HAL
open science

Collective destabilising phenomena in socio-economic systems

João Da Gama Batista

To cite this version:

João Da Gama Batista. Collective destabilising phenomena in socio-economic systems. Other. Université Paris Saclay (COmUE), 2015. English. NNT : 2015SACLC026 . tel-01307138

HAL Id: tel-01307138

https://theses.hal.science/tel-01307138

Submitted on 26 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT: 2015SACLC026

THÈSE DE DOCTORAT

DE

L'UNIVERSITÉ PARIS-SACLAY

PRÉPARÉE À

CENTRALESUPÉLEC

ÉCOLE DOCTORALE N° 573
INTERFACES : APPROCHES INTERDISCIPLINAIRES / FONDEMENTS, APPLICATIONS ET INNOVATION

MATHÉMATIQUES APPLIQUÉES

PAR

João da Gama Batista

Phénomènes collectifs déstabilisateurs dans les systèmes socio-économiques

THÈSE PRÉSENTÉE ET SOUTENUE À CHÂTENAY-MALABRY, LE 12 OCTOBRE 2015.

COMPOSITION DU JURY :

ALAN KIRMAN	UNIVERSITÉ D'AIX-MARSEILLE III	PRÉSIDENT RAPPORTEUR EXAMINATEUR
CEES DIKS	UNIVERSITY OF AMSTERDAM	RAPPORTEUR EXAMINATEUR
DAMIEN CHALLET	CENTRALESUPÉLEC	EXAMINATEUR DIRECTEUR DE THÈSE
JEAN-PHILIPPE BOUCHAUD	CAPITAL FUND MANAGEMENT	CO-DIRECTEUR DE THÈSE
FRÉDÉRIC ABERGEL	CENTRALESUPÉLEC	CO-DIRECTEUR DE THÈSE

ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

Abstract

Résumé en français

Cette thèse aborde deux sujets de recherche dans le cadre du même projet. La première voie de recherche, expliquée en détail au chapitre 3, est une approche de modélisation relative à la dynamique de confiance dans une société en réseau. La seconde voie de recherche, décrite au chapitre 4, est une approche expérimentale visant à étudier les décisions humaines lors de l'échange d'un actif avec une croissance moyenne positive par période dans un environnement de laboratoire contrôlé.

Un des liens communs entre ces deux thèmes est l'action collective, qui joue un rôle déterminant dans de nombreux phénomènes, par exemple la dynamique de la panique, les faillites et par conséquent le risque systémique. C'est pourquoi, j'espère que ce travail contribuera à l'étude des phénomènes d'actions collectives, en particulier dans la finance quantitative, où les conclusions spécifiques du modèle de confiance et l'expérience de trading en laboratoire mentionés ci-dessus pourront être utilisées dans leur état actuel.

Mots-clés : réseaux, complexité, confiance, laboratoire, trading.

Abstract in English

This thesis reports on two different research topics belonging to the same project. The first research avenue, which is thoroughly explained in chapter 3, is a modelling approach to the dynamics of trust in a networked society. The second, whose description can be found in chapter 4, is an experimental approach to study human decisions when people trade an asset with a positive average growth per period in a controlled laboratory environment.

One of the common links between these two topics is collective action, which is a key player in a number of phenomena, for example in the dynamics of panic, bankruptcies and, consequently, systemic risk. Therefore, the author hopes that this work will contribute to the study of collective action phenomena, especially in the field of quantitative finance, in which it is more likely that the specific findings from the above mentioned trust model and trading experiment can be used in their present form.

Keywords: networks, complexity, trust, experiment, trading.

Contents

Abstract 2

Acknowledgments 5

1 Introduction en français 7
- 1.1 Une nouvelle réponse . . . 7
- 1.2 Approche économique standard . . . 8
- 1.3 Etat de l'art . . . 10
- 1.4 Phénomènes collectifs déstabilisateurs . . . 14

2 Introduction 18
- 2.1 A new answer . . . 18
- 2.2 Standard economics approach . . . 19
- 2.3 State of the art . . . 21
- 2.4 Collective destabilizing phenomena . . . 25

3 Sudden trust collapse in networked societies 28
- 3.1 Introduction . . . 28
- 3.2 The model . . . 32
- 3.3 Numerical results . . . 37
- 3.4 A mean-field analysis . . . 39
- 3.5 Conclusion . . . 51

4 Do investors trade too much? A laboratory experiment. 53
- 4.1 Introduction and aim of the study . . . 54
- 4.2 Experiment motivation and set-up . . . 56
- 4.3 Rational Benchmark . . . 61
- 4.4 Results . . . 65
- 4.5 Conclusion . . . 89

5 Conclusion and future work 92

Appendix A Sudden trust collapse in networked societies 95
A.1 Model specifications 95
A.2 Panic factor d and stability 97

Appendix B Detrimental collective effects in trading experiments 100
B.1 Written laboratory materials 100

References 123

Listing of figures

3.1 TED spread, LIBOR and T-bill (2006-2009) 29
3.2 Average density of links for $N = 200$ and varying d and β 39
3.3 Hysteresis path along $d = 2\beta$ for $N = 50, 100, 200$ 40
3.4 Hysteresis path along μ for $N = 50, 100, 200$ 41
3.5 $\tilde{h}$ and $\tilde{L}$ (single run) for $N = 100$, $d = 3.9$ and $\beta = 0$ 41
3.6 Diagram of mean-field solutions for $d = 0$ and $d = 3$ and $\phi = 2$, $r = 1$ . . . 49

4.1 Utility function for bounded rationality 63
4.2 Critical values W^* for bounded rationality 64
4.3 Average price time series 66
4.4 Market positions time series 67
4.5 Activity rate and final wealth 68
4.6 Average wealth time series 69
4.7 Average wealth, cash and number of shares 70
4.8 Average skewness of price log returns 72
4.9 Two-by-two agent synchronisation for all orders 74
4.10 Two-by-two agent synchronisation for buying orders and for selling orders . 74
4.11 Clusters using FDR applied to market positions 76
4.12 Average trading activity and final wealth for each cluster 77
4.13 Risk aversion ECDF 79
4.14 Risk aversion probability density estimation 80
4.15 Average final wealth and activity rate versus risk aversion 80
4.16 Linear regressions of average final wealth and activity rate on risk aversion . 81
4.17 ECDF of price log returns predictions 84
4.18 Conditional densities of average return predictions ω_0 87
4.19 Conditional densities of average return predictions ω_1 88

A.1 Panic factor and stability, parametric representation of eigenvalues 99

Acknowledgments

This thesis would not have been possible without the contribution of a number of people, who have, directly or indirectly, lent their hand in this project.

First and foremost, I thank Jean-Philippe Bouchaud and Damien Challet, who provided invaluable help and guidance during my entire PhD at École Centrale Paris. I am particularly grateful for their incredible availability and enthusiasm throughout the projects we worked on together. They have always been very supportive and I could not possibly overstress how thankful I am to them both.

My sincere thanks also go to Fréderic Abergel, the director of the Chair of Quantitative Finance and Department of Mathematics Applied to Systems at École Centrale Paris, for accepting me as a PhD candidate in his research group and for ensuring that I had excellent work conditions at all times.

I wish to thank Alan Kirman and Cees Diks for showing interest in my work and promptly accepting to be part of the jury of this thesis.

I am grateful to the members of the CRISIS project[54], in which I was fortunate to be involved from the moment I started my PhD. I am particularly thankful to Cars Hommes and Domenico Massaro, without whom it would have been impossible to carry out the trading experiments discussed in chapter 4.

I take this opportunity to express gratitude to all the people who made my day-to-day at École Centrale Paris an experience that I will certainly miss. A special word here must go to Mehdi Lallouache, who has been not only a great office mate but also a friend.

Last but not least, I thank my close family and Mafalda, who have always managed to be close, despite the geographical distance.

1

Introduction en français

1.1 Une nouvelle réponse

Commençons par citer Jean-Claude Trichet, Président de la Banque Centrale Européenne (BCE) entre 2003 et 2011, lors d'une allocution du 18 novembre 2010 :

When the crisis came, the serious limitations of existing economic and financial models immediately became apparent. Arbitrage broke down in many market segments, as markets froze and market participants were gripped by panic. Macro models failed to predict the

crisis and seemed incapable of explaining what was happening to the economy in a convincing manner. As a policy-maker during the crisis, I found the available models of limited help. In fact, I would go further: in the face of the crisis, we felt abandoned by conventional tools [108].

Dans un souci de prévention, ou du moins dans un souci d'atténuation de la prochaine crise, il est essentiel d'élaborer d'autres approches et d'autres outils pour mieux comprendre la dynamique de l'économie dans son ensemble et en particulier la dynamique du système financier. Faisant écho à cet appel de Jean-Claude Trichet, cette thèse porte sur les phénomènes collectifs déstabilisateurs dans les systèmes socio-économiques. Bien que les phénomènes collectifs jouent un rôle considérable dans la dynamique des économies et des marchés financiers, ils sont omis dans la plupart des modèles et des outils classiques à la disposition des praticiens et des autorités de contrôle.

