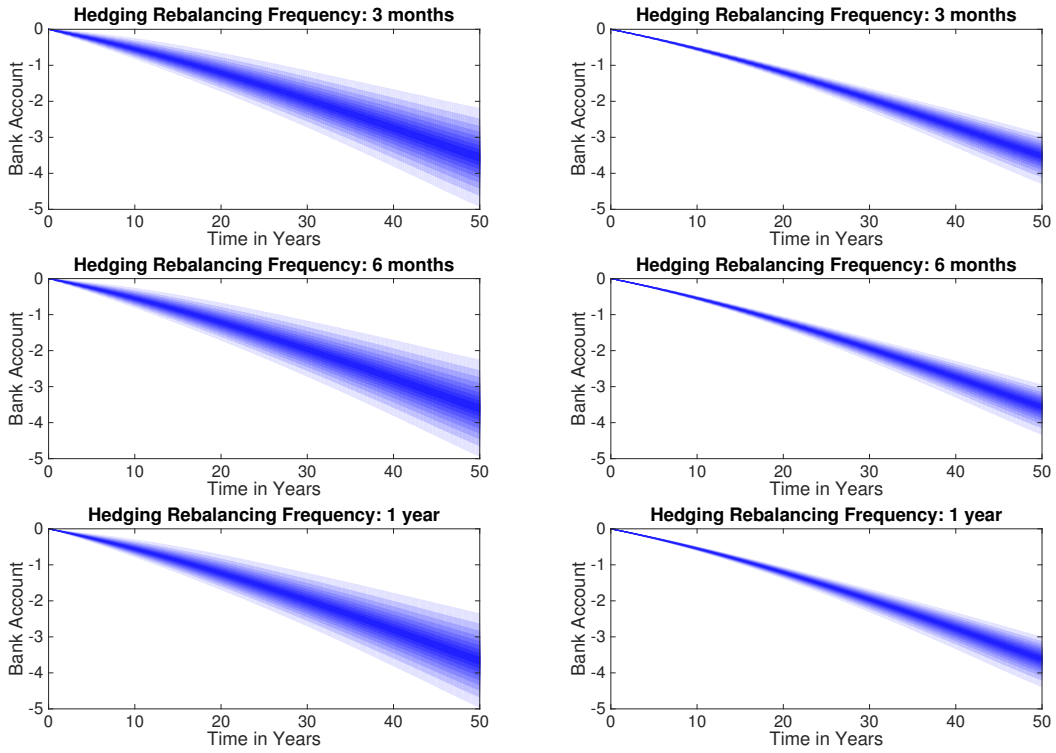
unknowns:

$$\begin{cases} -\frac{\partial N^{pp}(t)}{\partial I}dI + \sum_{i=1}^{4} n_i \frac{\partial M_i(t)}{\partial I}dI = 0, \\ -\frac{\partial^2 N^{pp}(t)}{\partial I^2}(dI)^2 + \sum_{i=1}^{4} n_i \frac{\partial^2 M_i(t)}{\partial I^2}(dI)^2 = 0, \\ -\frac{\partial N^{pp}(t)}{\partial J}dJ + \sum_{i=1}^{4} n_i \frac{\partial M_i(t)}{\partial J}dJ = 0, \\ -N^{pp}(t) + \sum_{i=1}^{4} n_i M_i(t) = 0. \end{cases} \tag{2.36}$$

The third equation in system (2.36) is the Interest Rate risk Delta condition, which sets the first order sensitivity of the hedging portfolio to the financial risk factor J equal to the one of the annuity.

Figure 15 shows the simulated percentiles of the Bank Account, from the 5th to the 95th, when longevity risk is Delta-Gamma hedged and interest rate risk is unhedged (Figure 15a) and Delta hedged (Figure 15b). In each case, the differences due to the rebalancing frequency are minimal, but we immediately see that Delta-hedging the interest rate risk substantially reduces the variability of the hedging error. To compare the two cases more in detail, Figure 16a and Figure 16b show, the distribution of the Bank Account at $t = 30$ years, for each rebalancing frequency, when longevity risk is Delta-Gamma hedged and interest rate risk is unhedged and when also Delta-hedging of the interest rate is performed, respectively.

We observe that, in both cases, the Bank account is not centered at zero and, even though the two distributions have similar means, when the financial risk factor is Delta hedged the standard deviation of the Bank Account is significantly lower. The moments of the two distributions are listed in Panel B and C of Table 12. If we compare them with those in Pane A, that refers to the case without interest rate risk, we see that when the financial risk factor is hedged, either fully or partially, the standard deviation of the Bank Account reduces by more than 50%.



(a) Δ-Γ Longevity + Unhedged Interest Rate (b) Δ-Γ Longevity + Δ Interest Rate

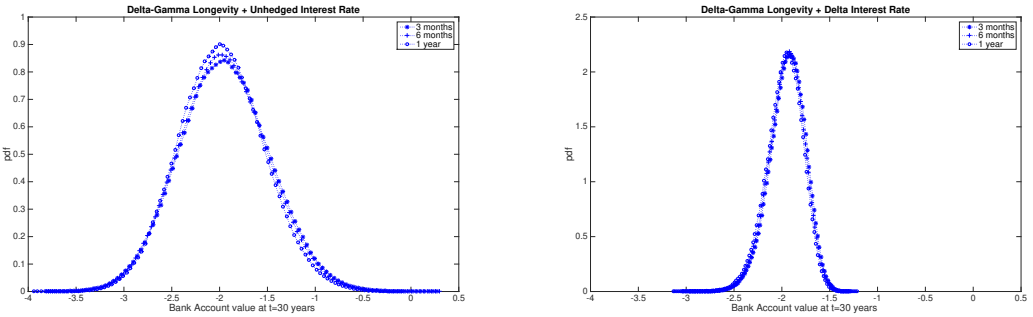
Figure 15: Percentiles, from the 5th to the 95th, of the Bank Account for different rebalancing frequencies.

2.4.3 Delta-Gamma-Theta hedging with Basis and Interest Rate Risk

We now turn our attention to the effects of interest rate risk on the effectiveness of the Delta-Gamma-Theta longevity risk hedging strategy. Again, our base-line case is the one presented in De Rosa et al. [2017], where the mortality risk factor is Delta-Gamma-Theta hedged and the interest rate risk is fully hedged and kept constant at $r_0 = 2\%$. Here, we analyze the performance of a mortality Delta-Gamma-Theta hedge when interest rate risk is unhedged or Delta-Theta hedged. In order to hedge the longevity risk factor, we use three longevity bonds with rolling maturities $T_i, i = 1, 2, 3$ of 10, 15, 20 years respectively and a deterministic deposit account, $K(t, dt)$. Its maturity is equal to the rebalancing frequency, allowing us to cover the deterministic changes in the value of the

Table 12: Moments of the hedging error of the Delta-Gamma strategy under different rebalancing frequencies and different assumptions on interest rate risk.

Panel A: Δ-Γ Longevity + No Interest Rate			
	3 months	6 months	1 year
Mean	2.08373	2.09597	2.11164
Std	0.20691	0.20715	0.20850
Panel B: Δ-Γ Longevity + Unhedged Interest Rate			
	3 months	6 months	1 year
Mean	1.9458	1.9605	1.9803
Std	0.4723	0.4642	0.4427
Panel C: Δ-Γ Longevity + Δ Interest Rate			
	3 months	6 months	1 year
Mean	1.9416	1.9560	1.9729
Std	0.1861	0.1866	0.1873



(a) Δ-Γ Longevity + Unhedged Interest Rate (b) Δ-Γ Longevity + Δ Interest Rate

Figure 16: (a) Distribution of the value of the Bank Account, under the assumption of basis risk, at t=30 years for different rebalancing frequencies with unhedged Interest Rate risk. (b) Distribution of the value of the Bank Account, under the assumption of basis risk, at t=30 years for different rebalancing frequencies with Delta hedged Interest Rate risk.

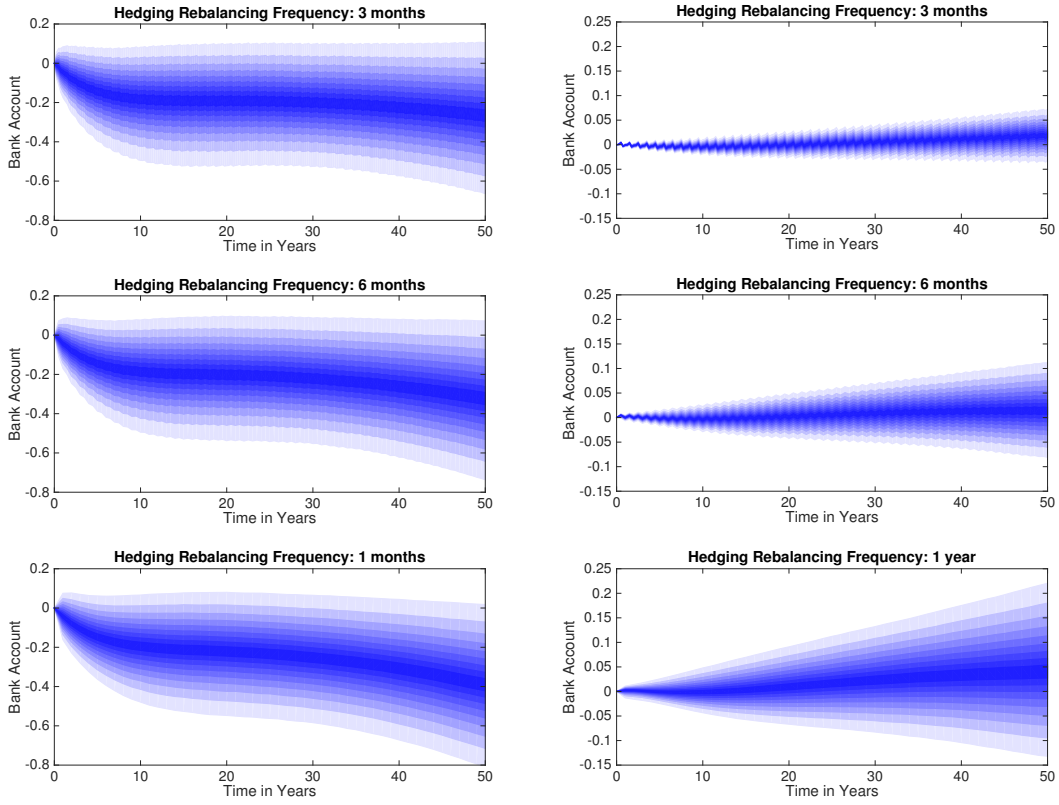
Annuity. The Delta-Theta hedging of the interest rate risk is, instead, performed using a zero-coupon bond $D(t,T)$, with a rolling maturity of $T = 10$ years. In this case the composition of the hedging portfolio can be determined, at each rebalancing date, by solving the following system of 5 equations in 5 unknowns:

$$
\begin{cases}
-\frac{\partial N^{pp}(t)}{\partial I}dI + \sum_{i=1}^{3} n_i \frac{\partial M_i(t)}{\partial I}dI = 0, \\
-\frac{\partial^2 N^{pp}(t)}{\partial I^2}(dI)^2 + \sum_{i=1}^{3} n_i \frac{\partial^2 M_i(t)}{\partial I^2}(dI)^2 = 0, \\
-\frac{\partial N^{pp}(t)}{\partial t}dt + \sum_{i=1}^{3} n_i \frac{\partial M_i(t)}{\partial t}dt + n_4 \frac{\partial K}{\partial t}dt + n_5 \frac{\partial D(t,T)}{\partial t}dt = 0, \\
-\frac{\partial N^{pp}(t)}{\partial J}dJ + \sum_{i=1}^{3} n_i \frac{\partial M_i(t)}{\partial J}dJ + n_5 \frac{\partial Z}{\partial J}dJ = 0, \\
-N^{pp}(t) + \sum_{i=1}^{3} n_i M_i(t) + n_4 K(t) + n_5 D(t,T) = 0.
\end{cases} \tag{2.37}
$$

The third equation of system (2.37) is the Theta condition, while the fourth one is the Delta condition with respect to interest rate risk. It is worth noticing that Delta-Theta hedging of the interest rate risk does not require any additional instrument with respect to the Delta hedging strategy only.

Figure 17 shows the evolution over time of the simulated percentiles of the Bank Account of the Delta-Gamma-Theta strategy, when the interest rate risk is left unhedged (left panel) and when it is instead Delta-Theta hedged (right panel). In the latter case, the variability of the Bank Account is considerably reduced, compared to the unhedged interest rate case. Moreover, the average hedging error is also closer to zero. It is clearly possible to identify a reduction in the variability of the hedging error, as the rebalancing frequency increases. In these figures, the percentiles of the Bank Account range between -0.8 and 0.2 when the interest rate risk is not hedged and between -0.15 and 0.25 when we cover it through Delta-Theta hedging.

Figures 18a and 18b show the distribution of the Bank Account at time $t = 30$ years for the two cases considered, while Panel B and C of Table 13 report the means and standard deviations of the corresponding hedging errors. If the rebalancing frequency is 1 year, not covering the interest rate increases both the mean and the variance of the hedging error by a factor 10, compared to the case when interest rate risk is fully hedged



(a) Δ-Γ-Θ Longevity + Unhedged Interest Rate (b) Δ-Γ-Θ Longevity + Δ-Θ Interest Rate

Figure 17: Percentiles, from the 5th to the 95th, of the Bank Account for different rebalancing frequencies.

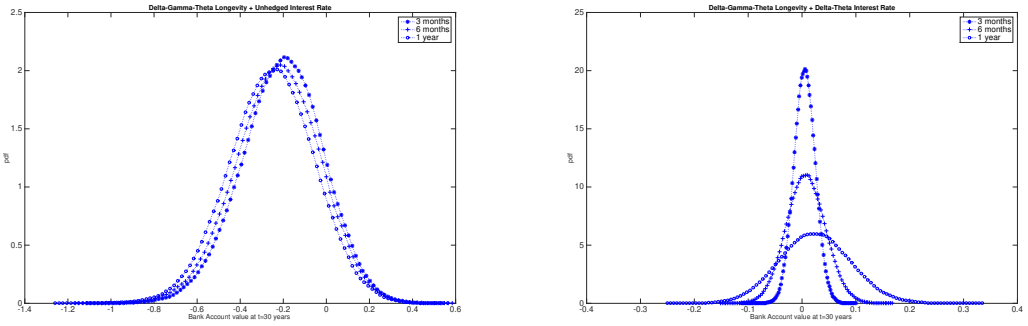
(Panel A Table 13). However, simply adding a zero-coupon bond to the hedging portfolio and performing a Delta-Theta hedge of J significantly improves the performances of the strategy.