Les résultats de cette thèse se basent sur deux approches distinctes bien que reliées, à savoir les modèles d'agents et les expériences de trading en laboratoire.

1.2 Approche économique standard

Bien que les modèles économiques standard soient largement utilisés par les praticiens et les décideurs politiques, ils présentent des limites indéniables [77, 80]. Premièrement, les modèles économétriques reposent entièrement sur des données historiques. C'est pourquoi, ils ne peuvent pas faire la lumière sur des conditions inédites. Deuxièmement, les modèles d'équilibre général stochastiques dynamiques (EGSD) doivent être résolus de manière analytique. Pour cette raison, ils exigent des suppositions extrêmement restrictives et sont fortement agrégés. Ces limitations sont d'autant plus marquées que ces modèles sont linéaires pour des raisons de commodité analytique. Ainsi, ils ne peuvent pas décrire les phénomènes procycliques. En effet, par définition, les modèles linéaires ne peuvent pas capturer des conclusions

non linéaires qui caractérisent de tels phénomènes. Par conséquent, si des cycles conjoncturels et de crédit sont occasionnés par des fluctuations non linéaires endogènes de l'économie, les modèles linéaires standard en économie ne pourront jamais modéliser ces cycles avec précision [77, 80, 101].

Parmi les hypothèses hautement restrictives qui caractérisent les modèles classiques en économie et en finance, l'une des plus notables est celle qui considère que les agents prennent des décisions parfaitement rationnelles et peuvent donc être représentés par un seul agent. Ces modèles de rationalité conduisent à des prévisions en contradiction avec la réalité [50], en particulier en période de stress, lorsque la confiance globale est généralement faible et lorsque les individus sont hypersensibles aux événements défavorables. Les données empiriques issues de l'économie comportementale montrent que les individus sont très souvent irrationnels [25] et que leur comportement est très hétérogène [68]. Par exemple, les individus sont disproportionnellement influencés par la crainte du regret [85], propices aux biais cognitifs [56] et souvent trop influencés par des suggestions externes (ancrage) [73]. En outre, les études indiquent que la plupart des individus font également l'objet d'une tendance au statu quo [74]. Pour finir, il existe une grande quantité de preuves pour affirmer que les individus sont trop confiants de manière uniforme et irrationnelle [14, 37, 38].

Outre la parfaite rationalité et à l'agrégation des agents, les modèles classiques ne prennent pas en considération le réseau complexe d'interactions et les mécanismes de rétroaction propres aux marchés financiers et aux économies. Qui plus est, l'hypothèse de la linéarité des interactions entrave également ces modèles classiques dès leur conception. Dans ce contexte, [18] déclare que les modèles économiques standard n'ont pas réussi à décrire la crise financière et économique car ils ne prennent pas en compte les « dark corners », qui expliquent le dysfonctionnement économique. Il explique comment de petites secousses peuvent ébranler les systèmes non linéaires, produire des effets considérables et se solder par des crises.

1.3 Etat de l'art

Depuis le début des années 90, les chercheurs ont de plus en plus recours à des méthodes bien établies en physique statistique pour expliquer les phénomènes provenant de systèmes complexes à grande échelle, qui comprennent souvent une rétroaction non linéaire et des phénomènes collectifs [1, 19]. Par exemple, la dynamique des avalanches observée dans certains systèmes complexes a été soigneusement étudiée dans [103], tandis que [90] a ajouté plusieurs applications à la dynamique sociale collective. L'économie en général et la finance de marché en particulier sont des domaines propices aux phénomènes procycliques, que les modèles linéaires standard ont du mal à traiter car ils ne prennent pas en compte les conséquences macroscopiques éventuelles de l'hétérogénéité dans comportement des agents, comme expliqué dans [77, 79, 80] et plus récemment face à la crise financière globale, comme décrit dans [29]. En outre, les marchés financiers offrent des données à différentes fréquences aux scientifiques et praticiens, ce qui facilite une comparaison entre théorie et données réelles [22].

Contrairement aux modèles économiques et financiers classiques, les modèles à agents décrivent les systèmes complexes comme un réseau d'agents en relation, qui adoptent une approche ascendante. Il n'existe aucune présomption concernant l'équilibre global ou l'optimisation individuelle. C'est pourquoi, contrairement à l'agent représentatif unique, les modèles à agents permettent de mieux comprendre comment les phénomènes macroscopiques complexes surgissent à partir du comportement de plus petits composants du système, d'après le modèle inspiré par Schelling dans [100]. De plus, cette approche peut prendre en considération les données historiques et les conclusions issues de l'économie comportementale via des expérimentations humaines, qui ont été conçues pour cerner le mode de décision de vrais individus dans un contexte économique donné. En réalité, ce processus d'étalonnage est la source du pouvoir prévisionnel des modèles à agents, aux dépens de la rigueur analytique, sensiblement

réduite par rapport aux modèles linéaires simplifiés, dotés d'un seul agent représentatif.

Le degré de complexité et le nombre de paramètres dépendent des phénomènes que le modèle à agents vise à décrire et reproduire. Il est essentiel de déterminer les relations informelles dans un modèle basé sur des agents afin d'en garantir une parfaite compréhension. Bien que cette tâche soit très simple dans un modèle à un ou deux paramètres et variables, elle est nettement plus complexe pour des modèles de dimensions supérieures. En fait, il convient de signaler qu'un modèle stylisé avec quelques paramètres peut être extrêmement difficile à analyser. En d'autres termes, il est généralement préférable de commencer par un simple modèle contenant l'essentiel des mécanismes sous-jacents, puis d'en augmenter la complexité et le nombre de paramètres uniquement après avoir compris parfaitement la version initiale du modèle. De cette façon, il est plus aisé d'assurer le suivi du modèle basé sur plusieurs agents, de bien comprendre ses propriétés dynamiques et de faciliter le choix de stratégies pour atténuer le sur-ajustement lors de l'étalonnage. Dans cet esprit, [8] et [27] sont deux exemples clés de modèles à agents, simples et hautement stylisés, qui traitent le problème de la rationalité et du raisonnement inductif dans le contexte d'un jeu de minorités. En revanche, [11] et [41] sont des exemples de modèles à agents, extrêmement complexes, dotés de plusieurs paramètres, variables et fonctions de rétroaction, pour obtenir des prévisions quantitatives réalistes avec des finalités pratiques.

La crise financière et économique a également prouvé que les modèles standard, bien qu'adaptés en période d'essor, sont totalement inappropriés pour prévoir des crises profondes. Elle a souligné la nécessité de définir une nouvelle approche pour cerner la dynamique de l'économie et elle a transformé la mentalité des économistes et des décideurs politiques. Durant son discours d'ouverture à la Conférence des banques centrales organisée par la BCE, en novembre 2010, le président de la BCE Jean-Claude Trichet a déclaré :

The atomistic, optimizing agents underlying existing models do not capture behaviour dur-

ing a crisis period. We need to deal better with heterogeneity across agents and the interaction among those heterogeneous agents. We need to entertain alternative motivations for economic choices. Behavioural economics draws on psychology to explain decisions made in crisis circumstances. Agent-based modelling dispenses with the optimization assumption and allows for more complex interactions between agents. Such approaches are worthy of our attention [108].

Certaines discussions récentes ainsi que des études de dernière génération sur les économies complexes sont [78] et [69], tandis que [19] porte sur les applications des modèles de physique statistique aux phénomènes socio-économiques. [81] et [71] analysent des modèles à agents financiers, tandis que [3] examine les contributions et les sondages. [48] met l'accent sur l'importance de la modélisation à agents dans le domaine de l'économie et [49] souligne le rôle des modèles à agents dans les décisions politiques. De plus, [43] et [39] discutent de l'application des modèles à agents dans l'analyse politique. Ainsi [84] et [51] analysent l'implication des nouvelles réglementations proposées dans le cadre de l'accord Basel III : le nouveau ratio de levier et les mécanismes de résolutions bancaires. Comme mentionné ci-dessus, l'un des problèmes lié aux modèles à agents avec des équations dynamiques complexes repose sur la difficulté à distinguer les causes et les effets. Pour cette raison, ces modèles finissent par être utilisés comme des boîtes noires dans de nombreux cas.

Dans ce contexte, [61] exploite les simples modèles stylisés de systèmes complexes pour identifier les types de phénomènes reproduisibles par les modèles à agents macroéconomiques. Dans cet article, Gualdi et al. utilisent des méthodes de physique statistique pour trouver les points critiques de transition d'un modèle à agents par le biais du diagramme de phase dans l'espace de paramètres et pour détecter les points de bascule qui occasionnent les transitions critiques.

Il s'agit de commencer par comprendre les fluctuations soudaines observées dans le modèle basé sur les agents proposé par [41]. La conclusion principale est l'existence (dans un modèle

stylisé simple) d'une transition de phase entre une « économie favorable », où le taux de chômage est bas, et une « économie défavorable », où le taux de chômage est élevé. En règle générale, cette transition est déclenchée par une asymétrie entre le taux de recrutement et le taux de licenciement dans les entreprises. Le niveau de chômage reste stable jusqu'au point de bascule, au-delà duquel l'économie s'effondre brutalement. Si les paramètres sont choisis pour que le système fasse l'objet d'une transition, toute fluctuation minime est amplifiée car le système fluctue entre ces deux équilibres. Plusieurs extensions du modèle sont analysées, notamment avec un seuil de faillite, limitant ainsi le niveau maximal du ratio endettement-ventes dans les entreprises. Ceci se solde par un riche diagramme de phases avec une région correspondant aux crises endogènes aiguës, au cours desquelles le taux de chômage monte en flèche avant le redressement de l'économie. L'effet de politiques monétaires simples, visant à contenir la hausse du chômage et à prévenir la crise, fait également l'objet d'une analyse.

L'étalonnage des règles décisionnelles individuelles représente un autre aspect crucial des modèles à agents. Comme mentionné ci-dessus, les anticipations rationnelles ne sont pas une approximation acceptable de la réalité, en particulier en temps de stress financier. Les expériences en laboratoire avec des sujets humains fournissent des bases empiriques indispensables pour créer et étalonner des modèles à agents, où la compréhension du comportement humain réel est essentielle pour former des prévisions. Une méthode clé de la macroéconomie expérimentale est l'étude en laboratoire des comportements collectifs pour analyser en simultané les règles décisionnelles micro-économiques, leurs interactions et le comportement macroéconomique agrégé.

En particulier, [9] a exploité les études en laboratoire pour savoir comment les entreprises définissent leurs prix et quantités dans un contexte de concurrence monopolistique standard, comme le modèle macroéconomique basé sur les agents dans [41]. Leur dispositif reproduit un marché dans lequel les entreprises décident à plusieurs reprises du prix et de la quantité

pour une denrée périssable. Les sujets doivent prendre des décisions relatives à la production et à la tarification compte tenu des informations sur les profits, l'excès de la demande, l'excès de l'offre et le niveau du prix agrégé. Ils confirment que l'hétérogénéité est une caractéristique stable au sein des groupes d'individus, avec environ 46% d'adeptes du marché, 28% d'ajusteurs de profits et 26% d'ajusteurs de la demande. En outre, les prix et les quantités convergent vers un équilibre concurrentiel monopolistique et les sujets se comportent en fonction d'heuristiques d'apprentissage adaptatif.

1.4 Phénomènes collectifs déstabilisateurs

Cette thèse aborde deux sujets de recherche dans le cadre du même projet. La première voie de recherche, expliquée en détail au chapitre 3, est une approche de modélisation relative à la dynamique de confiance dans une société en réseau. La seconde voie de recherche, décrite au chapitre 4, est une approche expérimentale visant à étudier les décisions humaines lors de l'échange d'un actif avec une croissance moyenne positive par période dans un environnement de laboratoire contrôlé.

La confiance est un phénomène collectif et auto-prophétique qui suggère des analogies avec les transitions de phase. Le chapitre 3 aborde les propriétés d'un modèle de réseau caractérisé par des échanges mutuels entre les liens et la fiabilité. Les liens du réseau peuvent représenter une relation commerciale entre deux entités alors que la fiabilité est définie pour chaque nœud dans le réseau en tant que variable réelle. Dans ce modèle stochastique, il est plus probable que deux agents entrent en relation si leur fiabilité est comparable, par exemple en cas d'homophilie, comme dans [30, 47, 87, 89, 97]. De plus, la fiabilité d'un nœud augmente de manière non linéaire avec la fiabilité des autres nœuds associés. Pour finir, une rétroaction positive amplifie les mouvements descendants, qui peuvent se solder

par une chute soudaine et catastrophique de la confiance collective. Il s'agit de reproduire le fait empirique que des événements brusques défavorables ont tendance à être plus drastiques que des événements brusques favorables, occasionnant ainsi le « skew » de volatilité, où les options bénéficiant des chutes du marché coûtent plus que les options bénéficiant des hausses du marché [44]. Les simulations numériques et l'approximation de champ-moyen révèlent l'existence de régions étendues de l'espace de paramètres pour ce modèle, où deux états d'équilibre coexistent : un réseau bien connecté où la confiance globale est élevée et un réseau mal connecté où la confiance globale est faible. Dans ces régions de coexistence, des sauts spontanés peuvent se produire de l'état bien connecté à l'état mal connecté, ce qui correspond à une chute brutale de la confiance non occasionnée par une catastrophe externe majeure. Dans des systèmes de grande taille, les crises spontanées sont remplacées par une dépendance historique : que le système soit dans un état ou dans un autre, le tout dépend essentiellement des conditions initiales. Pour finir, il existe également une phase où les agents sont bien connectés mais ne se font pas confiance mutuellement.

L'approche expérimentale, décrite au chapitre 4, repose sur une experience de trading, où les sujets échangent un seul actif sur un marché artificiel. En l'absence d'échange, il est indiqué aux participants que le prix de l'actif augmentera en moyenne de 2% par période. Bien que les participants soient conscients de cette croissance moyenne, ils savent également que les fluctuations sont aléatoires avec un impact sur les prix. Ainsi, le prix est entraîné à la hausse en cas d'achat et il est entraîné à la baisse en cas de vente. C'est pourquoi, lorsque le prix varie à l'écran, il est impossible de savoir s'il s'agit d'une fluctuation aléatoire ou si quelqu'un a effectué un échange avec une incidence sur le prix. Pour simplifier cette expérience, les sujets ne peuvent pas spéculer ou vendre à découvert, p. ex. ils sont soit sur le marché (détenteurs d'actions) soit en dehors (détenteurs de trésorerie) à tout moment. Sans transaction, les sujets devraient gagner au dela de 600%. Compte tenu de l'impact du marché, plus le taux

d'activité est élevé et plus le gain global des négociants est faible. Initialement, les sujets ont effectué tellement de transactions qu'ils ont atteint un gain moyen de 0.75% à peine. Lorsque cette expérience est renouvelée, les négociations ont nettement diminué et le taux moyen de gain s'élève à 92%. Les sujets sont regroupés dans plusieurs catégories, où des méthodes statistiques visant à réduire le taux de fausses découvertes (FDR) sont appliquées à leurs actions et à leurs positions sur le marché. En outre, les agents sont réfractaires aux risques, comme l'illustre une expérience parallèle basée sur des choix de loterie. En règle générale, une aversion accrue aux risques correspond à un taux réduit des activités et des richesses finales plus élevées. Pour finir, les anticipations de prix des négociants dépendent des investissements et des échanges. Bien que leurs actions ne soient pas rationnelles, elles sont parfaitement cohérentes avec leurs anticipations.

J'espère que les futurs modèles à agents étalonnés avec des données réelles deviendront des outils standard pour l'analyse des scénarios et les prévisions afin de soutenir les décisions politiques dans les domaines de l'économie et des finances. De plus, il serait optimal si les futurs modèles et outils prenaient en considération les conclusions des expérimentations relatives à la nature humaine, notamment en matière de rationalité et de prise de décision dans des scénarios réels. Comme avec la plupart des innovations scientifiques, par exemple avec les prévisions météorologiques [28, 34, 58], il faut s'armer de patience pour que des techniques puissantes puissent évoluer, s'imposer en tant que normes et faire une différence dans le monde réel. Quoi qu'il en soit, ce délai ne doit pas nous désespérer. Au contraire, il doit nous pousser à redoubler d'efforts afin de mieux comprendre l'économie, qui est essentielle au bon fonctionnement de la société.

Un des liens communs entre ces deux thèmes est l'action collective, qui joue un rôle déterminant dans de nombreux phénomènes, par exemple la dynamique de la panique, les faillites et par conséquent le risque systémique [15, 16]. C'est pourquoi, j'espère que ce travail con-

tribuera à l'étude des phénomènes d'actions collectives, en particulier dans la finance quantitative, où les conclusions spécifiques du modèle de confiance décrit ci-dessus et l'expérience de trading en laboratoire pourront être utilisées dans leur état actuel.

2

Introduction

2.1 A new answer

In the words of Jean-Claude Trichet, the President of the European Central Bank (ECB) from 2003 to 2011:

When the crisis came, the serious limitations of existing economic and financial models immediately became apparent. Arbitrage broke down in many market segments, as markets froze and market participants were gripped by panic. Macro models failed to predict the

crisis and seemed incapable of explaining what was happening to the economy in a convincing manner. As a policy-maker during the crisis, I found the available models of limited help. In fact, I would go further: in the face of the crisis, we felt abandoned by conventional tools [108].

In order to prevent or at least mitigate the next crisis, it is therefore paramount that we develop alternative approaches and tools to better understand the dynamics of the economy as a whole and the financial system in particular.