2.4.4 Risk Margins

We assume that, when an annuity contract is sold, the Insurer has to compute a risk margin, in order to ensure her solvency at a future instant $t = 30$ years. Consistently with the Solvency II framework, the risk margin is computed as the 99.5% confidence

Table 13: Moments of the hedging error of the Delta-Gamma-Theta strategy under different rebalancing frequencies and different assumptions on interest rate risk.

Panel A: Δ-Γ-Θ Longevity + No Interest Rate			
	3 months	6 months	1 year
Mean	0.00897	0.01146	0.02362
Std	0.00522	0.00797	0.01657
Panel B: Δ-Γ-Θ Longevity + Unhedged Interest Rate			
	3 months	6 months	1 year
Mean	0.2295	0.2483	0.2724
Std	0.1588	0.1665	0.1737
Panel C: Δ-Γ-Θ Longevity + Δ-Θ Interest Rate			
	3 months	6 months	1 year
Mean	0.0162	0.0292	0.0555
Std	0.0123	0.0222	0.0422



(a) Δ-Γ-Θ Longevity + Unhedged Interest Rate (b) Δ-Γ-Θ Longevity + Δ-Θ Interest Rate

Figure 18: (a) Distribution of the value of the Bank Account, under the assumption of basis risk, at t=30 years for different rebalancing frequencies with Δ-Γ-Θ hedged Longevity and unhedged Interest Rate risk. (b) Distribution of the value of the Bank Account, under the assumption of basis risk, at t=30 years for different rebalancing frequencies with Δ-Γ-Θ hedged Longevity and Δ-Θ hedged Interest Rate risk.

level Value-at-Risk of portfolio value. We compare the value of this risk margin under different assumptions regarding the hedging of longevity and interest rate risk. We express these risk margins as a percentage of the annuity price at time zero. Our results are summarized in Table 14. The risk margin for the annuity is about 27% either when the interest rate risk is fully hedged, either when it is completely unhedged. In both

Annuity Unhedged - No Interest Rate Risk		
3 months	6 months	1 year
27.85%	27.85%	27.63%
Annuity Δ-Γ Hedged - No Interest Rate Risk		
3 months	6 months	1 year
9.32%	9.34%	9.42%
Annuity Δ-Γ-Θ Hedged - No Interest Rate Risk		
3 months	6 months	1 year
0.077%	0.126%	0.263%
Annuity Unhedged - Interest Rate Risk Unhedged		
3 months	6 months	1 year
27.28%	27.49%	27.27%
Annuity Δ-Γ Hedged - Interest Rate Risk Unhedged		
3 months	6 months	1 year
11.25%	11.18%	11.04%
Annuity Δ-Γ Hedged - Interest Rate Risk Δ Hedged		
3 months	6 months	1 year
8.94%	9.02%	9.06%
Annuity Δ-Γ-Θ Hedged - Interest Rate Risk Unhedged		
3 months	6 months	1 year
2.60%	2.73%	2.84%
Annuity Δ-Γ-Θ Hedged - Interest Rate Risk Δ-Θ Hedged		
3 months	6 months	1 year
0.208%	0.374%	0.713%

Table 14: Risk Margins

cases, Delta-Gamma hedging the mortality risk factor reduces the risk margin by about two thirds. When interest rate risk is fully hedged, the Delta-Gamma-Theta strategy is able to reduce the risk margin to almost zero[3]. Therefore, from the case when the Delta-Gamma-Theta strategy is performed and the financial risk factor is left unhedged, we can quantify how much the interest rate risk is contributing to the total risk margin. In particular, we see that hedging the financial risk reduces the risk margin between 2%-3% of the initial price of the annuity. By comparing the totally unhedged strategy

[3]Especially when the rebalancing frequency is low, that is, more frequent rebalancing.

and the Delta-Gamma-Theta longevity hedged one, interestingly, we find that interest rate risk accounts only for 10% of the total risk margin. The remaining 90% of the risk margin is, instead, due to the uncertainty about the future mortality rates.

2.5 Summary and conclusions

The consideration of interest rate risk is important when insurance liabilities are evaluated at fair value because it affects the computation of actuarial reserves through the discount factors. Moreover, careful attention should be used when Asset-Liability management is performed, because both the Assets and the Liabilities of the annuity providers and insurers are affected by a common source of risk, namely, the financial risk factor.

In this paper we first studied the impact of interest rate risk on the effectiveness of dynamic longevity hedging strategies when also basis risk is present. Second, we investigated how risk margins change as consequence of different hedging choices.

We found that hedging only the longevity risk factor and leaving the financial one unhedged can significantly increase the variability of the hedging error of the strategy. The standard deviation of the hedging error of a longevity Delta-Gamma hedge doubles when interest rate risk is present and not managed. The impact is even higher for the Delta-Gamma-Theta hedge, where the standard deviation increases by more than 10 times. However, as shown in Tables 12 and 13, a partial interest rate hedge, such as a Delta or Delta-Theta hedge, can substantially improve the overall effectiveness of the strategy.

Even if the impact of interest rate risk on the effectiveness of longevity hedging strategies can be high, its contribution to the risk margin of an annuity contract can be significantly smaller than the mortality risk factor when the horizon of interest lenghtens. In particular, we found that, in the very long-run (30 years), the longevity risk factor accounts for 90% of the total risk margin, while the financial risk factor accounts only 10% of the risk margin.

Chapter 3

Geographical diversification and longevity risk mitigation in annuity portfolios[1]

3.1 Introduction

In the last twenty years, insurance companies have been expanding internationally, via subsidiaries operating in different countries or via cross-border mergers and acquisitions. The largest insurers and re-insurers are indeed multinational companies, with subsidiaries and branches located in several countries. Between 1990 and 2003, namely before the introduction of Solvency II, the internationalization of banks and insurance companies followed similar patterns, as argued by Focarelli and Pozzolo, 2008. As of 2012, namely after the adoption of Solvency II, instead, Schoenmaker and Sass [2016] argue that the share of cross-border activity in the insurance sector is higher than in the banking one, and that the degree of internationalization of the 25 largest European insurers increased over the period 2000-2012, despite the financial crisis.
A possible explanation for the higher level of internationalization of insurance companies

[1]The material for this chapter is taken from De Rosa, Luciano, and Regis [2019].

relative to banks after the adoption of Solvency II in 2009 is the fact that, on top of financial risks, whose diversification costs and benefits may be similar between banks and insurers, through internationalization insurers and re-insurers may reap also longevity risk diversification benefits, and the consequent regulatory burden relief.
Longevity risk is the risk of experiencing losses, due to unexpected fluctuations in mortality rates. It affects annuity providers in particular when, as occurred in the last decades, policyholders' longevity exceeds the expectations. Why is international diversification of life insurance portfolios beneficial in terms of risk? Because, even if – in expectation – longevity has been steadily increasing on a worldwide scale, idiosyncratic longevity risks of different populations may be non-perfectly correlated across countries. As a result, pooling portfolios of policies written on the lives of different populations allows to diversify longevity risk and reduce the regulatory capital required by Solvency to face it.

This work aims at filling a gap in the literature by providing three measures of the diversification effects deriving from longevity risk pooling across populations. As an example of the extra-benefit of international diversification proper of life insurers and re-insurers, we consider an annuity provider, who can decide to expand her portfolio by selling policies to members of a population different from the one to which she is currently exposed. Our goal is to quantify the diversification benefit deriving from such an expansion, relative to an expansion of the portfolio not involving internationalization. To this end, we first introduce a novel parsimonious model for the joint mortality dynamics of policyholders in different countries, which extends the model presented in De Rosa et al. [2017] to a multi-population setting. We set ourselves in the continuous-time framework, which has gained increasing popularity, alongside the more traditional discrete-time one, because of its analytical tractability. The model we propose is a stochastic, continuous-time multi-population extension of the deterministic Gompertz mortality law, a benchmark in the classical modelling of mortality arrival rates. It allows to compute survival probabilities and hedge ratios in closed form, differently from

models à la Lee-Carter. At the same time, it allows a rich description of the mortality dynamics of multiple populations, and generations within them. In the continuous-time framework, some models cope with the longevity risk of two cohorts or populations (Dahl et al., 2008), or several cohorts within one population (Blackburn and Sherris, 2013; Jevtić et al., 2013). Up to our knowledge, only Sherris et al. [2018] and Jevtić and Regis [2019] have attempted to combine the description of the mortality intensities of multiple populations and generations together in a continuous-time setting. Both these papers apply models driven by three independent Brownian motions (risk factors) and entail Gaussian intensities. Our model is richer, because we assume as many dependent risk factors as domestic generations and an idiosyncratic source that drives the mortality intensity of the foreign population. Thus, we are able to capture the correlation structure of different generations within and across populations accurately, while preserving a good level of parsimony. Our intensities follow square-root processes, and therefore can not become negative.

Building on our model, we provide three measures for the longevity risk of a portfolio, which we use to describe the effects of geographical diversification on a portfolio. The first is the percentage risk margin, computed à la Solvency II. To evaluate it, we define the value of the portfolio as the sum of the actuarial value of the policies (best estimate) and of a risk margin, i.e. an amount that the insurer has to set aside to cover up for the unhedgeable risks. The risk margin is defined as the value at risk (VaR) of the unexpected loss in the portfolio value at a certain confidence level. If it decreases, after an international expansion, its change provides a dollar-based measure of the benefits of diversification, because its reduction is the capital requirement relief for the insurer. The second measure is the standard deviation of the portfolio mortality intensity, which measures the volatility of the cohort-based intensity, weighted by the relevance of each cohort in the portfolio. It gives an indication of the longevity risk of the portfolio, because it measures the dispersion around the mean of the average mortality intensity. The third measure, that we call "Diversification Index" is an average of the dissimilarities between the same cohorts in different populations, present in both the initial portfolio

and in the portfolio after the expansion, weighted by the percentage of policies belonging to the cohort. The three measures, analyzed together, provide three complementary views to assess the risk effects of the international expansion. The risk margin reduction offers a measure of the mitigation of the "tail" risk, because it represents the loss in a worst-case scenario occurring with a low probability. The standard deviation of the portfolio mortality intensity provides a measure of volatility of the portfolio, due to longevity risk. The diversification index, finally, assesses in a simple way how dissimilar a portfolio after an international expansion is relative to the initial one.

In a numerical application, which portrays the situation of a UK annuity provider that can expand to Italy, we first assess that the model is able to fit well the observed mortality rates of individuals aged 65-75 in the two populations, while capturing, using the Gaussian mapping technique, the imperfect correlations observed across ages and populations. Based on our model estimates, we then compute our international diversification measures for different portfolio expansions. We show that the risk margin reduction can be as high as 3% as a proportion of the actuarial value, in the case of a foreign expansion, targeted to those cohorts in the Italian population who have low covariance with the initial annuity portfolio. We also highlight that longevity risk mitigation effects are more sizable when the interest rate – a flat term structure, for simplicity – is lower. The expansion can be performed, at a practical level, by starting foreign branches, acquiring foreign undertakings or, as shown in an appendix, through the use of longevity derivatives.

3.2 Set up

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, endowed with the usual properties, where $\mathcal{F}$ is the filtration containing the information regarding all the relevant variables and $\mathbb{P}$ is the historical probability measure. In this probability space, the mortality intensities of individuals are described as stochastic processes, and longevity risk, i.e. the risk of unexpected fluctuations in the likelihood of deaths of individuals, arises. In

what follows, we will consider longevity risk as the only source of risk in our setup.
We consider an Annuity Provider, or Life-Insurer, based in a certain country (that we call *Domestic*), having a portfolio of deferred annuities written on different cohorts belonging to the Domestic Population. Let $\mathcal{X} = \{x_1, \ldots, x_m\}$ be the set of annuitants' ages at time zero, and let n_i, for $i = 1, \ldots, m$, be the number of annuities sold to people aged x_i. When an annuity is sold at time zero, the annuitant pays an initial premium. We compute the actuarial value of the liabilities net of that premium. After signing the contract, the annuitant will receive a series of fixed annual instalments R, starting from the year-end of his 65-th birthday if $x_i < 65$, or immediately if $x_i \geq 65$, until his death, that may happen at most when he reaches a final age ω, at which he will die with probability 1.

3.2.1 Portfolio value

In Europe, the life-insurance business falls under the Solvency II regulation, that requires insurers to value their liabilities at market value and set aside VaR-based risk margins with respect to the sources of risk that affect these valuations. These risk margins are amounts prudentially set aside by the insurer, meant as financial covers for the unhedgeable risk that the insurer bears. We consider, then, that the overall value $\Pi^0(t)$ of the liability portfolio of a life insurer at time t is the sum of two components: the Actuarial Value $AV_{\Pi^0}(t)$, which is the sum of the actuarial values of each individual contract $N_i(t)$ and represents a best estimate of the liabilities of the insurer, and the Risk Margin $RM_{\Pi^0}(t)$ of the portfolio itself. In formulas, we have that:

$$\Pi^0(t) = AV_{\Pi^0}(t) + RM_{\Pi^0}(t) = \sum_{i=1}^{m} n_i N_i(t) + RM_{\Pi^0}(t). \tag{3.1}$$

We now detail further the assumptions we make to compute the two components. The actuarial value of the contract is its fair premium. To compute it, we first define the number of years before the individual i aged x_i reaches age 65 as $\tau = max(65 - x_i, 0)$. If $\tau > 0$, then the contract is a deferred annuity, while if $\tau = 0$ the contract is an

immediate annuity. Because we consider no risk source other then longevity risk, the actuarial value of an annuity can be expressed as

$$N_i(t) = D(t, t+\tau)S_i(t, t+\tau)\left[R\sum_{u=1}^{\omega-t-\tau} D(t+\tau, t+\tau+u)S_i(t+\tau, t+\tau+u)\right], \quad (3.2)$$

where $D(t,s)$, $s \geq t$ denotes the deterministic financial discount factor, $D(t,s) = e^{-r(s-t)}, r \in \mathbb{R}$ and $S_i(t,\cdot)$ is the time-t survival probability curve of the individual aged x_i at time t.

We define the portfolio risk margin $RM_{\Pi^0}(t)$ as the discounted Value-at-Risk, at a given confidence level $\alpha \in (0,1)$, of the unexpected portfolio's future actuarial value at a given time horizon T:

$$RM_{\Pi^0}(t) = D(t, t+T) \cdot VaR_\alpha\big(AV_{\Pi^0}(t+T) - \mathbb{E}_t[AV_{\Pi^0}(t+T)]\big), \quad (3.3)$$

$$= D(t, t+T) \cdot inf\{l \in \mathbb{R}^+ : \mathbb{P}(AV_{\Pi^0}(t+T) - \mathbb{E}_t[AV_{\Pi^0}(t+T)] > l) < 1-\alpha\}, \quad (3.4)$$

where $\mathbb{P}(\cdot)$ denotes the probability of the event that the future actuarial value exceeds its time-t expected value by more than l.