Following the call of Jean-Claude Trichet, this thesis looks into collective destabilizing phenomena in socio-economic systems. Although collective phenomena play a significant role in the dynamics of economies and financial markets, they are often ignored in most conventional models and tools available to practitioners and supervisory authorities. The results presented in this thesis, as in state of the art studies, stem from two different yet connected perspectives: agent-based modelling and trading experiments.

2.2 Standard economics approach

Although they are widely used by practitioners and policy makers, standard models in economics have strong limitations [77, 80]. On the one hand, econometric models are based entirely on historical data, therefore they cannot provide insight into completely new conditions. On the other hand, dynamic stochastic general equilibrium models are to be solved analytically and for this very reason they require highly restrictive assumptions and are highly aggregated. It is especially limiting that, for reasons of analytic tractability, these models are linear. Therefore, they cannot describe procyclical phenomena because, by definition, linear models cannot capture the nonlinear feedbacks which characterize such phenomena. Consequently, if business and credit cycles are caused by endogenous nonlinear oscillations of the economy, standard linear models in economics will never be able to accurately model these

cycles [77, 80, 101].

Amongst the highly restrictive assumptions which characterize traditional models in economics and finance, one of the most notable is the assumption that agents are perfectly rational decision makers which can be linearly aggregated into a single representative agent. Such models of rationality lead to models of expectation formation which do not match what we observe in the real world [50], especially in times of stress, when overall trust is typically low and people tend to be over-sensitive to negative events. This disagreement between theory and facts is to be expected, given the assumptions of these traditional models, which are not observed in reality. In fact, empirical evidence from the field of behavioural economics shows that people very often stray away from rationality [25] and that their behaviour is markedly heterogeneous [68]. For example, people are disproportionally influenced by the fear of feeling regret [85], are prone to cognitive dissonance [56] and are often overly influenced by outside suggestion (anchoring) [73]. Furthermore, experiments show that most people also suffer from status quo bias [74]. Finally, there is a large amount of evidence that people are consistently and irrationally overconfident [14, 37, 38].

In addition to full rationality and aggregating agents, traditional models do not take into account the complex network of interactions and feedback mechanisms typical of financial markets and economies. Furthermore, the above mentioned assumption of linearity in interactions also hampers these traditional models right at their very conception. From this perspective, [18] states that standard economic models have failed to describe the financial-economic crisis because these models did not pay attention to *the dark corners where the economy can malfunction badly*. He emphasizes how small shocks to nonlinear systems can have large effects and lead to crises.

2.3 State of the art

Since the early 1990s researchers have been resorting more and more to well-established methods in statistical physics to explain phenomena arising from large-scale complex systems, which often include nonlinear feedbacks and collective phenomena [1, 19]. For example, the avalanche dynamics observed in certain complex systems has been thoroughly studied in [103], while [90] brought up a few applications to collective social dynamics.

Economics in general, and finance in particular, are fields rich in pro-cyclical phenomena. These phenomena are not easily addressed by standard linear models, which do not account for the macroscopic consequences of agent-level heterogeneity, as explained in [77, 79, 80] and, more recently and in the face of the global financial crisis, in [29]. In addition, finance is a field in which data from a wide range of frequencies is increasingly available to scientists and practitioners, allowing for systematic comparison between theories and real data [22].

Contrary to conventional models in economics and finance, agent-based models describe complex systems as a network of interacting agents using a bottom-up approach. There are no *a priori* assumptions with regards to global equilibria or individual optimization. Therefore, contrary to the single representative agent framework, agent-based models allow for a detailed understanding of how complex macroscopic phenomena emerge from the behaviour of the smaller elements of the system, very much in the spirit introduced by Schelling in [100]. Moreover, this approach can take into account both historical data and findings from the field of behavioural economics through human experiments, which are designed to understand how real people make decisions in an economic context. In fact, this calibration process is precisely where the predictive power of agent-based models stems from, at the expense of analytic tractability, which is significantly reduced compared to simplified, linear models with a single representative agent.

The degree of complexity and the number of parameters of a particular agent-based model depend on the features it is expected to describe and reproduce. In order to fully understand an agent-based model, it is crucial to establish what the causal relationships in the model are. Although this might be a trivial task in a simple model with one or two parameters and variables, it grows significantly harder for more complex, higher-dimensional models. In fact, it is worth noting that even a stylized model with a few parameters may be incredibly hard to analyse. This means that it is usually a sound idea to start off with a simple model containing the essence of the underlying mechanisms, and increase the model complexity and the number of variables and parameters only after the initial version of the model is fully understood. This way, it is easier to keep track of what is causing what in the agent-based model, which allows for a proper understanding of the dynamical properties of the model and makes it easier to adopt strategies that mitigate overfitting at the calibration stage. In this spirit, [8] and [27] are two notable examples of simple and highly stylized agent-based models, which tackle the issue of bounded rationality and inductive reasoning in the context of a minority game. On the other hand, [11] and [41] are examples of very complex agent-based models, with several parameters, variables and feedback functions, which aim at realistic quantitative predictions with practical applications.

The financial-economic crisis brought additional evidence that standard models, while adequate for the good times, are utterly inadequate for predicting major crises. It has highlighted the need for an alternative approach to understand the dynamics of the economy and it changed the way of thinking of economists and policymakers. In the opening address to the ECB Central Banking Conference, November 2010, then president of the ECB Jean-Claude Trichet said:

The atomistic, optimizing agents underlying existing models do not capture behaviour during a crisis period. We need to deal better with heterogeneity across agents and the interaction

among those heterogeneous agents. We need to entertain alternative motivations for economic choices. Behavioural economics draws on psychology to explain decisions made in crisis circumstances. Agent-based modelling dispenses with the optimization assumption and allows for more complex interactions between agents. Such approaches are worthy of our attention [108].

Following the above call of Jean-Claude Trichet, then president of the ECB, the Complexity Research Initiative for Systemic Instabilities (CRISIS), in which I took part, was set up in 2010. It consisted in a multi-disciplinary group of universities, private firms and policymakers whose goal was to build an agent-based model of the European economy and financial system based on how people and institutions behave in the real world [54].

Recent discussions and up to date surveys of complexity economics are [78] and [69], while [19] discusses applications of statistical physics models to socio-economic phenomena. In addition, [81] and [71] survey financial agent-based models, while [3] reviews further contributions and surveys. Moreover, [48] stresses the importance of agent-based modelling for economics and [49] focuses on the role of agent-based models in policy decision making. Furthermore, [43] and [39] discuss the application of agent-based models in policy analysis. In particular, [84] and [51] analyse the implication of new regulations proposed under the Basel III agreement: the new leverage ratio and the banks resolution mechanisms.

As mentioned above, one of the problems with agent-based models with many agents and complex dynamical equations is the difficulty to distinguish cause from effect. Due to this issue, such agent-based models turn out to be used as black boxes in many cases.

With this in mind, [61] uses simple stylized models of complex systems to explore the possible types of phenomena that macroeconomic agent-based models can reproduce. In this paper, Gualdi et al. use methods from the field of statistical physics to characterize an agent-based model through its phase diagram in the parameter space and to detect the tipping points that cause critical transitions.

Their starting point is to understand the sudden fluctuations observed in the agent-based model put forward by [41] and their major finding is the existence, in a simple stylized framework, of a phase transition between a "good economy", where unemployment is low, and a "bad economy", where unemployment is high. This transition is typically induced by an asymmetry between the rate of hiring and the rate of firing of the companies. The unemployment level remains small until a tipping point, beyond which the economy suddenly collapses. If the parameters are such that the system is close to this transition, any small fluctuation is amplified as the system jumps between the two equilibria. A number of extensions of the model are explored, e.g. with bankruptcy threshold, limiting the firms maximum level of debt-to-sales ratio. This leads to a rich phase diagram with a region corresponding to acute endogenous crises, during which the unemployment rate shoots up before the economy can recover. The effect of simple monetary policies that attempt to contain rising unemployment and prevent crises are also explored.

Another crucial aspect of agent-based models is the calibration of individual decision rules. As stated above, rational expectations are not a good approximation of reality, especially in times of financial stress. Laboratory experiments with human subjects provide the empirical guidance needed to construct and calibrate agent-based models in which the understanding of real human behaviour is vital to predictions. A key feature of experimental macroeconomics is the laboratory study of group behaviour, to investigate simultaneously individual decision rules at the micro level, their interactions and the aggregate macro behaviour.

In particular, [9] used laboratory experiments to study how firms set prices and quantities in a standard monopolistic competition setting, such as the macro agent-based model proposed in [41]. Their set-up considers experimental markets in which firms decide repeatedly both on price and quantity of a perishable good. Subjects are asked to make both production and pricing decisions given different information sets on individual profits, excess de-

mand and excess supply, and on the aggregate price level. They confirm that persistent heterogeneity is a characteristic feature within groups of individuals, with about 46% of market followers, 28% profit-adjusters and 26% demand adjusters. Moreover, prices and quantities converge to the monopolistic competitive equilibrium and subjects behave according to adaptive learning heuristics.

2.4 Collective destabilizing phenomena

This thesis reports on two different research topics belonging to the same project. The first research avenue, which is thoroughly explained in chapter 3, is a modelling approach to the dynamics of trust in a networked society. The second, whose description can be found in chapter 4, is an experimental approach to study human decisions when people trade an asset with a positive average growth per period in a controlled laboratory environment.

Trust is a collective, self-fulfilling phenomenon that suggests analogies with phase transitions. Chapter 3 looks into the properties of a network model in which there is a mutual feedback between links and trustworthiness. Links in the network can represent, for example, a business relationship between two entities, while trustworthiness is defined for each node in the network as a real-valued variable. In this stochastic model, two agents are more likely to link if they have more similar trustworthiness, i.e. there is homophily as in [30, 47, 87, 89, 97]. In addition, the trustworthiness of a node increases nonlinearly with the trustworthiness of the other nodes it is connected to. Finally, there is a positive feedback loop that amplifies downward movements, which can lead to a sudden and catastrophic drop of collective trust. This aims to reproduce the empirical fact that sudden negative events tend to be more drastic than sudden positive events, causing the so called "volatility skew," in which options that benefit from market drops cost more than options that benefit from market rises [44]. Nu-

merical simulations and mean-field analytic arguments reveal that, for this model, there are extended regions of the parameter space where two equilibrium states coexist: a well-connected network where global confidence is high, and a poorly connected network where global confidence is low. In these coexistence regions, spontaneous jumps from the well-connected state to the poorly connected state can occur, corresponding to a sudden collapse of trust that is not caused by any major external catastrophe. In large systems, spontaneous crises are replaced by *history dependence*: whether the system is found in one state or in the other essentially depends on initial conditions. Finally, there is also a phase in which agents are well connected yet distrustful.

The experimental approach described in chapter 4 consists of a trading experiment in which subjects trade a single asset in an artificial market. In the absence of trading, the price of the asset is guaranteed to grow 2% per period on average. Although the participants are well aware of this average growth, they also know that there are random fluctuations and price impact, i.e., the price is pushed up whenever they decide to buy and the price is pushed down whenever they decide to sell. Therefore, when a change in price shows up on the screen, it is not possible for them to know if it is a random fluctuation or the result of someone in the room executing a trade and impacting price in the process. In order to keep the experiment simple, subjects are not able to hedge or short-sell, i.e., they are either in the market (holding shares) or out of the market (holding cash) at each point in time. Without trading, the subjects would make a net profit larger than 600%. Due to market impact, the higher their activity, the smaller the overall gain of the traders. At first, subjects traded so much that they achieved a meagre 0.75% average gain. When the experiment is repeated, trading activity is much decreased and traders earn 92% on average. Traders cluster into different categories when false discovery rate statistical methods are applied to their actions and market positions. In addition, agents are shown to be risk averse through a parallel experiment based

on lottery choices. Moreover, in general, a higher risk aversion corresponds to a lower activity rate and a higher final wealth. Finally, traders' price expectations depend on whether they are invested or not, and on their trading actions. Although this is not rational, their actions are fully consistent with their expectations.

It is the hope of the author that in the future agent-based models calibrated with real data become standard tools in scenario analysis and forecasting to support policy making in the fields of economics and finance. Furthermore, it would be optimal if future models and tools took into account findings from experiments in regards to human nature, especially concerning the issues of rationality and decision making in real-life scenarios. As it happened in the past with most scientific innovations, for example in the field of weather forecasting [28, 34, 58], it can take a while before powerful techniques can evolve, become the norm and make a difference in the real world. In any case, this should not cause us to lose hope but rather motivate us to redouble our efforts to better understanding the economy, which is, after all, essential to the understanding of society as we know it.

One of the common links between these two topics is collective action, which is a key player in a number of phenomena, for example in the dynamics of panic, bankruptcies and, consequently, systemic risk [15, 16]. Therefore, the author hopes that this work will contribute to the study of collective action phenomena, especially in the field of quantitative finance, in which it is more likely that the specific findings from the above described trust model and trading experiment can be used in their present form.

3

Sudden trust collapse in networked societies

The contents of this chapter are the result of joint work with Jean-Philippe Bouchaud and Damien Challet [35].

3.1 Introduction

In the wake of the 2008 crisis, President Barack Obama declared: *Our workers are no less productive than when this crisis began. Our minds are no less inventive, our goods and services no less needed than they were last week, or last month, or last year* [93]. So what had happened

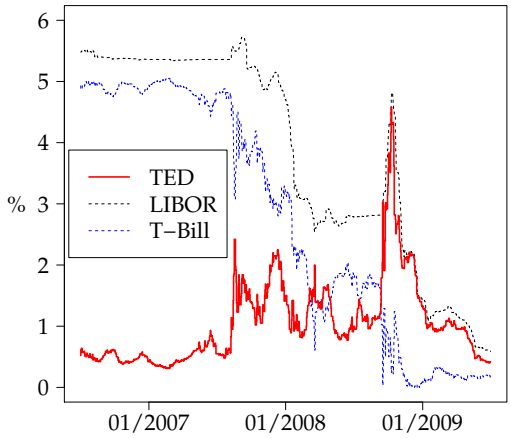


Figure 3.1: TED spread, three-month LIBOR and three-month T-bill interest rate (Jul. 2006 – Jul.2009). The TED spread is the difference between the three-month LIBOR and the three-month T-bill interest rate. Taking into account that T-bills are considered risk-free, an increase in the TED spread is an indicator of higher perceived credit risk in the overall economy. In 10th October 2008, in the wake of the bankruptcy of Lehman Brothers, the TED spread reached 4.57%, several times above the long term average of 0.30%!

that made the world so different from a few months before? No war or physical catastrophe had occurred that would have destroyed tangible assets, infrastructures or knowledge. As implied by President Obama's comment, the damage seems to have been, at least partially, self-inflicted by a sudden collapse of trust that led to a "freeze" of the interbank lending network (evidenced by soaring interbank rates, see Fig. 3.1) and, nearly immediately afterwards, to a collapse of confidence of all economic actors – investors, firms, households interrupted projects and reduced consumption, driving the economy to a grinding halt.* The bewildering aspect of such a crisis (as well as many previous ones) is the speed at which financial markets, or the economy as a whole, can shift from a relatively efficient state to a completely dysfunctional one. Whereas most "real" economic factors (technology, workforce, R&D) usually change relatively slowly, trust or subjective expectations seem to have no inertia, no anchor to their past values, and can swing from high to low in a matter of days, hours or even minutes.

Trust is critical in determining the prosperity of human societies and to secure a well-

*Those who were in New York at the end of Sept. 2008 will remember the sight of completely empty retail stores and the stories of people emptying their bank accounts and going home with cash in plastic bags.

functioning economy and orderly financial markets. Moreover, trust is a collective asset that allows efficient coordination and cooperation, and tremendously accelerates business. It allows for the emergence of genuinely collective figments, such as money and other social conventions. Fiat money is a perfect example: a piece of paper can only be valuable if everybody believes that it will not be worthless tomorrow, and if everybody does, bank notes indeed become valuable.

The fact that trust is (as we view it) a collective, self-fulfilling phenomenon suggests analogies with phase transition phenomena, where collective properties emerge that cannot exist at the individual level, like magnetism, superfluidity, etc. Magnets, for example, arise because the spin of each atom acquires a favoured orientation, imposed by the favoured orientation of neighbouring atoms. This occurs when the interaction between spins becomes strong enough. Clearly, trust emergence is similar, and follows from positive feedback loops such as *I trust you because he trusts you because I trust you*. The most important aspect of the analogy with phase transition is the possible coexistence of very different equilibrium states, which leads to dis-equilibrium phenomena like *history dependence* or "hysteresis", when the system is trapped in one equilibrium while another is more favourable, and discontinuities, when the system jumps from one state to the other. This is an interesting scenario as it opens a path to explain the sudden swings of trust that seem to underpin many economic, financial or political crises.

Several models for trust collapse have been studied along these lines in the past few years, see e.g. [2, 4, 5, 16, 17, 19, 23, 24, 31, 32, 52, 53, 63, 66, 96, 104] and references therein. The common crucial feature is the coexistence of two (or more) equilibrium states in a region of the parameter space, and therefore the possibility of a sudden jump between a favourable, high-confidence state to an unfavourable, low-confidence state. In these models, the jump is not induced by a major catastrophe (that would *replace* the favourable equilibrium by an

unfavourable one) but rather by some anecdotal random fluctuation, which can induce a transition toward an *already pre-existing* low-confidence equilibrium.

Here, we introduce and study a highly stylized model for the build-up and collapse of collective trust in a dynamically evolving network, which generically displays a first order transition with possible coexistence of different equilibria. The nodes of the network can represent individuals, firms, banks, etc. Each node is assigned a real number that measures its (perceived) trustworthiness.