3.2.2 Portfolio Expansion

In our setup, we consider the case in which the Insurer wants to expand the size of her annuity portfolio and can choose between two alternative strategies. The first one consists simply in selling new contracts to her own Domestic population. In this case, we denote with n_i' the number of new contracts sold to individuals aged x_i, with Π^D the portfolio composed of just these new annuities, and with Π^1 the portfolio after the expansion, composed of the old and the new contracts. The actuarial value of the new portfolio is simply

$$AV_{\Pi^D}(t) = \sum_{i=1}^{m} n_i' N_i(t), \quad (3.5)$$

and

$$AV_{\Pi^1}(t) = AV_{\Pi^0}(t) + AV_{\Pi^D}(t). \tag{3.6}$$

The value of the total portfolio Π^1 is the sum of the actuarial value of the old portfolio, the actuarial value of the new portfolio and the risk margin of the total portfolio:

$$\Pi^1(t) = AV_{\Pi^1}(t) + RM_{\Pi^1}(t) = AV_{\Pi^0}(t) + AV_{\Pi^D}(t) + RM_{\Pi^1}(t). \tag{3.7}$$

The second possible strategy is to acquire a new portfolio of annuities Π^F, written on a foreign population. We assume that, for each age x_i, the number of annuities written on people aged x_i in the foreign population is n_i^f. The actuarial value of portfolio Π^F is

$$AV_{\Pi^F}(t) = \sum_{i=1}^{m} n_i^f N_i^F(t). \tag{3.8}$$

We denote with Π^2 the portfolio obtained after the expansion towards the foreign country. The actuarial value of such portfolio is

$$AV_{\Pi^2}(t) = AV_{\Pi^0}(t) + AV_{\Pi^F}(t) \tag{3.9}$$

and its overall value is

$$\Pi^2(t) = AV_{\Pi^2}(t) + RM_{\Pi^2}(t) = AV_{\Pi^0}(t) + AV_{\Pi^F}(t) + RM_{\Pi^2}(t). \tag{3.10}$$

Notice that the original portfolio and the one obtained after the expansion do not have the same actuarial value, neither when the expansion is domestic nor foreign. The risk margin of the two portfolios is different as well. Our aim is to measure the effects of the two alternative strategies on the longevity risk profile of the insurer. To this end, in the next sections we introduce a novel longevity risk model and three measures of the diversification effects.

3.3 Longevity Risk Modeling

We now turn to the description of the source of uncertainty that affects the value of the Insurer's portfolio: the risk of longevity, i.e. the risk that her policyholders live longer than expected. We set ourselves in the well-established continuous-time stochastic mortality setting initiated by Milevsky and Promislow [2001] that models the death of individuals as a Cox process. The time to death of an individual belonging to cohort x_i is the first jump time of a Poisson process with stochastic intensity. This intensity is indeed the force of mortality of the individual. When we consider different populations, and different cohorts within each population, it is reasonable to assume that their mortality intensities processes will be different, even though they may be (even closely) related one another. In this section, we propose a novel, parsimonious model to describe the evolution of the mortality intensities of several cohorts in two different populations. The parsimony of our approach stems from making the intensity of one population (the "foreign" one) a linear combination of the other, benchmark, population's intensity (the "domestic" one) and of an idiosyncratic risk factor. This makes the whole correlation structure across populations dependent on the weight of the linear combination.

To preserve tractability, allowing for closed form expressions for the survival probabilities, but at the same time ensuring non-negativity of the intensities, we adopt stochastic processes belonging to the affine family, of the Cox et al. [1985b] type. These models have been used in single-country longevity modeling by Dahl et al. [2008] and Luciano et al. [2012b].

3.3.1 Mortality intensities and survival probabilities

Let us consider two populations, each containing m different cohorts. The first population is called the *Domestic population* and the second one is called the *Foreign population*. A given cohort i, with $i = 1, \ldots, m$, belonging to one of the two populations, is identified by the (common) initial age x_i at time zero. The set $\mathcal{X}$ of initial ages is common to the two populations.

Domestic Population

The mortality intensity of each cohort x_i, for $i = 1, \ldots, m$, belonging to the Domestic population is denoted with λ_i^d, and follows a non-mean reverting CIR process:

$$d\lambda_i^d(t) = (a_i + b_i\lambda_i^d(t))dt + \sigma_i\sqrt{\lambda_i^d(t)}dW_i(t), \tag{3.11}$$

where $a_i, b_i, \sigma_i,\ \lambda_i^d(0) \in \mathbb{R}^{++}$ are strictly positive real constants and the W_i's are instantaneously correlated standard Brownian Motions: $dW_i(t)dW_j(t) = \rho_{ij}dt$ with $i, j \in \{1, \ldots, m\}$. As a consequence, the mortality intensities of two different cohorts belonging to the Domestic Population are instantaneously correlated, as soon as $\rho_{i,j} \neq 0$.

Foreign Population

The mortality intensity of cohort x_i belonging to the Foreign population is denoted with λ_i^f, and is given by the convex combination of the mortality intensity of the corresponding cohort belonging to the Domestic population λ_i^d and an idiosyncratic component λ', which affects the Foreign population only and that depends on the initial age x_i in a deterministic way[2], i.e.

$$\lambda_i^f(t) = \delta_i\lambda_i^d(t) + (1 - \delta_i)\lambda'(t; x_i), \tag{3.12}$$

where

$$d\lambda'(t; x_i) = (a' + b'd\lambda'(t; x_i))dt + \sigma'\sqrt{d\lambda'(t; x_i)}dW'(t), \tag{3.13}$$

with $\delta_i \in [0, 1]$.[3] The parameters a', b' and σ' are positive constants, while W' is a standard Brownian Motion, that is assumed to be independent of W_i for each $i = 1, \ldots, N$.

Intuitively, the idiosyncratic risk source W' is population-specific, in the sense that it is

[2]For the empirical application in Section 3.5 we will consider the easiest case of no dependence on the initial age x_i.

[3]In principle, linear affine coefficients a', b' and σ' could be chosen.

common to all the cohorts of the Foreign population. Nonetheless, each foreign cohort x_i has a specific sensitivity to the idiosyncratic component $\lambda'(t;x_i)$, that is given by the parameter δ_i, which is, instead, cohort-specific. The mortality intensities of two different cohorts of the Foreign population are correlated, and the correlation between λ_i^f and λ_j^f depends both on the correlation between λ_i^d and λ_j^d and on the weights δ_i and δ_j. Moreover, thanks to the presence of the idiosyncratic component λ' affecting the Foreign population, our model allows to account for the non-perfect correlation between cohorts across the two populations. The correlation structure among the different cohorts of the two populations will be derived in Section 3.5.2.

From (3.11) we have that the survival probability of generation x_i in the Domestic population is given by:

$$S_i^d(t,T) = A_i^d(t,T)e^{-B_i^d(t,T)\lambda_i^d(t)}, \tag{3.14}$$

where

$$A_i^d(t,T) = \left(\frac{2\gamma_i e^{\frac{1}{2}(\gamma_i - b_i)(T-t)}}{(\gamma_i - b_i)\big(e^{\gamma_i(T-t)}-1\big)+2\gamma_i}\right)^{\frac{2a_i}{\sigma_i^2}}, \tag{3.15}$$

$$B_i^d(t,T) = \frac{2\big(e^{\gamma_i(T-t)}-1\big)}{(\gamma_i - b_i)\big(e^{\gamma_i(T-t)}-1\big)+2\gamma_i}, \tag{3.16}$$

with $\gamma_i = \sqrt{b_i^2 + 2\sigma_i^2}$. Similarly, for the Foreign population we have:

$$S_i^f(t,T) = A_i^d(t,T)A'(t,T)e^{-B_i^d(t,T)\delta_i\lambda_i^d(t) - B'(t,T)(1-\delta_i)\lambda_i'(t)}, \tag{3.17}$$

where

$$A'(t,T) = \left(\frac{2\gamma' e^{\frac{1}{2}(\gamma' - b')(T-t)}}{(\gamma' - b')\big(e^{\gamma'(T-t)}-1\big)+2\gamma'}\right)^{\frac{2a'}{(\sigma')^2}}, \tag{3.18}$$

$$B'(t,T) = \frac{2\big(e^{\gamma'(T-t)}-1\big)}{(\gamma' - b')\big(e^{\gamma'(T-t)}-1\big)+2\gamma'}, \tag{3.19}$$

with $\gamma' = \sqrt{(b')^2 + 2(\sigma')^2}$. The time-t survival probability curves of the two populations

are thus both available in closed form, and depend on the parameters of the model.

3.3.2 Variance Covariance Structure

The model we proposed allows the computation of the variance of each generation's mortality intensity, as well as the covariance between generations within and across population.

The time-t variance of the intensity of a generation i belonging to the Domestic population, λ_i^d, conditional on the information at time 0 is available in closed form and is equal to

$$Var_0(\lambda_i^d(t)) = \frac{a_i\sigma_i^2}{2b_i^2}\left(e^{b_i t}-1\right)^2 + \frac{\sigma_i^2}{b_i}e^{b_i t}\left(e^{b_i t}-1\right)\lambda_i^d(0). \tag{3.20}$$

Similarly, the conditional variance of $\lambda_i^f(t)$ is

$$Var_0(\lambda_i^f(t)) = \delta_i^2 Var_0(\lambda_i^d(t)) + (1-\delta_i)^2 Var_0(\lambda'(t)), \tag{3.21}$$

where

$$Var_0(\lambda'(t)) = \frac{a'(\sigma')^2}{2(b')^2}\left(e^{b' t}-1\right)^2 + \frac{(\sigma')^2}{b'}e^{b' t}\left(e^{b' t}-1\right)\lambda'(0). \tag{3.22}$$

Since the mortality of the Domestic generations follow a square-root process, there is no closed form expression for the covariance between the intensities two generations i and j. However, we can obtain a closed form approximation using the Gaussian Mapping technique described in Section 3.5.2 and in Appendix3.A. Indeed, referring the reader to those sections for further details, we have that

$$Cov_0(\lambda_i^d(t), \lambda_j^d(t)) = \frac{\sigma_i^V \sigma_j^V \rho_{ij}}{b_i + b_j}\Big(e^{(b_i+b_j)t}-1\Big), \tag{3.23}$$

where σ_i^V and σ_j^V are the instantaneous volatilities resulting from the mapping of λ_i^d and λ_j^d into Gaussian processes.

From (3.11) and (3.12) we have that the covariance between the same generation i

belonging to the Domestic and Foreign population can be written as:

$$Cov_0(\lambda_i^d(t), \lambda_i^f(t)) = \delta_i Var_0(\lambda_i^d(t)). \tag{3.24}$$

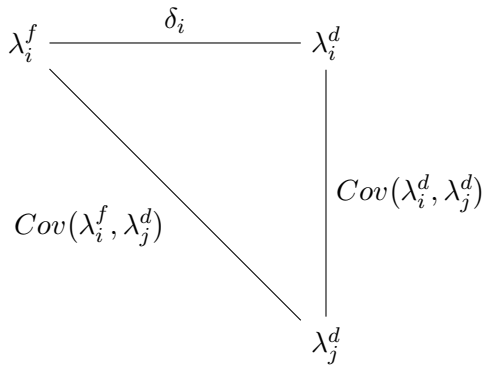
Considering, instead, two different generations i and j belonging to the Foreign population, we have that

$$Cov_0(\lambda_i^f(t), \lambda_j^f(t)) = \delta_i\delta_j Cov_0(\lambda_i^d(t), \lambda_j^d(t)) + (1-\delta_i)(1-\delta_j)Var_0(\lambda'(t)). \tag{3.25}$$

Finally, the covariance between the mortality intensity of generation i belonging to the Foreign population and generation j belonging to the Domestic is given by:

$$Cov_0(\lambda_i^f(t), \lambda_j^d(t)) = \delta_i Cov_0(\lambda_i^d(t), \lambda_j^d(t)). \tag{3.26}$$

From (3.26), it is interesting to notice that the covariance between λ_i^f and λ_j^d depends both on δ_i, which measures the dependence between the same generation i across the two populations, and on $Cov_0(\lambda_i^d, \lambda_j^d)$, which instead measures the dependence between the generations i and j within the Domestic population.



(3.27)

Diagram (3.27) visualizes that, when computing the covariance λ_i^f and λ_j^d, we are able to disentangle the effect of the two types of dependence: the within-population and the cross-population ones. The importance of (3.27) can be explained with a simple

example. Suppose there are two portfolios belonging to two populations f_1 and f_2 that are competing targets of a foreign expansion, and suppose that each portfolio is composed of annuities sold only to one generation k. The objective of the expansion is to find a foreign portfolio that minimizes $Cov(\lambda_k^{f_*}, \lambda_j^d)$, to obtain the maximum level of longevity risk diversification. Since we cannot change the covariance structure of the Domestic population, the solution to the problem is to find the portfolio Π^{f_*} such that

$$f_* = \arg \min_{x \in \{f_1, f_2\}} (\delta_k^x). \tag{3.28}$$

Then, it is sufficient to compare the δ's of the two competing foreign populations. For instance, if $\delta_k^{f_1} < \delta_k^{f_2}$, then the optimal foreign expansion target portfolio is Π^{f_1}.

3.4 Measuring the longevity risk effects of geographical diversification

In the following paragraphs we introduce some measures of longevity risk in a portfolio, which allow us to appreciate the degree of geographical diversification achieved through a foreign expansion of the annuity portfolio. The first measure is the *Percentage Risk Margin* of the portfolio, computed à la Solvency II. Comparing this measure before and after a portfolio expansion allows to appreciate the economic benefit of a foreign expansion. A reduction in the percentage risk margin is connected with a reduction of tail risk, evaluated as the portfolio losses in a worst-case scenario. This measure follows the principle with which capital requirements are computed in the current regulation. Reducing the Percentage Risk Margin of a portfolio can thus be connected to a reduction in the regulatory capital requirement for longevity risk.

The second measure we propose is the *Standard Deviation of the Portfolio Mortality Intensity*. We define the portfolio mortality intensity as a weighted average of the cohort-based mortality intensities entering a portfolio, with weights equal to the percentage of policies written on each generation. A reduction of this quantity indicates a stronger

concentration of the distribution of the portfolio mortality intensity around its mean, denoting a reduction of longevity risk. Finally, the *Diversification Index* is an average of the degree of dissimilarity of the mortality intensities of the cohorts in different populations. This measure is a synthetic way of quantifying the level of diversification achieved by a foreign expansion.