The presence of an undirected link between two nodes indicates an established relationship of some kind (business, loan, collaboration, etc.) resulting from some common rational benefit, but only possible if the perceived trustworthiness of the partner is high enough. Links are thus created or destroyed depending on the trustworthiness of the nodes and their dynamics; conversely, the trustworthiness of a node depends on that of its neighbours. The network and the trustworthiness therefore co-evolve and, depending on the precise specification of the model (see below), this leads to a rich dynamics with crises where the network disintegrates and the collective trust collapses. We solve our model within a mean-field approximation and find, as anticipated, that there is a region of parameters where different equilibria indeed coexist.

Our model and results are in several ways similar to those obtained by M. Marsili and associates in two very inspiring papers [46, 88]. They also study the coupled dynamics of links and nodes and find generic phase coexistence and hysteresis. One new aspect of our work is to consider that the *speed of change* of trustworthiness is itself a piece of information which agents strongly react to, in particular when it is negative – in a "panic feeds panic" spirit. Our mean-field analysis describes the phenomena induced by this effect and predicts phases which had not been considered before, such as a connected yet distrustful phase. In a sense, our model is a stylized version of [4, 5] that removes all the specifics of the interbank lending

network, and a generalized version of [46], where some ingredients specific to the dynamics of trustworthiness are introduced, leading to new effects. The possible coexistence of different states has also been noted in the context of epidemic propagation on networks which may be rewired so as to avoid infected nodes. In this case, infected network situations may indeed coexist with healthy networks [95]. This is similar to our model, where agents/firms/banks tend to cut their links with degraded nodes.

3.2 The model

3.2.1 Trustworthiness of the nodes

The nodes in the network are agents which can represent individuals, companies, banks or other institutions. We make the strong assumption that the *perceived* trustworthiness of a node i, which determines its propensity to link with other nodes, can be summarized by the value of a real number $-\infty < h_i < +\infty$. That real number may depend on a variety of factors, which can be deemed either objective or subjective depending on their underlying nature. The balance sheet of a bank or the health of a business are examples of objective or "intrinsic" factors. Subjective factors come into play, for instance, when one needs to assess how trustworthy the counterparties or business partners of i are. Clearly, if the debtors of i are close to bankruptcy, they endanger the balance sheet of i itself – this mechanism is at the core of many recent models of bankruptcy cascades such as [16, 24, 32, 33, 53, 86, 104]. But one can imagine different, less mechanical channels of propagation. A good example for our purpose is reputation risk. In fact, if node j is caught up in a scandal while making business with i, other partners of i might become wary that i is also involved and decide to end their business with i, unless i reacts immediately and severs its own link with j.

Another important factor is the speed of variation of the trustworthiness itself. Imagine a

highly respected bank or institution i that rapidly loses many of its partners. This will be interpreted as worrying news by the remaining partners who, as a precautionary measure, will be tempted to cut their relation as well, even if the trustworthiness of i is still high. This "bank run" or "panic" type of feedback loop can be amplified by the existence of a CDS (Credit Default Swap) market, which is supposed to price the default probability of firms and banks (and countries) and thus a proxy for h_i. The very fact that the price of the CDS increases (and thus the perceived default probability) can trigger a crash-type dynamics. These avalanches of sell-offs when the perceived risk increases are often observed in financial markets as a consequence of a highly conservative management of "Black Swan" events – that, ironically, may result from these risk management policies!

Mathematically, we therefore write the trustworthiness h_i of each node i as the sum of three terms:

$$h_i = h_{i,0} + fh^* k_i \tanh\left(\frac{\overline{h_i}}{h^*}\right) + d \cdot \min\left(0,\ \delta h_i\right), \tag{3.1}$$

where f, h^*, d are positive constants, k_i is the degree of node i, $\overline{h_i} = (\sum_{j \in V_i} h_j)/k_i$ is the average trustworthiness of the nodes $j \in V_i$ that are connected to i (with $\overline{h_i} \equiv 0$ if $k_i = 0$), and δh_i is the variation of h_i over the last time step.

The first term $h_{i,0}$ is the intrinsic trustworthiness of node i, assumed here to be time-independent, IID random variables with mean m and variance σ^2. More specifically we will choose $h_{i,0}$ to be uniformly distributed in the interval $[0, 2]$, corresponding to a positive mean $m = 1$ and $\sigma^2 = \frac{1}{3}$.

The second term describes how much of the trustworthiness of the peers of i is bequeathed to i. When $\overline{h_i}$ is much smaller than a characteristic value h^*, expanding $\tanh(x)$ for small

arguments gives the following contribution:

$$fh^* k_i \tanh\left(\frac{\overline{h_i}}{h^*}\right) \approx f \sum_{j \in V_i} h_j, \tag{3.2}$$

which means that a fraction f of the total trustworthiness of the business partners of i is transferred to i itself. The tanh function imposes a saturation: for large average trustworthiness, node i only receives a quantity $fh^* k_i$ that grows with the number of neighbours but not with the value of $\overline{h_i}$.

Finally, the third term accounts for the dependence of the current trustworthiness on its *speed of change*. In particular, δh_i increases with the difference between the current and previous trustworthiness values, while the minimum operator implies that only negative recent changes are considered. Therefore, the coefficient d tunes the amplification of negative events and introduces an asymmetry between positive and negative trustworthiness variations. In a sense, it measures the susceptibility of a population to panic. For simplicity, we shall refer to d as "panic factor". The exact definition of δh_i can be found in appendix A.1.

How real is our notion of perceived trustworthiness h_i? How could it be measured, for example? As mentioned above, one clear example are the CDSs of companies, which directly price the default probability as seen by market participants. Another possibility is to gauge the trustworthiness of individuals and firms through surveys, as discussed in [55], echoing a concern expressed by Putnam [98]: *since trust is so central to the theory of social capital, it would be desirable to have strong behavioural indicators of trends in social trust or misanthropy. I have discovered no such behavioural measures.* Even if there is still a lot to be done in order to devise faithful, quantitative indicators of trustworthiness in general, it is highly plausible that the final answer will not be a single real variable as we assume, but a more complex, higher dimensional object. Nevertheless, we believe that the results obtained below, in particular

those pertaining to the co-existence of different equilibria where collective trust is present or absent, will survive in more elaborate models of trustworthiness.

3.2.2 Network dynamics

We now specify how links in the network are created or broken depending on the trustworthiness of the nodes. Since the latter depends itself on the degree of the nodes and on its dynamics, we end up with a model of coupled trustworthiness/network dynamics which shows interesting properties, much as in [46].

At each time step, we choose a pair of nodes at random, say (i, j), characterized by their trustworthiness h_i and h_j. The total number of nodes is constant in time and equal to N. The global average (over all nodes) of h_i, which characterizes the overall confidence level in the network, is denoted by $\overline{h} = \sum_i h_i/N$.

Link creation

If there are no links between i and j, the probability Π_{ij}^{+} that they decide to do business together is

$$\Pi_{ij}^{+} = \frac{r}{N} \frac{z_{ij}}{1 + z_{ij}}, \tag{3.3}$$

where $0 < r < N$ is the a priori propensity to enter into a business relation (the factor $1/N$ is discussed below) and $z_{ij} \geq 0$ is a modulating factor that depends on the trustworthiness h_i and h_j as follows:

$$z_{ij} = \exp\left(\alpha\overline{h} - \beta|h_i - h_j|\right), \tag{3.4}$$

where α, β are two positive parameters. Therefore, a small value of z implies a small probability of link formation. The term $\alpha\overline{h}$ attempts to capture the idea that a trustful society eases the creation of new collaborations or business relations, i.e. that *a rising tide lifts all boats*.

This is the essential virtue of trust that we discussed in the introduction: it acts as a catalyst to exchange and activity, an effect that we attempt to model through α. It is quite clear that together with Eq. (3.1) above, this term can lead to a virtuous circle – more confidence leads to a more connected society which in turn leads to more confidence.

The second term $-\beta|h_i - h_j|$ decreases z and is consequently detrimental to link creation. This attempts to account for "homophily", i.e. the intuitive fact that two entities with very similar credit level are more likely to conduct business together than less comparable peers [30, 47, 87, 89, 97].

Instead of coupling z_{ij} to the overall confidence level $\overline{h}$, one could have imagined to use only the "local" trustworthiness $h_i + h_j$. We have in fact investigated a generalized model in which

$$z_{ij} = \exp\left[\alpha\overline{h} + \alpha'(h_i + h_j - 2\overline{h}) - \beta|h_i - h_j|\right], \tag{3.5}$$

where the α' term captures deviations from the global average. We have found numerically that the new α' term does not change much the phenomenology of the model. This will be confirmed by the mean-field approximation below. We will thus set henceforth $\alpha' = 0$.

Link destruction

If there is a link between the chosen pair (i, j), it is destroyed with probability

$$\Pi_{ij}^{-} = \frac{1}{1 + z_{ij}} \in [0, 1], \tag{3.6}$$

which tends to unity when $z \ll 1$, i.e. when average confidence is very negative, or when homophily is strong ($\beta \gg 1$), both being detrimental to maintaining relationships. The specific choice for $\Pi_{ij}^{\pm}$, and the factor N^{-1} in front of Π_{ij}^{+}, can be understood by calculating the probability P_{ij} that the link between i and j exists in the stationary state. Assuming z_{ij} to

be time independent, P_{ij} is the solution of

$$\Pi_{ij}^{+}(1 - P_{ij}) - \Pi_{ij}^{-} P_{ij} = 0 \implies P_{ij} = \frac{rz_{ij}}{rz_{ij} + N} \underset{r \ll N}{\approx} \frac{rz_{ij}}{N}. \tag{3.7}$$

Therefore, when r, z are both of order unity, the probability that a link exists is of order $1/N$ and the typical degree of a node is itself of order $zr = O(1)$. This is the scaling needed in order to have a non trivial dynamics in the limit $N \to \infty$.

3.3 Numerical results

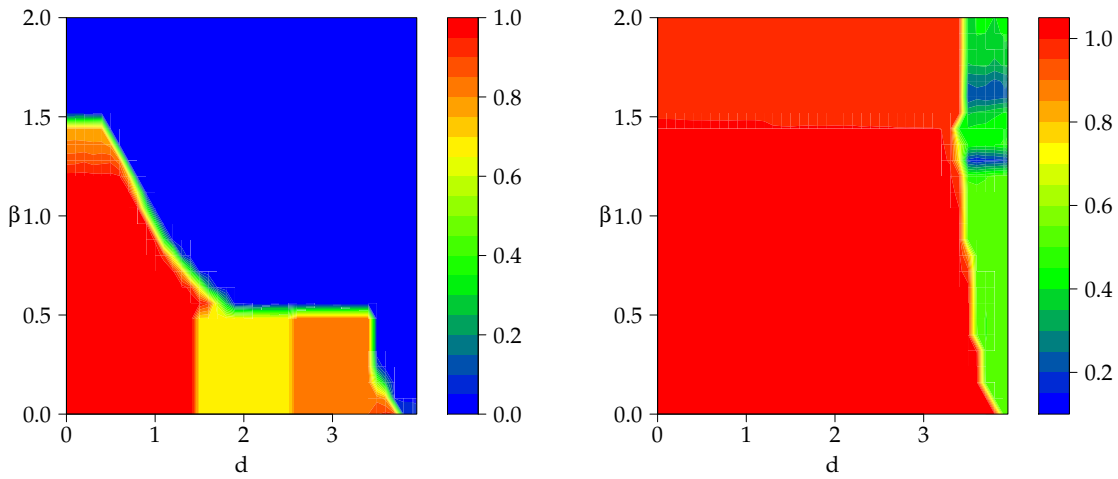
We have numerically investigated this model in detail for various values of its six parameters: f, h^*, d for trustworthiness and α, β, r for link creation/destruction. Some initial conditions for the h's and for the state of the network also need to be specified to run the dynamics. It turns out that as soon as N is somewhat large (i.e. $N \gtrapprox 50$), and for some regions in parameter space, the dynamics of the model becomes *history dependent*, in the sense that starting from an empty network (no links at all) or a full network (all links are present) leads to completely different stationary states – at least over time scales that can be reached in simulations and hence in reality as well (if our model captures anything of reality).

The most important parameters of our model appear to be the homophily parameter β and the panic factor d. This will be justified within our mean-field approximation below: as long as the confidence parameter α is not vanishingly small and r is large enough, the phenomenology of the model is mostly determined by β and d. We have therefore plotted the phase diagram of the model in the (d, β) plane, and the results are shown in Fig. 3.2. We represent the average density of links $\tilde{L} = \langle k \rangle / N$ of the network in a color code, starting from an empty network at $t = 0$ (Fig. 3.2a) or from a densely connected network (Fig. 3.2b). Similar patterns appear when one represents the average confidence $\overline{h}$ instead. One observes a

clear boundary line $\beta_c(d)$ separating two distinct phases: one in which the network is sparse in the stationary state, corresponding to a low average confidence $\overline{h}$, and another in which the network is dense, corresponding to a high average confidence $\overline{h}$. However, this boundary line shifts to significantly higher values when one starts from an already dense network. In other words, there is a large crescent region in phase space where the two outcomes (sparse or dense) are possible, and where the initial condition determines the fate of the network. Another way to illustrate this is to show the evolution of the density of links and of the average trustworthiness $\overline{h}$ as a function of d as one cycles along the line $d = 2\beta$ as in Fig. 3.3a and Fig. 3.3b.

For small N (but still large enough to be of practical interest, say $N \lesssim 100$) the system can in fact alternate between these two states, leading to interesting endogenous crises – i.e. large swings between high confidence and low confidence that are not due to any particular event, but are the result of the noisy evolution of a system for which two very different equilibrium states coexist – see Fig. 3.5. As N grows larger and larger, the probability to jump from one state to another becomes exponentially small, a typical behaviour of physical systems undergoing a first order phase transition (see below for a discussion of this point within a mean-field approximation). However, interesting dynamics will follow from the time-variation of parameters. A suggestive numerical experiment is to let the average value m of the intrinsic trustworthiness $h_{i,0}$ slowly evolve with time, in order to model a progressive shift of the objective state of the economy. When the system is in the coexistence region, one observes a succession of booms and crises, corresponding to jumps between the two underlying equilibrium states – see Fig. 3.4a and Fig. 3.4b.

An analytic description of the dynamics of crisis and recovery can be performed, in particular when $\beta = 0$ and close to the complete instability limit $d = 4$, which is derived in appendix A.2. We now turn to a mean-field approximation that accounts relatively well for



(a) Initial conditions: $\tilde{L}(t=0)=0$. When $d=0$, the network is mostly dense for $\beta \lessapprox 1.50$. If d increases, the maximum β which allows for the dense state decreases. When $d \gtrapprox 2$, crash phenomena start to take place and $\tilde{L}$ in the dense state is lower than before. As $d \to 4$, the dense state eventually becomes unreachable.

(b) Initial conditions: $\tilde{L}(t=0)=0.9$. The region of the parameter space where the dense state is not the preferred stationary state is smaller than in Fig. 3.2a and the absolute sparse state ($\tilde{L} \approx 0$) is not clearly visible. Further numerical calculations indicate that we would observe the absolute sparse state with these initial conditions beyond $\beta \approx 10$.

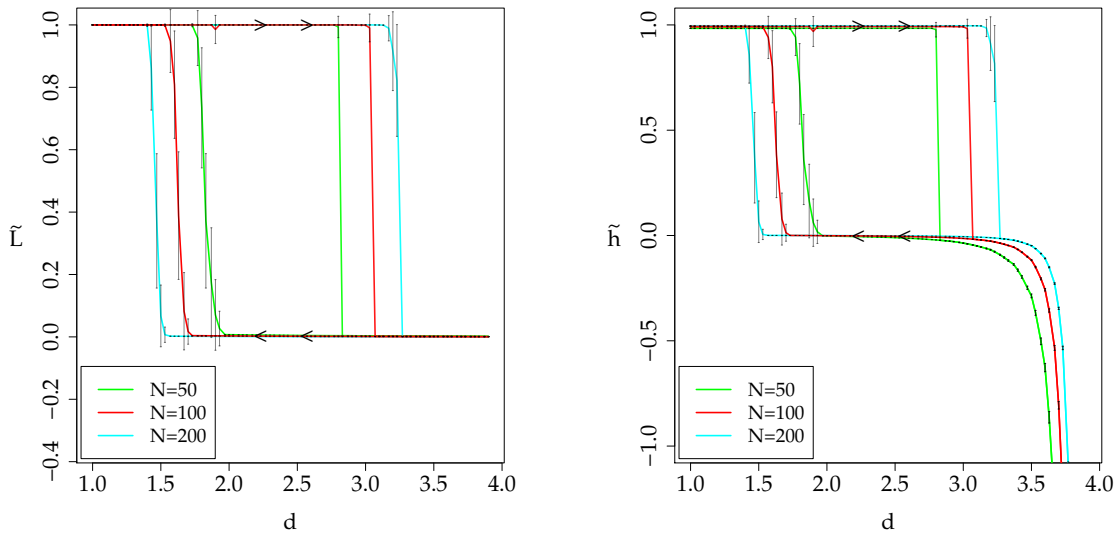
Figure 3.2: Average density of links $\tilde{L}$ for varying d and β and for two different initial conditions (sparse and dense). $N=200$, $\alpha=1$, $f=1$, $h^*=5$, $r=1$, 100 runs and after 10^5 time steps. Regardless of the initial condition, there are two distinct regions in the parameter space, which correspond to two different stationary states, with a sharp transition in between. The red area in the plot corresponds to a dense network ($\tilde{L} \approx 1$) and the blue area corresponds to a sparse network ($\tilde{L} \approx 0$).

our numerical observations.