3.4.1 Percentage Risk Margin

To be able to compare the effects of an expansion, we consider first a normalized quantity, i.e. the ratio of the risk margin and the actuarial value of a portfolio Π, which we call percentage risk margin:

$$\%RM_{\Pi} = \frac{RM_{\Pi}(t)}{AV_{\Pi}(t)}. \tag{3.29}$$

A lower percentage risk margin denotes a lower percentage loss in the worst-case scenario, relative to portfolio value. Hence, reducing this measure is beneficial for the company in two respects. First, it indicates a mitigation in the risk connected to adverse scenarios. In this sense, the risk margin can be considered as a measure of the systemic risk that the company may generate, by triggering losses that will hit its creditors. Second, it represents a capital requirement reduction, which frees up resources. Because the risk margin can be interpreted as both a capital requirement and a measure of the loss the company can generate – at a given level of confidence – among its creditors, it is then conceivable that minimizing the percentage risk margin aligns the interests of both the insurance company and its regulators. In what follows we take the point of view of the insurer, taking for granted the alignment of her interest with the ones of the regulator.

3.4.2 Standard Deviation of the Portfolio Mortality Intensity

Another measure of the diversification effects deriving from longevity risk pooling across populations can be derived by looking at the change in the standard deviation of the portfolio mortality intensity pre- and post-foreign expansion. Given an annuity portfolio Π, we define its portfolio mortality intensity λ^{Π} as the weighted average of the mortality

intensities of each generation in the portfolio, where the weights are the percentages of contracts written on each generation. Considering the initial Domestic portfolio Π^0, let n_i^d be the number of contracts sold to generation i belonging to the Domestic population, and let $n^d = \sum_{i=1}^m n_i^d$ be the total number of contracts in the portfolio. Then, we define λ^{Π^0} as:

$$\lambda^{\Pi^0}(t) = \sum_{i=1}^{m} \frac{n_i^d}{n^d} \lambda_i^d(t) = \sum_{i=1}^{m} w_i^d \lambda_i^d(t), \tag{3.30}$$

where $\omega_i^d = \frac{n_i^d}{n^d}$ is the weight for each generation i of the domestic population. Similarly, let $n = n^d + n^f$ be the total number of contracts in the portfolio, Π^2, after a foreign expansion in which n^f contracts are written on the target foreign population, n_i^f on each generation i. The mortality intensity of the portfolio Π^2 is given by:

$$\begin{aligned}
\lambda^{\Pi^2}(t) &= \sum_{i=1}^{m} \frac{n_i^d}{n} \lambda_i^d(t) + \frac{n_i^f}{n} \lambda_i^f(t) \\
&= \sum_{i=1}^{m} \frac{n_i^d}{n} \lambda_i^d(t) + \sum_{i=1}^{m} \frac{n_i^f}{n} \lambda_i^f(t) \\
&= \sum_{i=1}^{m} w_i^{d,\Pi^2} \lambda_i^d(t) + \sum_{i=1}^{m} w_i^{f,\Pi^2} \lambda_i^f(t),
\end{aligned} \tag{3.31}$$

where $w_i^{d,\Pi^2} = \frac{n_i^d}{n}$ and $w_i^{f,\Pi^2} = \frac{n_i^f}{n}$ represent the weights in the portfolio for each generation of the domestic and foreign population, respectively. Starting with the initial Domestic portfolio Π^0 and its mortality intensity λ^{Π^0} defined in (3.30), we have that:

$$Var_0(\lambda^{\Pi^0}(t)) = Var_0\Big(\sum_{i=1}^{m} w_i^d \lambda_i^d(t)\Big) \tag{3.32}$$

$$= \sum_{i=1}^{m} (w_i^d)^2 Var_0(\lambda_i^d(t)) + 2 \sum_{i<j} w_i^d w_j^d Cov_0(\lambda_i^d(t), \lambda_j^d(t)). \tag{3.33}$$

Thus we define the standard deviation of the portfolio Π^0 mortality intensity as:

$$\sigma^{\lambda}(\Pi^0) = \sqrt{Var_0(\lambda^{\Pi^0}(t))}. \tag{3.34}$$

Similarly, considering the post expansion portfolio Π^2, we have that:

$$Var_0(\lambda^{\Pi^2}(t)) = Var_0\Big(\sum_{i=1}^{m} w_i^{d,\Pi^2}\lambda_i^d(t) + \sum_{i=1}^{m} w_i^{f,\Pi^2}\lambda_i^f(t)\Big) \tag{3.35}$$

$$= Var_0\Big(\sum_{i=1}^{m} w_i^{d,\Pi^2}\lambda_i^d(t)\Big) + Var_0\Big(\sum_{i=1}^{m} w_i^{f,\Pi^2}\lambda_i^f(t)\Big) + \tag{3.36}$$

$$+ 2Cov_0\Big(\sum_{i=1}^{m} w_i^{d,\Pi^2}\lambda_i^d(t), \sum_{i=1}^{m} w_i^{f,\Pi^2}\lambda_i^f(t)\Big) \tag{3.37}$$

$$= \sum_{i=1}^{m} [(w_i^{d,\Pi^2})^2 + (w_i^{f,\Pi^2})^2\delta_i^2]Var_0(\lambda_i^d(t)) + \tag{3.38}$$

$$+ \sum_{i=1}^{m}(w_i^{f,\Pi^2})^2(1-\delta_i)^2 Var_0(\lambda'(t)) + \tag{3.39}$$

$$+ 2\sum_{i<j}(w_i^{d,\Pi^2}w_j^{d,\Pi^2} + w_i^{f,\Pi^2}w_j^{f,\Pi^2}\delta_i\delta_j)Cov_0(\lambda_i^d(t), \lambda_j^d(t)) + \tag{3.40}$$

$$+ 2\sum_{i<j} w_i^{f,\Pi^2}w_j^{f,\Pi^2}(1-\delta_i)(1-\delta_j)Var_0(\lambda'(t)) + \tag{3.41}$$

$$+ 2\sum_{i=1}^{m}\sum_{j=1}^{m} w_i^{d,\Pi^2}w_j^{f,\Pi^2}\delta_j Cov_0(\lambda_i^d(t), \lambda_j^d(t)), \tag{3.42}$$

and

$$\sigma^{\lambda}(\Pi^2) = \sqrt{Var_0(\lambda^{\Pi^2}(t))}. \tag{3.43}$$

A foreign expansion provides a diversification benefit if

$$\sigma^{\lambda}(\Pi^2) < \sigma^{\lambda}(\Pi^0).$$

This can happen because, after the expansion, $\lambda^{\Pi^2}(t)$ depends on λ_i^d, but also on the different risk source λ_i^f that may be non perfectly correlated with λ_i^d for $i = 1, \ldots, m$. Moreover, if there are multiple target portfolios for a foreign expansion, a possible way to decide about the optimal expansion target would be for instance to choose the portfolio that provides the lowest $\sigma^{\lambda}(\Pi^*)$.

3.4.3 Similarity/Diversification index

Building up on the characteristics of the longevity model described in the previous section, finally, we propose a synthetic measure to describe the similarity/dissimilarity between the annuity portfolios written on two populations, that we define as *Similarity* and *Diversification* index. Let n_i^d be the number of annuities written on cohort x_i belonging to the domestic population, n_i^f the number of annuities written on cohort x_i belonging to the foreign population, $n_i = n_i^d + n_i^f$ and m the number of generations in the initial, domestic portfolio. Then the Diversification Index (DI) is equal to:

$$DI = \frac{1}{m} \sum_{i=1}^{m} \frac{n_i^f (1 - \delta_i)}{n_i}, \tag{3.44}$$

and the Similarity Index (SI)[4] is:

$$SI = 1 - DI. \tag{3.45}$$

The Diversification Index represents a weighted average of the dissimilarities between the same cohorts in different populations[5], present in both the initial portfolio and in the portfolio after the expansion. Dissimilarities are captured by the complement to 1 of δ_i, the generation-specific parameter that captures the degree of correlation between the same generation of the different populations. The weights, n_i^f/n_i, are given, for each cohort in the initial portfolio, by the number of annuities in the foreign population (after the expansion) relative to the total number of annuities written on that cohort in both populations. We average the weighted dissimilarities across all the m cohorts of the domestic population initially present in the annuity portfolio.

Our proposed indicator has the following properties. First, $0 \leq DI \leq 1$. If $\delta_i = 1$ for every i, i.e. the two portfolios are written on perfectly correlated populations, then,

[4]Since we are only averaging over the generations belonging to the domestic portfolio, the Similarity Index defined in equation (3.45) should be interpreted as a synthetic measure of the similarity relative to the domestic population. If instead m is defined as the number of generations in the foreign portfolio, the resulting measure should be interpreted as the similarity with respect to the foreign portfolio.

[5]Indeed, by construction, the Diversification Index does not take into account the diversification benefit across different generations in the two populations.

obviously, $SI = 1$ and $DI = 0$. On the other hand, if $\delta_i = 0$, for every i, which means that the intensities of the foreign population are independent of the risk factor of the domestic, the DI does not go to 1 independently of the portfolio composition. If $n_i^f \to \infty$ and n_i^d remains constant, then $SI \to 0$, $DI \to 1$. This happens because the longevity risk of the foreign population is completely idiosyncratic and therefore diversification is reaped only enlarging the foreign portfolio as much as possible. This shows that the Diversification Index appropriately reflects both the properties of the intensity correlation structure and the portfolio mix chosen by the underwriter.

Let us conclude this section with some intuition behind the derivation of the DI and a comparison with $\sigma^{\lambda}(\Pi^2)$. From the definition of $\lambda^{\Pi^2}(t)$ and from (3.12) we observe that:

$$\begin{aligned}\lambda^{\Pi^2}(t) &= \sum_{i=1}^{m} w_i^{d,\Pi^2}\lambda_i^d(t) + \sum_{i=1}^{m} w_i^{f,\Pi^2}\lambda_i^f(t) \\ &= \sum_{i=1}^{m} w_i^{d,\Pi^2}\lambda_i^d(t) + \sum_{i=1}^{m} w_i^{f,\Pi^2}\delta_i\lambda_i^d(t) + \sum_{i=1}^{m} w_i^{f,\Pi^2}(1-\delta_i)\lambda'(t). \qquad (3.46)\end{aligned}$$

The last term in (3.46) can be interpreted as the source of the diversification benefit, and each coefficient of the summation $w_i^{f,\Pi^2}(1-\delta_i)$ can be interpreted as the diversification contribution of each foreign generation i. Hence the DI can be seen as the average diversification contribution of each generation in the foreign portfolio.

Recalling diagram (3.27) for the dependence structure between the foreign and domestic generations, we could say that $\sigma^{\lambda}(\Pi^2)$ captures both the horizontal and the vertical dependence, while DI only focuses on the first one.

Further insights on the properties and indications deriving from the three measures presented will emerge from the application in Section 3.5, but let us comment briefly on them before going on.

The percentage risk margin has the advantage of being expressed in economic terms, allowing a comparison between the economic benefit of a foreign expansion and its implementation cost, and between the benefits of competing target portfolios. Among the three measures, the percentage risk margin is the only one that can capture the impact of the term structure of interest rates on the economic benefit of geographical longevity

risk diversification. However, computing it – for our proposed model, at least – requires Monte Carlo simulations, making it the most computationally expensive measure among the ones presented.

The standard deviation of the portfolio mortality intensity does not require Monte Carlo simulations and can provide similar information to the risk margin when comparing different expansion strategies. It is able to capture the entire dependence structure between the domestic and foreign generations, but it is simply a distributional property and not a monetary measure.

The Diversification Index is the easiest measure to compute, because it does not require Monte Carlo simulations or the estimation of a correlation matrix. However, it does not capture the entire dependence structure between the domestic and foreign generations and, therefore, can provide useless indications when the target foreign portfolio population shows low dependence across generations and when the diversification benefit of grouping different cohorts belonging to different populations is large.

3.5 Application

In this section, we calibrate our proposed model and try to quantify the diversification gains deriving from an international expansion towards Italy of an initially UK-based annuity portfolio. The situation we consider is that of a UK annuity provider who has the option of expanding her business either in her home country or abroad, selling additional policies to Italian policyholders. In practice, this expansion can be performed by creating an Italian branch or acquiring an Italian undertaking. As an alternative, the geographical diversification can be obtained through the use of longevity derivatives (see Blake et al., 2006b for instance), which allow the insurer to gain some exposure to the mortality development of a different population. We explore this case in Appendix A.

3.5.1 Mortality intensities estimation

To calibrate our model, we proceed in two steps. First, we calibrate the parameters of the two intensity processes, of the domestic and of the foreign population respectively. Then, in a second step, we calibrate the correlation parameters ρ_{ij}. We calibrate the parameters of the mortality model to the generations of UK and Italian males whose age, at 31/12/2012, is between 65 and 75, that is, the cohorts born between 1937 and 1947. We consider thus 11 different cohorts present in the initial portfolio: $x_i = 65, \ldots, 75$. We use the 1-year×1-year cohort death rates data provided by the Human Mortality Database and recover, using the 20 observations from 1993 and 2012[6] the observed conditional survival probabilities, for each cohort, for the individuals alive in 1993. The estimation of the parameters is performed minimizing the Rooted Mean Squared Error (RMSE) between the observed and the model-implied survival probabilities. Tables 1 and 2 report the calibrated parameters for the two populations, while Figures 1 and 2 report the actual and fitted survival probabilities and the calibration errors, respectively. The model, although parsimonious, is able to capture well the survival probability curves of the two populations, for all the cohorts considered.

Table 15: Domestic Population (UK) calibration results.

Age	a	b	σ	λ_0	RMSE
65	$2.7878 \cdot 10^{-5}$	0.0723	0.0075	0.0116	0.00035
66	$6.5423 \cdot 10^{-5}$	0.0652	0.0059	0.0124	0.00028
67	$1.8424 \cdot 10^{-5}$	0.0740	0.0080	0.0135	0.00035
68	$5.3144 \cdot 10^{-5}$	0.0685	0.0084	0.0160	0.00043
69	$1.2500 \cdot 10^{-4}$	0.0589	0.0091	0.0164	0.00039
70	$8.4734 \cdot 10^{-5}$	0.0646	0.0108	0.0189	0.00056
71	$7.1323 \cdot 10^{-5}$	0.0667	0.0106	0.0212	0.00038
72	$4.1759 \cdot 10^{-5}$	0.0688	0.0073	0.0239	0.00040
73	$2.2984 \cdot 10^{-5}$	0.0689	0.0066	0.0262	0.00063
74	$9.6036 \cdot 10^{-5}$	0.0663	0.0131	0.0282	0.00040
75	$3.3898 \cdot 10^{-5}$	0.0684	0.0077	0.0316	0.00049

[6]These correspond to the last 20 observations available to date for the Italian males. However, since the UK dataset is updated until 31/12/2013, we have excluded the last available observation for the UK cohorts.

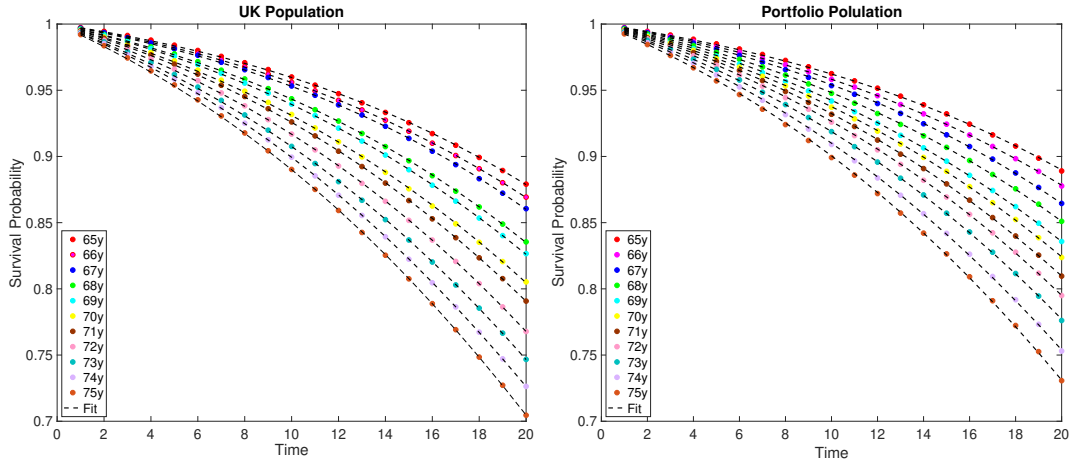


Figure 19: Observed and theoretical survival probabilities. The left panel shows the observed vs. fitted survival probabilities for the Foreign population, while the right reports the figures for the Domestic population.

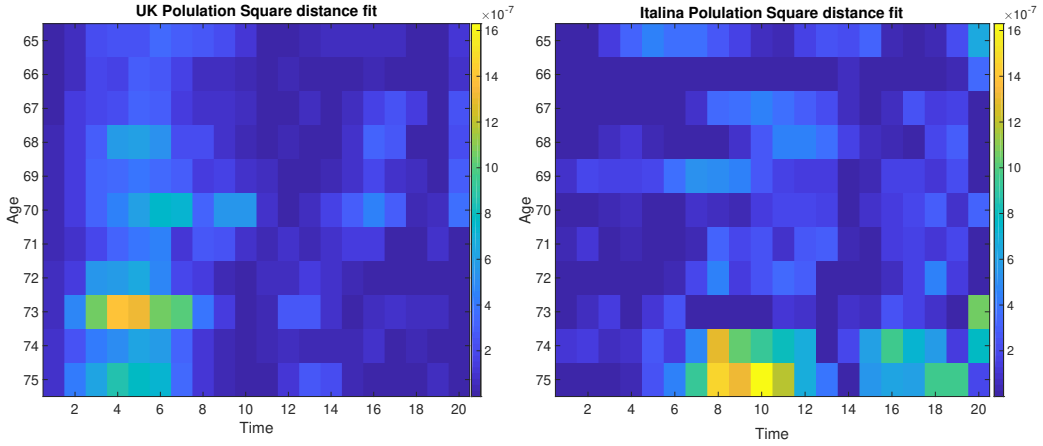


Figure 20: Calibration errors.

Table 16: Foreign Population (IT) calibration results.

Age	a'	b'	σ'	δ	RMSE	λ'_0
65				0.7939	0.00045	
66				0.8528	0.00017	
67				0.9369	0.00038	
68				0.8289	0.00036	
69				0.9039	0.00045	
70	$1.1785 \cdot 10^{-4}$	$4.7825 \cdot 10^{-7}$	0.0153	0.8362	0.00032	0.0022
71				0.8548	0.00034	
72				0.8210	0.00036	
73				0.8203	0.00036	
74				0.8484	0.00071	
75				0.8683	0.00078	

3.5.2 Correlation matrix estimation

After having estimated the cohort-specific parameters of the two populations, we turn to the estimation of their correlation structure. Having chosen a non-Gaussian process for the mortality intensities of the cohorts, we are not able to derive a formula for their correlations in closed form. However, to estimate correlations, we can apply the Gaussian Mapping technique, which has been used extensively in the pricing of Credit Default Swaps (see Brigo and Mercurio, 2001). Such technique allows to obtain a closed-form approximation of the correlations between the intensities of the different cohorts, in turn permitting the direct estimate of the correlation parameters ρ_{ij}. Technically, it consists in mapping a CIR process into a Vasicek process that is as close as possible to the original one, i.e. returning the same survival probability. Since we are able to compute analytically the correlations between each λ_i^d and λ_j^d, with $i, j = 1, \dots, N$ in the mapped Vasicek process, we can then retrieve our desired parameters in closed-form.
Starting from the CIR process (3.11) describing the mortality intensity of cohort x_i belonging to the domestic population, we consider a Vasicek process driven by the same Brownian Motion $W_i(t)$, having the same drift and the same initial point:

$$d\lambda_i^V(t) = (a_i + b_i\lambda_i^V(t))dt + \sigma_i^V dW_i(t), \quad \lambda_i^V(0) = \lambda_i^d(0). \tag{3.47}$$

The instantaneous volatility coefficient σ_i^V of (3.47) is then determined by making the two processes as close as possible. Here, having fixed a maturity T, by close we mean that the two processes return the same survival probability:

$$S_i^d(t,T) = S_i^V(t,T;\sigma_i^V). \tag{3.48}$$

Then, we approximate the correlation between $\lambda_i^d(t)$ and $\lambda_j^d(t)$ by the correlation between $\lambda_i^V(t)$ and $\lambda_j^V(t)$:

$$Corr_0(\lambda_i^d(t), \lambda_j^d(t)) \approx Corr_0(\lambda_i^V(t), \lambda_j^V(t)), \tag{3.49}$$

since this last correlation can be computed analytically. Each pair-wise correlation is a function of the parameters b_i and b_j of the mapped Vasicek process and of ρ_{ij}:

$$\begin{aligned} Corr_0(\lambda_i^V(t), \lambda_j^V(t)) &= \frac{Cov_0(\lambda_i^V(t), \lambda_j^V(t))}{\sqrt{Var_0[\lambda_i^V(t)]Var_0[\lambda_j^V(t)]}} \\ &= \frac{2\rho_{ij}}{b_i + b_j} \cdot \frac{e^{(b_i+b_j)t} - 1}{\sqrt{\frac{(e^{2b_i t}-1)(e^{2b_j t}-1)}{b_i b_j}}} \end{aligned} \tag{3.50}$$

To estimate the correlation parameters, first the parameters of the process described by (3.47) are recovered. Then, using the central mortality rates data available in the UK life tables[7], we estimate the instantaneous correlations ρ_{ij} between $d\lambda_i$ and $d\lambda_j$ by inverting the approximated correlation expression (3.50). To compute the correlations between the 11 cohorts involved, we start from the central mortality rates in 1968 of the people aged between 1 and 11, and we follow the diagonal of the life table until we reach the central mortality rates of the people aged between 65 and 75 in 2012. The central mortality rates table constructed this way has dimension 65×11 and allows to estimate the correlation coefficients which we report in Table 17. The upper and lower confidence bounds are computed with bootstrapping from $10,000$ resampled samples with replacement. Each sample has dimension 65×11, and is obtained by randomly choosing 65 times with

[7]Source: Human Mortality Database.

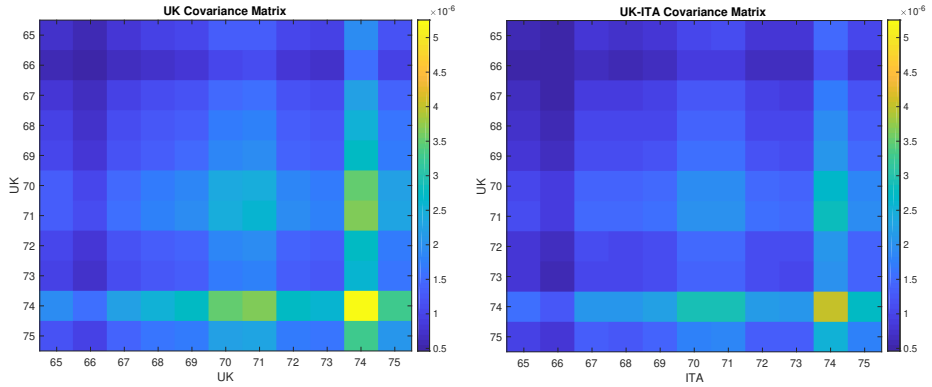


Figure 21: Left Panel: Covariance matrix between UK generations. Right Panel: Covariance matrix between Italian and UK generations.

replacement a row of our original central mortality table. As expected, because of the similarity between the UK and the Italian populations, correlations are close to 1 with tight 95% confidence bounds, but they tend to decrease with the distance between the initial ages of the two considered cohorts. This behaviour aligns with the intuition that the changes leading to longevity improvements (such as healthy habits or medical advancements) have different impact on different generations and that cohort effects are at play. Table 18 reports instead the correlations across the two populations. Also in this case, the correlations appearing in the diagonals are the highest, and they tend to decrease along the rows and column dimensions, indicating the presence of common cohort effects across populations. Figure 21 shows the covariances between the different UK cohorts, and between the UK and Italian cohorts. The two generations with the lowest covariance are the 66 years old UK and 66 years old Italian cohorts. Both the UK and Italian 66 years old cohorts are the ones with the lowest covariance with all other generations in the other country.

3.5.3 Evaluating the diversification gains in terms of risk margin

Because the oldest cohort considered in our application is 75, and we assume a maximum life span of $\omega = 105$ years, we fix the time horizon of our simulations to 30 years. Consistent with this choice, we consider a constant interest rate of 2%, matching the 30-year risk-free-rate indicated by EIOPA for the calculation of technical provisions. The choice of a constant interest rate term structure allows us to isolate and capture any possible added benefit specifically due to the geographical diversification of an annuity portfolio. The time horizon at which the Risk Margin is computed is 15 years. This choice is justified because we want to focus on the medium-long term benefits of geographical diversification. Consistently with the Solvency II regulation, we select a confidence level $\alpha = 99.5\%$ when calculating the Risk Margin associated to the portfolio.

Initial Portfolio

We consider a UK Insurer with an initial portfolio Π^0, made of 1000 contracts sold to males whose age, at 31/12/2012, is between 65 and 75. The distribution of contracts among ages reflects the proportions of individuals aged between 65 and 75 in the UK national population. For instance, since in the general UK population 69 years old constitute 11.00% of all the people aged between 65 and 75, the domestic portfolio contains 110 contracts sold to 69 years old (see Table 19). The initial Actuarial Value $AV_{\Pi^0}(0)$ of the portfolio is:

$$AV_{\Pi^0}(0) = 1.4104 \cdot 10^4, \tag{3.51}$$

while the Risk Margin computed at time 0 is

$$RM_{\Pi^0}(0) = 1.1838 \cdot 10^3. \tag{3.52}$$

Hence, the initial portfolio value is

$$\Pi^0(0) = AV_{\Pi^0}(0) + RM_{\Pi^0}(0) = 1.5288 \cdot 10^4. \tag{3.53}$$

The Risk Margin accounts for 8.39% of the initial portfolio Actuarial Value and $\sigma^{\lambda}(\Pi^0) = 0.00124$.

Portfolio Π^F is exposed to the foreign population only, distributed among ages according to Table 20, useful for comparison. As we did for the initial portfolio, we assume that the policyholders' distribution reflects the proportion of individuals belonging to each generation between 65 and 75 in the Italian population (see Table 20). Figure 22 shows the different percentage of individuals per cohort in the UK and Italian population. For the foreign portfolio Π^F, the risk margin is 7.39% and $\sigma^{\lambda}(\Pi^F) = 0.00107$. One could guess that by expanding towards Italy, the UK underwriter could, at most, reduce his risk margin to this level. However, we will show later on that, thanks to the diversification effect, the risk margin of the underwriter can be even lower.

Domestic Expansion

With a Domestic Expansion, we assume that the Insurer doubles the size of her annuity portfolio, selling additional policies to her domestic population, i.e. the UK population. The new portfolio Π^1 is therefore composed of 2000 contracts and is obtained by simply doubling the number of contracts for each generation. Hence,

$$AV_{\Pi^1}(0) = 2.8208 \cdot 10^4, \tag{3.54}$$

$$RM_{\Pi^1}(0) = 2.3676 \cdot 10^3, \tag{3.55}$$

$$\Pi^1(0) = 3.0576 \cdot 10^4. \tag{3.56}$$

The Risk Margin proportion relative to actuarial value is unaffected by the size of the portfolio, and still accounts for 8.39% of the Actuarial Value of the Domestically Expanded portfolio. Similarly, also the portfolio mortality standard deviation remains unchanged. In this case, the diversification index between Π^0 and $\Pi^1 - \Pi^0$ is 0, as no diversification gain can be obtained. However, some diversification gains could be obtained through a domestic expansion, in case the new portfolio had a different composition, in

terms of policyholders' ages, than the initial one.

Foreign Expansion

In case of a Foreign Expansion, we assume that the Insurer doubles the number of policies in its annuity portfolio by selling contracts written on policyholders belonging to the Foreign population. The composition of the Foreign portfolio per cohort is assumed to follow the same proportions of the Italian population[8] (see Figure 22). The new portfolio Π^2 is, therefore, composed of 1000 contracts sold to the UK population composing the initial portfolio and of 1000 contracts written on the Italian population, $\Pi^2 = \Pi^0 + \Pi^F$ (both distributed as described in Table 19). It has the following actuarial value and risk margin:

$$AV_{\Pi^2}(0) = 2.8872 \cdot 10^4, \tag{3.57}$$

$$RM_{\Pi^2}(0) = 2.2749 \cdot 10^3, \tag{3.58}$$

As a consequence,

$$\Pi^2(0) = 3.1147 \cdot 10^4. \tag{3.59}$$

For this portfolio, the Risk Margin accounts for 7.87% of the Actuarial Value, reduced, as expected, by 0.52 percentage points relative to the one of the initial portfolio. The portfolio mortality standard deviation $\sigma^\lambda(\Pi^2)$ consistently decreases to 0.00115, and the diversification index increases to 0.0746.

The diversification gain provided by the Foreign portfolio just described can be further exploited. We then explore alternative portfolios and summarize the results in terms of actuarial values, risk margins and total values in Table 21. Portfolio Π^3 represents a more aggressive foreign expansion, where the number of policies sold to each generation of foreign policyholders is twice the number of policies in Π^F. Tilting the portfolio towards the foreign population has the effect of decreasing the percentage risk margin

[8] Inserting the exact composition of the UK population is a trivial extension.

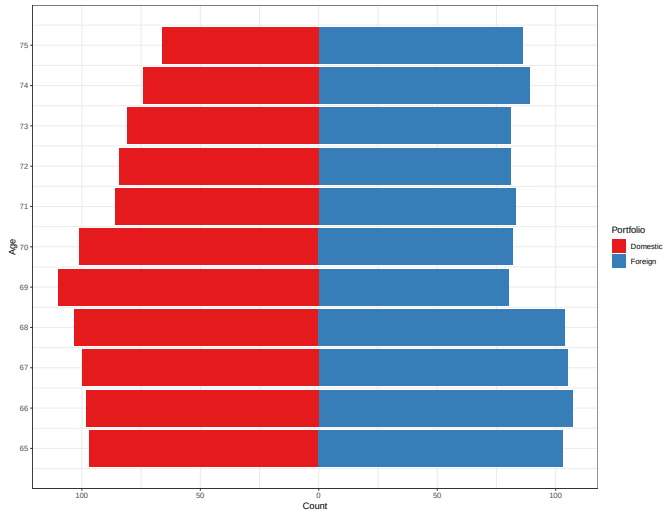


Figure 22: Domestic and Foreign Portfolio Composition.

(7.71%) and the portfolio mortality standard deviation (0.00112), while increasing the diversification index (0.0992). However, it is evident that, at most, by increasing the exposure to the Italian population, the risk margin can not be lower than 7.39%, which is the risk margin of the Foreign portfolio. This suggests to optimize the portfolio mix using not only the diversification across populations, but also across generations.

The portfolio Π^1_{opt} is obtained diversifying within the UK population. Its composition is optimized to obtain the minimum risk margin achievable, under the constraint that the number of new contracts is 1000. It can then be considered as the maximally diversified portfolio, in the absence of geographical diversification. The maximum diversification is thus obtained by selling 1000 annuities to the UK 66 years old, whose mortality intensity process shows the minimum covariance with the other UK cohorts (see the left panel of Figure 3). Notice that the percentage risk margin of this portfolio is 6.61%, which is lower than 7.39%. Being entirely composed of UK annuitants, this portfolio has a null Diversification Index, but due to its composition, is able to reduce σ^λ to 0.00096.

Similarly, Π^2_{opt} is obtained allowing for geographical diversification and optimizing the composition of the foreign portfolio. The optimization is performed by looking at the covariance matrix between the two populations(see right panel of Figure 3) and choosing to

concentrate the foreign expansion on the Italian 66 years old males, who have the lowest covariance with all the cohorts of the UK population. The risk margin of Π^2_{opt} is 6.17% and σ^λ is 0.00091. The percentage risk margin of portfolio Π^2_{opt} and its σ^λ are the lowest among the portfolios we have considered. The DI of this last portfolio is small compared to the DIs of the other portfolios involving an international expansion, being 0.0121. It is small because the expansion is performed by concentrating the sales of policies in the foreign population in one generation only. Notice that the DI and $\%RM$ reduction differ more when the portfolio added to the initial one is optimized across generations than when it is not. This happens because the DI - by definition, to be kept simple - does not capture the effects of putting different weights on generations with low within population covariance, while the percentage risk margin and the portfolio mortality standard deviation capture the entire dependence structure between populations and generations. Indeed, the Diversification Index provides a non-dollar measure of diversification which "averages" the contributions of different generations and penalizes any concentration in a particular one, even though the latter is justified by a strategy which aims at minimizing the risk margin reduction. This is why we presented all the three measures.

3.5.4 Sensitivity Analysis

Table 22 reports the results for the different portfolios considered in Section 3.5.3, under the assumption of a zero interest rate, i.e. $r = 0\%$. Under this lower interest rate level, the magnitude of longevity risk is more severe, as expected: the percentage Risk Margins are higher for all portfolios, increasing in the best-case scenario to 8.07%, up from 6.17%. However, diversification as measured by the $\%RM$ is even more valuable, because the reduction from the initial portfolio to Π^2_{opt} portfolio is almost 4 percentage points. The Diversification Index and portfolio mortality standard deviation, instead, by definition, are not affected by the change in the interest rate, because the weights appearing in (3.31) and (3.44) are expressed in nominal terms (number of annuities written on a generation) rather than in value terms (value of the annuity portfolios on the different generations, for instance). We finally assess the impact of the parameter δ_i

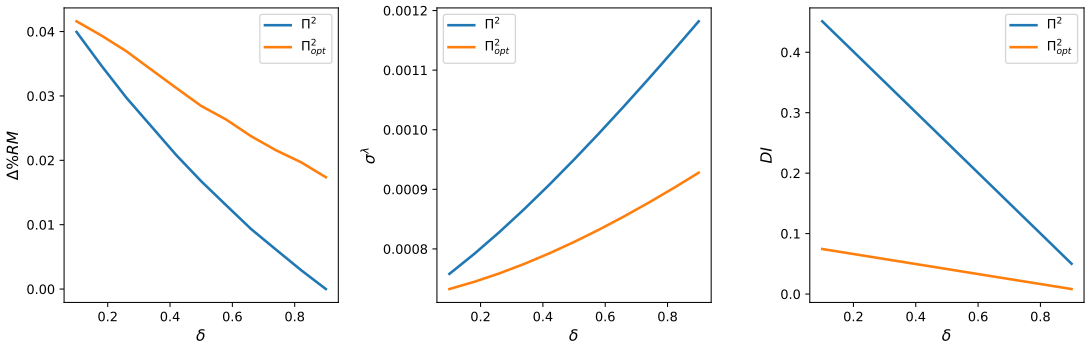


Figure 23: Left Panel: Percentage Risk Margin reduction under different exogenous $\delta_i = \delta$ for every i assumption. Center Panel: Portfolio Mortality Standard Deviation under different exogenous $\delta_i = \delta$ for every i. Right Panel: Diversification under different exogenous $\delta_i = \delta$ for every i.

on portfolio diversification following an expansion. In Section 3.5.3, we considered two countries, the UK and Italy, that belong to the same continent and share many similar features. As a consequence, also their past mortality dynamics were not so dissimilar. We expect, however, that more different countries show way lower similarity, and thus lower δ's between cohort intensities. We perform, then, a simulation study where the parameters of the foreign population are set as in Table 16, with the only exception of the parameters δ_i, which we assume to be a constant δ for every generation i. The interest rate is set to $r = 2\%$, as in our base case. We exogenously set δ to a value that ranges from 0.1 to 0.9. When δ is close to 0, the dynamics of the mortality intensities of the domestic and the foreign populations are orthogonal. Thus, the international expansion targets a foreign population whose mortality dynamics is very different from the domestic one. In this case, we expect the maximum level of diversification gains from an international expansion. As δ increases, the correlation between the mortality intensities of the two populations increases as well. When δ is close to 1, the mortality dynamics of the domestic and foreign population are perfectly correlated. In this last case, we can expect the lowest level of longevity risk diversification gains from an international expansion strategy. We compute, for each level of δ_i, the DI, the portfolio mortality standard deviation and the percentage risk margin reduction, for the portfolios Π^2 and

Π^2_{opt} described in Section 3.5.3.

As expected, the highest values of both the percentage risk margin reduction and the Diversification Index, for both portfolios, are achieved when δ_i is close to 0. The percentage risk margin reduction is 4% in this case, showing that sizable benefits from geographical diversification are possible. Such benefits, measured in terms of either the risk margin reduction or the Diversification Index, decrease as δ approaches 1. The optimal portfolio expansion Π^2_{opt} provides consistently higher risk margin reduction than Π^2, and the gap between the two strategies widens as δ increases (see the left panel of Figure 23). On the contrary, strategy Π^2 shows a higher DI with respect to Π^2_{opt}, for every δ.

The Diversification Index tends to 0 for both portfolios as δ goes to 1. Instead, while the percentage risk margin reduction for Π^2 goes to zero when δ is 1, portfolio expansion Π^2_{opt} offers a diversification benefit relative to the initial portfolio even in that case. This happens because the expansion is targeted in this case to a specific generation. The effects of the international expansion are analogous to those that can be obtained by targeting the domestic expansion to the generation which shows the lowest covariance with the others: Π^2_{opt} reduces the percentage risk margin as much as Π^1_{opt}.

Given their properties, and the evidence from this sensitivity analysis, the Diversification Index and the standard deviation of the portfolio mortality intensity can be an extremely easy-to-handle and useful tools when choosing among competing target foreign populations in an international expansion. The percentage risk margin reduction, being a monetary measure of the diversification gains, is better suited, instead, to select the best strategy when different alternative foreign portfolio compositions can be targeted, once the candidate foreign population has been selected.

3.6 Conclusions

In this paper, we discussed the benefits of geographically diversified portfolios, due to the non-perfect correlation between the dynamics of the mortality rates of different populations. We have considered the problem of an insurer who has to decide whether to

expand his portfolio in the country where it is based or in a foreign country. Some diversification gains can be realized when expanding internationally, due to the mitigation of the exposure to domestic longevity risk. To discuss whether these gains may be sizable in an annuity portfolio, we built a longevity risk model that, while being parsimonious, can capture the non-perfect correlations among the different cohorts of two different populations. We then provided three indicators of the diversification of an international expansion. The percentage risk margin reduction is computed coherently with the Solvency II modeling approach. The standard deviation of portfolio mortality intensity is the volatility of the distribution of a weighted average of the cohort-based mortality intensities, where the weights are the relative contributions of each cohort to the portfolio. The Diversification Index is a weighted average which depends on both the portfolio mix and the weight of the idiosyncratic foreign risk factor. This last measure is a very easy-to-handle indicator which, however may be unreliable when different cohorts compose the domestic portfolio and the foreign one. It is instead very useful when comparing expansions towards the same target portfolios in different populations. Our application, based on an annuity portfolio written on the UK and the Italian populations, shows that the effects of an international diversification are sizable. Expanding internationally decreases the volatility of the portfolio mortality intensity up to 26%. Under a 0% interest rate assumption, we showed that an optimally designed expansion can lower the percentage risk margin, relative to the actuarial value of the portfolio, by almost 4 percentage points. The example in the paper can be considered as conservative, since the two populations of UK and Italy present rather similar historical mortality dynamics. The diversification effect is shown to be more relevant the lower the correlation between intensities.

The diversification benefits of an international expansion may happen to be counterbalanced by the costs connected to the foreign portfolio acquisition process. These costs, that are - say - the fixed costs of opening a foreign affiliate, or the fees required by the agents involved in the M & A operation, etc., may be substantial. As an alternative to a physical expansion, the insurer may obtain the same diversification benefit operating on

the longevity derivatives market. Longevity derivatives, and longevity swaps in particular, are bespoke transactions between (re)insurers and funds or companies, that agree to exchange fixed cash flows and cash flows linked to the survivorship of a particular population (see Blake et al., 2006a for instance). The buyer of the protection provided by a longevity swap transfers the longevity risk linked to a given reference population to the seller, who in turn becomes exposed to such risk. In our case, the insurer can expand internationally by receiving a fixed periodical fee and paying the realized survivorship of the foreign cohorts. Thus, the risk margin reduction benefits of a foreign expansion can be replicated by selling protection through a swap. Even in this case, however, the costs of structuring the agreement and coping with informational asymmetries (Biffis et al., 2016), can substantially reduce the diversification gains. We interpret our results as a possible explanation of the higher degree of internationalization of insurance companies with respect to banks after the adoption of Solvency II. Because of the synthetic possibility to diversify through longevity transfer agreements and longevity swaps, our results may also explain the high number of such contracts recently signed in the marketplace and the attention dedicated to the growth of the market capacity (Blake et al., 2018).

Appendix 3.A Gaussian Mapping Covariance

A simple application of Itô's Lemma allows us to show that the solution to the SDE (3.47) is given by:

$$\lambda_i^V(t) = \lambda_i^V(0)e^{b_i t} + \frac{a_i}{b_i}(1 - e^{b_i t}) + \sigma_i^V \int_0^t e^{b_i(t-s)}\, dW_i(s). \tag{3.60}$$

Therefore, we have that:

$$\mathbb{E}_0[\lambda_i^V(t)] = \lambda_i^V(0)e^{b_i t} + \frac{a_i}{b_i}(1 - e^{b_i t}) \tag{3.61}$$

$$Var_0[\lambda_i^V(t)] = \frac{(\sigma_i^V)^2}{2b_i}\sqrt{e^{2b_i t} - 1}. \tag{3.62}$$

Since $\lambda_i^V(t) - \mathbb{E}_0[\lambda_i^V(t)] = \sigma_i^V \int_0^t e^{b_i(t-s)}\, dW_i(s)$, the covariance between $\lambda_i^V(t)$ and $\lambda_j^V(t)$ is:

$$\begin{aligned} Cov_0(\lambda_i^V(t), \lambda_j^V(t)) &= \mathbb{E}_0\left[\sigma_i^V \sigma_j^V \Big(\int_0^t e^{b_i(t-s)}\, dW_i(s)\Big)\Big(\int_0^t e^{b_j(t-s)}\, dW_j(s)\Big)\right] \\ &= \mathbb{E}_0\left[\sigma_i^V \sigma_j^V \rho_{ij} \int_0^t e^{(b_i+b_j)(t-s)}\, ds\right] \\ &= \sigma_i^V \sigma_j^V \rho_{ij} \int_0^t e^{(b_i+b_j)(t-s)}\, ds \\ &= \frac{\sigma_i^V \sigma_j^V \rho_{ij}}{b_i + b_j}\Big(e^{(b_i+b_j)t} - 1\Big). \end{aligned} \tag{3.63}$$

Finally, we have:

$$\begin{aligned} Corr_0(\lambda_i^V(t), \lambda_j^V(t)) &= \frac{Cov_0(\lambda_i^V(t), \lambda_j^V(t))}{\sqrt{Var_0[\lambda_i^V(t)] Var_0[\lambda_j^V(t)]}} \\ &= \frac{2\rho_{ij}}{b_i + b_j} \cdot \frac{e^{(b_i+b_j)t} - 1}{\sqrt{\frac{(e^{2b_i t}-1)(e^{2b_j t}-1)}{b_i b_j}}} \end{aligned} \tag{3.64}$$

Thanks to the Gaussian Mapping technique we can also compute the conditional correlation between two generations belonging to two different populations. Considering $0 \le u \le t$, the conditional correlation between $\lambda_{x_i}^d(t)$ and $\lambda_{x_j}^f(t)$ is given by:

$$Corr_u[\lambda_{x_i}^d(t), \lambda_{x_j}^f(t)] = \delta_j \frac{Cov_u(\lambda_{x_i}^d(t), \lambda_{x_j}^d(t))}{\sqrt{Var_u(\lambda_{x_i}^d(t)) \cdot Var_u(\lambda_{x_j}^f(t))}}, \tag{3.65}$$

where $Cov_u(\lambda_{x_i}^d(t), \lambda_{x_j}^d(t))$ is computed using the Gaussian mapping technique, and

$$Var_u(\lambda_{x_j}^f(t)) = \delta_j^2 Var_u(\lambda_{x_j}^d(t)) + (1-\delta_j)^2 Var_u(\lambda'(t; x_j)). \tag{3.66}$$

Table 17: Instantaneous correlation matrix UK population with upper and lower 95% confidence bounds from boostrapping.

	65	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**
UB		(0.99938)	(0.999000)	(0.998716)	(0.998124)	(0.997751)	(0.997741)	(0.997332)	(0.997197)	(0.99709)	(0.996878)
65	1	0.996614	0.993845	0.992351	0.990865	0.989791	0.988837	0.987733	0.987419	0.987183	0.986708
LB		(0.9909)	(0.983627)	(0.980001)	(0.977076)	(0.974617)	(0.97186)	(0.969869)	(0.968956)	(0.968354)	(0.966995)
UB	(0.999378)		(0.999558)	(0.999358)	(0.999023)	(0.99862)	(0.998523)	(0.998186)	(0.998116)	(0.998188)	(0.997826)
66	0.996614	1	0.999089	0.998491	0.997805	0.99704	0.996548	0.995866	0.995682	0.995693	0.995174
LB	(0.991186)		(0.998287)	(0.997043)	(0.995688)	(0.994296)	(0.993232)	(0.991792)	(0.991525)	(0.991261)	(0.990592)
UB	(0.999017)	(0.999559)		(0.999578)	(0.999456)	(0.999219)	(0.999085)	(0.998907)	(0.998915)	(0.998813)	(0.998598)
67	0.993845	0.999089	1	0.999317	0.999113	0.998707	0.998423	0.998021	0.997997	0.997849	0.997468
LB	(0.983628)	(0.998268)		(0.999077)	(0.998673)	(0.997955)	(0.997325)	(0.996572)	(0.996455)	(0.996298)	(0.995615)
UB	(0.998679)	(0.999353)	(0.99958)		(0.99971)	(0.999449)	(0.999293)	(0.999087)	(0.999063)	(0.999027)	(0.998571)
68	0.992351	0.998491	0.999317	1	0.999553	0.999168	0.998931	0.998549	0.99852	0.998456	0.997807
LB	(0.979943)	(0.997032)	(0.999073)		(0.999417)	(0.998807)	(0.998385)	(0.997821)	(0.997697)	(0.997582)	(0.99674)
UB	(0.998116)	(0.99904)	(0.999447)	(0.99971)		(0.999733)	(0.999568)	(0.999369)	(0.999376)	(0.999438)	(0.9989)
69	0.990865	0.997805	0.999113	0.999553	1	0.999569	0.999189	0.999009	0.999047	0.999114	0.998333
LB	(0.97749)	(0.995701)	(0.998669)	(0.999414)		(0.999343)	(0.998844)	(0.998593)	(0.998596)	(0.998586)	(0.997635)
UB	(0.997754)	(0.998619)	(0.999221)	(0.999444)	(0.999732)		(0.999625)	(0.99954)	(0.999555)	(0.999654)	(0.999236)
70	0.989791	0.99704	0.998707	0.999168	0.999569	1	0.999263	0.99927	0.999281	0.999448	0.998791
LB	(0.974657)	(0.99433)	(0.998)	(0.998816)	(0.999344)		(0.998901)	(0.998917)	(0.998931)	(0.999094)	(0.998258)
UB	(0.997714)	(0.998512)	(0.999078)	(0.999293)	(0.999566)	(0.999631)		(0.999724)	(0.999777)	(0.999767)	(0.999632)
71	0.988837	0.996548	0.998423	0.998931	0.999189	0.999263	1	0.999609	0.999689	0.999442	0.999444
LB	(0.972173)	(0.993218)	(0.997347)	(0.998366)	(0.998853)	(0.998884)		(0.999512)	(0.999578)	(0.999165)	(0.999215)
UB	(0.997296)	(0.998214)	(0.998897)	(0.999096)	(0.999366)	(0.999541)	(0.9997259		(0.999917)	(0.999875)	(0.999819)
72	0.969828	0.991711	0.996482	0.997807	0.99859	0.998918	0.999511	1	0.999779	0.999616	0.999536
LB		(0.9909)	(0.983627)	(0.980001)	(0.977076)	(0.974617)	(0.97186)	(0.969869)	(0.968956)	(0.968354)	(0.966995)
UB	(0.997148)	(0.998104)	(0.99892)	(0.999065)	(0.999377)	(0.999555)	(0.999777)	(0.999919)		(0.9999)	(0.999822)
73	0.987419	0.995682	0.997997	0.99852	0.999047	0.999281	0.999689	0.999862	1	0.999732	0.999698
LB	(0.969095)	(0.991471)	(0.996389)	(0.997707)	(0.998618)	(0.998915)	(0.999575)	(0.99978)		(0.999586)	(0.99955)
UB	(0.997123)	(0.998201)	(0.998815)	(0.999026)	(0.999433)	(0.999655)	(0.999763)	(0.999876)	(0.999898)		(0.999815)
74	0.987183	0.995693	0.997849	0.998456	0.999114	0.999448	0.999442	0.999746	0.999732	1	0.999547
LB	(0.968577)	(0.991343)	(0.996281)	(0.997581)	(0.998579)	(0.999085)	(0.999171)	(0.999618)	(0.999585)		(0.99928)
UB	(0.996825)	(0.997807)	(0.998599)	(0.998586)	(0.998896)	(0.999231)	(0.999633)	(0.999817)	(0.999819)	(0.999811)	
75	0.986708	0.995174	0.997468	0.997807	0.998333	0.998791	0.999444	0.999701	0.999698	0.999547	1
LB	(0.968182)	(0.99056)	(0.995587)	(0.9967)	(0.997637)	(0.998261)	(0.999215)	(0.999532)	(0.999552)	(0.999285)	

Table 18: Correlation between populations with upper and lower 95% confidence bounds from boostrapping. Rows are UK generations while columns are ITA generations.

	65	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**
UB	(0.958536)	(0.979156)	(0.980102)	(0.97169)	(0.972043)	(0.97753)	(0.976328)	(0.978398)	(0.970923)	(0.977916)	(0.963227)
65	0.958536	0.976454	0.97503	0.965521	0.964995	0.969704	0.967647	0.968971	0.961354	0.968185	0.953476
LB	(0.958536)	(0.971023)	(0.965148)	(0.953324)	(0.951457)	(0.955025)	(0.950993)	(0.951074)	(0.943368)	(0.950217)	(0.934664)
UB	(0.958504)	(0.980357)	(0.981229)	(0.972916)	(0.973531)	(0.978934)	(0.977709)	(0.979876)	(0.972373)	(0.979556)	(0.964782)
66	0.955861	0.980357	0.980761	0.972075	0.972336	0.97739	0.975777	0.977536	0.969978	0.977114	0.962233
LB	(0.950595)	(0.980357)	(0.979968)	(0.970672)	(0.970352)	(0.974748)	(0.972428)	(0.973557)	(0.965942)	(0.972905)	(0.957943)
UB	(0.958127)	(0.9799)	(0.98163)	(0.973106)	(0.973914)	(0.979498)	(0.978242)	(0.980498)	(0.973106)	(0.980139)	(0.965516)
67	0.953179	0.979438	0.98163	0.972854	0.973585	0.978999	0.977588	0.979627	0.972208	0.979203	0.964426
LB	(0.943524)	(0.978657)	(0.98163)	(0.972614)	(0.97315)	(0.978292)	(0.976524)	(0.97819)	(0.970704)	(0.977606)	(0.962632)
UB	(0.957858)	(0.979697)	(0.98122)	(0.973526)	(0.974171)	(0.979732)	(0.978443)	(0.980683)	(0.973254)	(0.980351)	(0.965512)
68	0.951754	0.978859	0.980966	0.973526	0.97402	0.979458	0.978092	0.980151	0.972725	0.979807	0.964761
LB	(0.940025)	(0.977399)	(0.980725)	(0.973526)	(0.97388)	(0.979109)	(0.977554)	(0.979404)	(0.971887)	(0.978966)	(0.963694)
UB	(0.957178)	(0.979304)	(0.981013)	(0.973156)	(0.974367)	(0.979922)	(0.978629)	(0.980861)	(0.973472)	(0.980681)	(0.965729)
69	0.950242	0.978098	0.980677	0.973002	0.974367	0.979762	0.978255	0.980513	0.97315	0.980363	0.965181
LB	(0.936819)	(0.976141)	(0.980238)	(0.972864)	(0.974367)	(0.979541)	(0.977932)	(0.980107)	(0.972719)	(0.979836)	(0.964518)
UB	(0.956916)	(0.978959)	(0.980822)	(0.972937)	(0.974142)	(0.980224)	(0.978725)	(0.981078)	(0.973685)	(0.980934)	(0.966091)
70	0.949249	0.977386	0.980317	0.972666	0.973986	0.980224	0.978367	0.98081	0.973417	0.980729	0.965663
LB	(0.934764)	(0.974749)	(0.979582)	(0.972317)	(0.973764)	(0.980224)	(0.978001)	(0.980467)	(0.973071)	(0.980377)	(0.965158)
UB	(0.956931)	(0.978951)	(0.980825)	(0.972915)	(0.974108)	(0.979987)	(0.979213)	(0.981381)	(0.974022)	(0.981165)	(0.966604)
71	0.948454	0.977027	0.980163	0.972558	0.973739	0.979625	0.979213	0.981268	0.973937	0.980846	0.966417
LB	(0.932229)	(0.973718)	(0.979085)	(0.972032)	(0.973406)	(0.979262)	(0.979213)	(0.981174)	(0.973829)	(0.980572)	(0.966199)
UB	(0.956826)	(0.978961)	(0.980921)	(0.973006)	(0.97419)	(0.980186)	(0.979232)	(0.981944)	(0.974446)	(0.981556)	(0.967063)
72	0.947671	0.976645	0.980056	0.972471	0.973848	0.97992	0.979119	0.981944	0.974391	0.981429	0.96695
LB	(0.930941)	(0.972758)	(0.978568)	(0.971745)	(0.973446)	(0.979576)	(0.979026)	(0.981944)	(0.97431)	(0.9813)	(0.966792)
UB	(0.956183)	(0.978327)	(0.980331)	(0.972373)	(0.973617)	(0.979602)	(0.978686)	(0.981261)	(0.973933)	(0.980987)	(0.966475)
73	0.946796	0.975871	0.979435	0.971852	0.973294	0.979335	0.978599	0.981204	0.973933	0.980824	0.966357
LB	(0.928455)	(0.971803)	(0.977957)	(0.971019)	(0.97287)	(0.978985)	(0.978489)	(0.981125)	(0.973933)	(0.980677)	(0.966211)
UB	(0.955689)	(0.977943)	(0.979815)	(0.971933)	(0.973264)	(0.979293)	(0.978265)	(0.980801)	(0.973428)	(0.980679)	(0.96606)
74	0.946176	0.975475	0.97888	0.971384	0.972952	0.979088	0.977947	0.980676	0.973266	0.980679	0.965806
LB	(0.928609)	(0.971309)	(0.977319)	(0.970554)	(0.972423)	(0.978737)	(0.977676)	(0.980553)	(0.973122)	(0.980679)	(0.965556)
UB	(0.955982)	(0.978132)	(0.980179)	(0.972063)	(0.973293)	(0.979438)	(0.978696)	(0.981318)	(0.973913)	(0.981055)	(0.966801)
75	0.946262	0.975526	0.97907	0.971311	0.97275	0.979008	0.978514	0.981203	0.973792	0.980795	0.966801
LB	(0.927589)	(0.971166)	(0.977298)	(0.970242)	(0.972053)	(0.978498)	(0.978285)	(0.981041)	(0.97364)	(0.980536)	(0.966801)

Table 19: Domestic portfolio composition.

	65	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**
% in the national population	9.68%	9.79%	9.98%	10.34%	11.00%	10.10%	10.10%	8.43%	8.10%	7.36%	6.56%
Π^0	97	98	100	103	110	101	86	84	81	74	66

Table 20: Foreign portfolio composition.

	65	**66**	**67**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**
% in the national population	10.31%	10.73%	10.48%	10.39%	8.01%	8.18%	8.26%	8.09%	8.09%	8.86%	8.60%
Π^F	103	107	105	104	80	82	83	81	81	89	86

Table 21: Effects of geographical diversification ($r = 2\%$)

Portfolio	AV	RM	Π	$\%RM$	$\sigma^\lambda(\Pi)$	DI
Π^0	$1.4104 \cdot 10^4$	$1.1838 \cdot 10^3$	$1.5288 \cdot 10^4$	8.39%	0.00124	-
Π^F	$1.4768 \cdot 10^4$	$1.0912 \cdot 10^3$	$1.5586 \cdot 10^4$	7.39%	0.00107	-
Π^1	$2.8208 \cdot 10^4$	$2.3676 \cdot 10^3$	$3.0576 \cdot 10^4$	8.39%	0.00124	0
Π^2	$2.8872 \cdot 10^4$	$2.2749 \cdot 10^3$	$3.1147 \cdot 10^4$	7.87%	0.00115	0.0746
Π^3	$4.3631 \cdot 10^4$	$3.3646 \cdot 10^3$	$4.7005 \cdot 10^4$	7.71%	0.00112	0.0992
Π^1_{opt}	$3.0187 \cdot 10^4$	$1.9950 \cdot 10^3$	$3.2182 \cdot 10^4$	6.61%	0.00096	0
Π^2_{opt}	$3.0790 \cdot 10^4$	$1.8998 \cdot 10^3$	$3.2690 \cdot 10^4$	6.17%	0.00091	0.0121

Table 22: Effects of geographical diversification ($r = 0\%$)

Portfolio	AV	RM	Π	$\%RM$	$\sigma^\lambda(\Pi)$	DI
Π^0	$1.7656 \cdot 10^4$	$1.9408 \cdot 10^3$	$1.9596 \cdot 10^4$	10.99%	0.00124	-
Π^F	$1.8614 \cdot 10^4$	$1.8109 \cdot 10^3$	$2.0425 \cdot 10^4$	9.72%	0.00107	-
Π^1	$3.5311 \cdot 10^4$	$3.8815 \cdot 10^3$	$3.9193 \cdot 10^4$	10.99%	0.00124	0
Π^2	$3.6270 \cdot 10^4$	$3.7701 \cdot 10^3$	$4.0040 \cdot 10^4$	10.39%	0.00115	0.0746
Π^3	$5.4885 \cdot 10^4$	$5.5786 \cdot 10^3$	$6.0464 \cdot 10^4$	10.16%	0.00112	0.0992
Π^1_{opt}	$3.8182 \cdot 10^4$	$3.3155 \cdot 10^3$	$4.1497 \cdot 10^4$	8.68%	0.00096	0
Π^2_{opt}	$3.9079 \cdot 10^4$	$3.1565 \cdot 10^3$	$4.2235 \cdot 10^4$	8.07%	0.00091	0.0121

List of Figures

List of Tables

Bibliography

Gara Afonso, Marco Cipriani, Adam Copeland, Anna Kovner, Gabriele La Spada, and Antoine Martin. The Market Events of Mid-September 2019. Staff Reports 918, Federal Reserve Bank of New York, March 2020. URL `https://ideas.repec.org/p/fip/fednsr/87585.html`.

Giuseppe Arbia, Anil K Bera, Osman Doğan, and Süleyman Taşpınar. Testing impact measures in spatial autoregressive models. *International Regional Science Review*, 1 2019. ISSN 0160-0176. doi: 10.1177/0160017619826264.

Manuel Arellano and Olympia Bover. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51, July 1995. URL `https://ideas.repec.org/a/eee/econom/v68y1995i1p29-51.html`.

Enrico Biffis, David Blake, Lorenzo Pitotti, and Ariel Sun. The cost of counterparty risk and collateralization in longevity swaps. *Journal of Risk and Insurance*, 83(2): 387–419, 2016.

Craig Blackburn and Michael Sherris. Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics*, 53(1):64–73, 2013.

David Blake and William Burrows. Survivor bonds: Helping to hedge mortality risk. *Journal of Risk and Insurance*, pages 339–348, 2001.

David Blake, Andrew Cairns, Kevin Dowd, and Richard MacMinn. Longevity bonds:

financial engineering, valuation, and hedging. *Journal of Risk and Insurance*, 73(4): 647–672, 2006a.

David Blake, Andrew JG Cairns, and Kevin Dowd. Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, 12(1):153–197, 2006b.

David Blake, Andrew John George Cairns, Kevin Dowd, and Amy R Kessler. Still living with mortality: The longevity risk transfer market after one decade. *British Actuarial Journal*, 24, 2018.

Steffanie Brady, Kenechukwu Anadu, and Nathaniel Cooper. The stability of prime money market mutual funds: sponsor support from 2007 to 2011. *Available at SSRN 3015986*, 2012.

Yann Bramoullé, Habiba Djebbari, and Bernard Fortin. Identification of peer effects through social networks. *Journal of econometrics*, 150(1):41–55, 2009.

Damiano Brigo and Fabio Mercurio. *Interest rate models : theory and practice*. Springer finance. Springer, Berlin, Heidelberg, Paris, 2001. ISBN 3-540-41772-9. URL `http://opac.inria.fr/record=b1097778`.

A.J.G. Cairns, D. Blake, and K. Dowd. Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, (2):79–113, 2008. ISSN 0346-1238.

Andrew JG Cairns. Robust hedging of longevity risk. *Journal of Risk and Insurance*, 80(3):621–648, 2013.

Andrew JG Cairns, David Blake, and Kevin Dowd. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718, 2006.

Sergey Chernenko and Adi Sunderam. Frictions in Shadow Banking: Evidence from the Lending Behavior of Money Market Mutual Funds. *The Review of Financial*

Studies, 27(6):1717–1750, 04 2014. ISSN 0893-9454. doi: 10.1093/rfs/hhu025. URL https://doi.org/10.1093/rfs/hhu025.

Sergey Chernenko and Adi Sunderam. Do fire sales create externalities? *Journal of Financial Economics*, 135(3):602–628, 2020.

Joshua Coval and Erik Stafford. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, 86(2):479–512, 2007.

John C Cox, Ingersoll Jonathan E, and Stephen Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985a. URL https://EconPapers.repec.org/RePEc:ecm:emetrp:v:53:y:1985:i:2:p:385-407.

John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. An intertemporal general equilibrium model of asset prices. *Econometrica: Journal of the Econometric Society*, pages 363–384, 1985b.

J. David Cummins and Mary A. Weiss. Convergence of insurance and financial markets: Hybrid and securitized risk-transfer solutions. *The Journal of Risk and Insurance*, 76 (3):493–545, 2009. ISSN 00224367, 15396975. URL http://www.jstor.org/stable/40247567.

M. Dahl. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics*, 35(1):113–136, 2004.

Mikkel Dahl, Martin Melchior, and Thomas Møller. On systematic mortality risk and risk-minimization with survivor swaps. *Scandinavian Actuarial Journal*, (2-3):114–146, 2008.

Clemente De Rosa, Elisa Luciano, and Luca Regis. Annuity portfolios: risk margin and longevity and interest rate hedging. 2016. URL https://globalriskinstitute.org/wp-content/uploads/2017/02/hedging_longevity_interest_rate_14_4.pdf.

Clemente De Rosa, Elisa Luciano, and Luca Regis. Basis risk in static versus dynamic longevity-risk hedging. *Scandinavian Actuarial Journal*, 2017(4):343–365, 2017.

Clemente De Rosa, Elisa Luciano, and Luca Regis. Geographical diversification and longevity risk mitigation in annuity portfolios. *Carlo Alberto Notebooks*, (546), 2019. URL https://www.carloalberto.org/wp-content/uploads/2017/12/no.546.pdf.

Clemente De Rosa, Gabriele La Spada, and Fabrizio Lillo. Peer effects and spillovers in money market mutual funds flow. *Working paper*, 2020.

Kevin Dowd, David Blake, Andrew JG Cairns, and Paul Dawson. Survivor swaps. *Journal of Risk and Insurance*, 73(1):1–17, 2006.

John Driscoll and Aart Kraay. Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4):549–560, 1998. URL https://EconPapers.repec.org/RePEc:tpr:restat:v:80:y:1998:i:4:p:549-560.

J.P. Elhorst. *Spatial econometrics: from cross-sectional data to spatial panels*. Springer, 2014. ISBN 9783642403408.

Andrew Ellul, Chotibhak Jotikasthira, and Christian T. Lundblad. Regulatory pressure and fire sales in the corporate bond market. *Journal of Financial Economics*, 101(3):596 – 620, 2011. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2011.03.020. URL http://www.sciencedirect.com/science/article/pii/S0304405X11000857.

Paul Erdös et al. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.

Antonio Falato, Ali Hortacsu, Dan Li, and Chaehee Shin. Fire-sale spillovers in debt markets. 2016.

Peter Feldhütter. The Same Bond at Different Prices: Identifying Search Frictions and Selling Pressures. *The Review of Financial Studies*, 25(4):1155–1206, 04 2012. ISSN 0893-9454. doi: 10.1093/rfs/hhr093. URL https://doi.org/10.1093/rfs/hhr093.

Dario Focarelli and Alberto Franco Pozzolo. Cross-border m&as in the financial sector: Is banking different from insurance? *Journal of Banking & Finance*, 32(1):15–29, 2008.

Man Chung Fung, Katja Ignatieva, and Michael Sherris. Systematic mortality risk: An analysis of guaranteed lifetime withdrawal benefits in variable annuities. *Insurance: Mathematics and Economics*, 58:103 – 115, 2014. ISSN 0167-6687. doi: http://dx.doi.org/10.1016/j.insmatheco.2014.06.010. URL `http://www.sciencedirect.com/science/article/pii/S0167668714000821`.

Robin Greenwood and David Thesmar. Stock price fragility. *Journal of Financial Economics*, 102(3):471–490, 2011. URL `https://EconPapers.repec.org/RePEc:eee:jfinec:v:102:y:2011:i:3:p:471-490`.

Samuel G Hanson, David S Scharfstein, and Adi Sunderam. An evaluation of money market fund reform proposals. *IMF Economic Review*, 63(4):984–1023, 2015. doi: 10.1057/imfer.2015.14. URL `https://doi.org/10.1057/imfer.2015.14`.

International Monetary Fund (IMF). Global Financial Stability Report. (IMF April): 123–154, 2012.

R.A. Jarrow and S.M. Turnbull. Delta, gamma and bucket hedging of interest rate derivatives. *Applied Mathematical Finance*, 1:21–48, 1994. ISSN 0012-9682.

Petar Jevtić and Luca Regis. A continuous-time stochastic model for the mortality surface of multiple populations. *Insurance: Mathematics and Economics*, 88:181–195, 2019.

Petar Jevtić, Elisa Luciano, and Elena Vigna. Mortality surface by means of continuous time cohort models. *Insurance: Mathematics and Economics*, 53(1):122–133, 2013.

Petar Jevtić and Luca Regis. Assessing the solvency of insurance portfolios via a continuous-time cohort model. *Insurance: Mathematics and Economics*, 61:36 – 47,

2015. ISSN 0167-6687. doi: https://doi.org/10.1016/j.insmatheco.2014.12.002. URL http://www.sciencedirect.com/science/article/pii/S0167668714001632.

Marcin Kacperczyk and Philipp Schnabl. How safe are money market funds? *The Quarterly Journal of Economics*, 128(3):1073–1122, 2013.

Harry H. Kelejian and Ingmar R. Prucha. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1):99–121, Jul 1998. ISSN 1573-045X. doi: 10.1023/A:1007707430416. URL https://doi.org/10.1023/A:1007707430416.

Gabriele La Spada. Competition, reach for yield, and money market funds. *Journal of Financial Economics*, 129(1):87 – 110, 2018. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2018.04.006. URL http://www.sciencedirect.com/science/article/pii/S0304405X18301004.

Lung-fei Lee and Jihai Yu. A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory*, 26(2):564–597, 2010.

Lung-fei Lee and Jihai Yu. Efficient gmm estimation of spatial dynamic panel data models with fixed effects. *Journal of Econometrics*, 180(2):174 – 197, 2014. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2014.03.003. URL http://www.sciencedirect.com/science/article/pii/S0304407614000438.

Lung-fei Lee. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews*, 22(4):307–335, 2003. doi: 10.1081/ETC-120025891. URL https://doi.org/10.1081/ETC-120025891.

Ronald D Lee and Lawrence R Carter. Modeling and forecasting us mortality. *Journal of the American statistical association*, 87(419):659–671, 1992.

J. LeSage and R.K. Pace. *Introduction to Spatial Econometrics*. Statistics: A Series of

Textbooks and Monographs. CRC Press, 2009. ISBN 9781420064254. URL https://books.google.it/books?id=EKiKXcgL-D4C.

E. Luciano and E. Vigna. Mortality risk via affine stochastic intensities: calibration and empirical relevance. *Belgian Actuarial Bulletin*, 8(1):5–16, 2008.

Elisa Luciano, Luca Regis, and Elena Vigna. Delta and gamma hedging of mortality and interest-rate risk. *Insurance: Mathematics and Economics*, (50):402–412, 2012a.

Elisa Luciano, Luca Regis, and Elena Vigna. Delta–gamma hedging of mortality and interest rate risk. *Insurance: Mathematics and Economics*, 50(3):402–412, 2012b.

Charles F Manski. Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542, 1993.

Patrick E. McCabe. The cross section of money market fund risks and financial crises. Technical report, 2010.

Moshe A Milevsky and S David Promislow. Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29(3):299–318, 2001.

Andrew Ngai and Michael Sherris. Longevity risk management for life and variable annuities: The effectiveness of static hedging using longevity bonds and derivatives. *Insurance: Mathematics and Economics*, (49):100–114, 2011.

Vincenzo Nicosia, John Tang, Cecilia Mascolo, Mirco Musolesi, Giovanni Russo, and Vito Latora. Graph metrics for temporal networks. In *Temporal networks*, pages 15–40. Springer, 2013.

Arthur E Renshaw and Steven Haberman. A cohort-based extension to the lee–carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38(3): 556–570, 2006.

Bruce Sacerdote. Peer Effects with Random Assignment: Results for Dartmouth Roommates*. *The Quarterly Journal of Economics*, 116(2):681–704, 05 2001. ISSN

0033-5533. doi: 10.1162/00335530151144131. URL https://doi.org/10.1162/00335530151144131.

Lawrence Schmidt, Allan Timmermann, and Russ Wermers. Runs on money market mutual funds. *American Economic Review*, 106(9):2625–57, September 2016. doi: 10.1257/aer.20140678. URL http://www.aeaweb.org/articles?id=10.1257/aer.20140678.

Dirk Schoenmaker and Jan Sass. Cross-border insurance in europe: Challenges for supervision. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 41(3): 351–377, 2016.

Michael Sherris, Yajing Xu, and Jonathan Ziveyi. Cohort and value-based multi-country longevity risk management. *Working Paper Cepar 2018/1*, 2018.

RSP Stevens, AMB De Waegenaere, B Melenberg, et al. Longevity risk and natural hedge potential in portfolios of life insurance products: The effect of investment risk. Technical report, 2011.

Philip E. Strahan and BaÅŸak Tanyeri. Once burned, twice shy: Money market fund responses to a systemic liquidity shock. *Journal of Financial and Quantitative Analysis*, 50(1-2):119–144, 2015. URL https://EconPapers.repec.org/RePEc:cup:jfinqa:v:50:y:2015:i:1-2:p:119-144_00.

John Tang, Salvatore Scellato, Mirco Musolesi, Cecilia Mascolo, and Vito Latora. Small-world behavior in time-varying graphs. *Physical Review E*, 81(5):055101, 2010.

Jeffrey T. Tsai, Larry Y. Tzeng, and Jennifer L. Wang. Hedging longevity risk when interest rates are uncertain. *North American Actuarial Journal*, 15(2):201–211, 2011. doi: 10.1080/10920277.2011.10597617. URL https://doi.org/10.1080/10920277.2011.10597617.

O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, (5):177–188, 1977.