3.4 A mean-field analysis

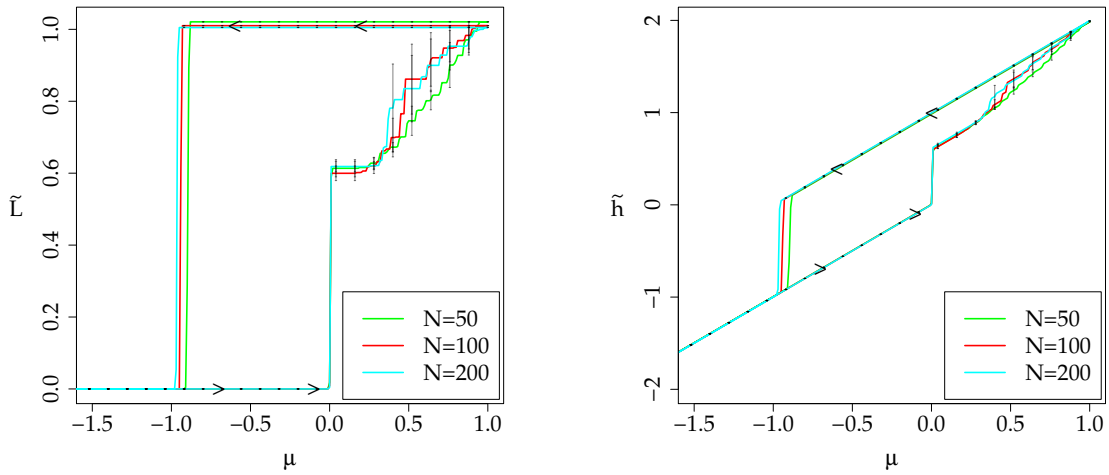
3.4.1 Warm-up: Erdös-Rényi

Let us start by adopting a kinetic view of the standard Erdös-Rényi network with N nodes. At each time step t, a link is randomly chosen among the $\frac{N(N-1)}{2} \approx \frac{N^2}{2}$ possible links. Following the same notation as before, the probability to create a link is $\Pi^+ = \frac{r}{N}\frac{z}{1+z}$, where, for the time being, r and z are constants. If the link is already present, the probability that it is destroyed is $\Pi^- = \frac{1}{1+z}$. We introduce the time-dependent degree distribution $P(k,t)$, i.e. the probability



(a) $\tilde{L}$ **(b)** $\tilde{h} = \frac{\bar{h}}{fh^* N}$

Figure 3.3: Path along $d = 2\beta$ for $N = 50, 100, 200$, $\alpha = 1$, $f = 1$, $h^* = 5$, $r = 1$, 100 runs and 10^6 time steps per point. The average density of links $\tilde{L}$ is shown on the left. The direction along the hysteresis path, in which $\tilde{L}$ ranges from 0 to 1, is represented by black arrows. The plot of the average trustworthiness $\bar{h}$ scaled by $fh^* N$, which we call $\tilde{h}$, is on the right. When d is small and the number of links of the network approaches $\frac{1}{2}N(N-1)$, i.e. $\tilde{L} \to 1$, $\tilde{h} = \frac{\bar{h}}{fh^* N} \to 1$. When $d \to 4$, $\tilde{h} \to -\infty$ and $\tilde{L} \to 0$. The change from one state to the other occurs discontinuously, as observed in first order phase transitions.



(a) $\tilde{L}$ **(b)** $\tilde{h} = \frac{\bar{h}}{fh^* N}$

Figure 3.4: Path along $\mu = m \cdot fh^* N$ for $N = 50, 100, 200$, $\alpha = 1$, $f = 1$, $h^* = 5$, $r = 1$, $d = 2$, $\beta = 2$, 100 runs and 10^6 time steps per point. m is the (time-dependent) common shift added to the original intrinsic trustworthiness of each node $h_{i,0}$. The direction along the hysteresis path, in which $\tilde{L}$ ranges from 0 to 1, is represented by black arrows. When we start at $\mu = 1$, the network is dense and $\tilde{L} \approx 1$. If we continuously decrease μ the network disintegrates ($\tilde{L} \approx 0$) when $\mu \lessapprox -1$. Then, if we increase μ the network will switch back to the dense state ($\tilde{L} \approx 1$) only when $\mu \gtrapprox 0$. The coexistence of two different equilibria allows the system to be trapped in one of these states even if the other is more favourable. Besides, we observe discontinuities when the system jumps from one state to the other.

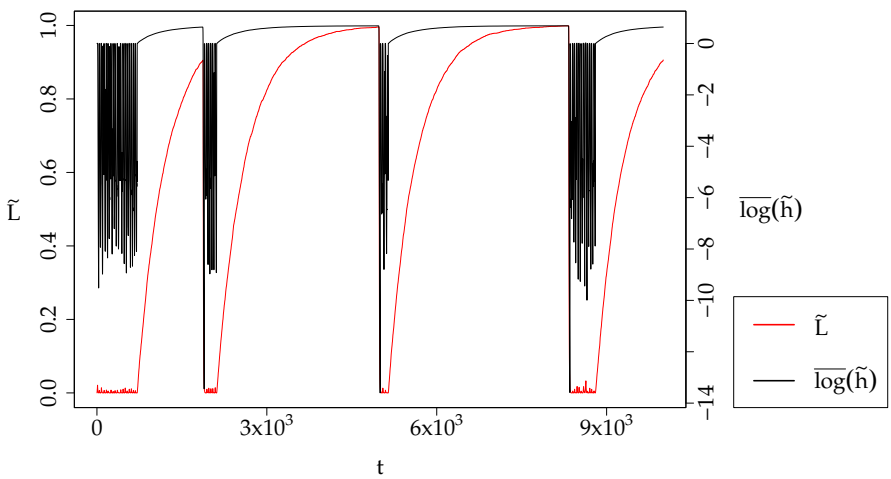


Figure 3.5: $\tilde{h}$ and $\tilde{L}$ (single run) for $N = 100$, $\alpha = 1$, $f = 1$, $h^* = 5$, $d = 3.9$, $\beta = 0$ and $r = 1$. $\tilde{L}(t = 0) = 0$ and the function $\overline{\log}$ is defined as $\overline{\log}(x) = \text{sign}(x).\overline{\log}(|x| + 1)$. We confirm that the system keeps switching between two distinct equilibria: one consisting in a dense network ($\tilde{L} \approx 1$) and the other in a sparse network ($\tilde{L} \approx 0$). The dense state is typically characterized by positive average trustworthiness $\tilde{h}$ which grows steadily in time towards the asymptotic value. On the other hand, the average trustworthiness $\tilde{h}$ in the sparse state tends to be negative and oscillates wildly. The transitions from the sparse state to the dense state are smooth and steady, while the transitions from the dense state to the sparse state, which are triggered by random fluctuations that break links and cause cascade phenomena, are quick and abrupt.

that a randomly chosen node has exactly k outgoing links at time t. The probability that this node changes from $k \to k+1$ in the next time step $t+1$ is

$$W^{+}(k) = \frac{2}{N^2}(N-k)\frac{rz}{N(1+z)}, \tag{3.8}$$

while the probability to change from $k \to k-1$ in the next time step is

$$W^{-}(k) = \frac{2}{N^2}\frac{k}{(1+z)}. \tag{3.9}$$

Making time a continuous variable leads to the following Master equation for $P(k,t)$:

$$\frac{\partial P(k,t)}{\partial t} = \frac{2}{N^2(1+z)}\left[rz\frac{N-k+1}{N}P(k-1,t) + (k+1)P(k+1,t) - \left(rz\frac{N-k}{N}+k\right)P(k,t)\right]. \tag{3.10}$$

By inspection, one finds that $P_0(k) = C_N^k q^k (1-q)^{N-k}$ is a stationary solution of Eq. (3.10), as it should be, provided

$$q = \frac{zr}{zr+N}. \tag{3.11}$$

The average degree $\langle k \rangle$ and the corresponding variance are then given by:

$$\langle k \rangle = Nq = \frac{Nzr}{zr+N} \underset{N\to\infty}{\approx} zr \tag{3.12}$$

$$\langle k^2 \rangle - \langle k \rangle^2 = Nq(1-q) \tag{3.13}$$

The following sections extend the above calculation to the case where z self-consistently depends on the trustworthiness of the nodes.

3.4.2 Coupling with the average trustworthiness $\overline{h}$

We now consider the baseline case where $z = e^{\overline{h}}$, with $\alpha > 0$ and $\overline{h}$ the average trustworthiness of the population. For the time being, we discard all homophily effects or feedback loops (i.e. $\beta = d = 0$).

We first assume that the average intrinsic trustworthiness $h_{i,0}$ has a zero mean, $m = 0$. This is an interesting situation since it does not break the $h \to -h$ symmetry, i.e. collective trust or distrust are a priori equally probable outcomes. Averaging Eq. (3.1) over all nodes and using a mean field argument, i.e neglecting all fluctuations making all h_i different, we find

$$\overline{h} = fh^* \langle k \rangle \tanh\left(\frac{\overline{h}}{h^*}\right). \tag{3.14}$$

This approximation is certainly justified in the dense limit $\langle k \rangle \gg 1$, but breaks down for small $\langle k \rangle$, in particular when $\langle k \rangle < 1$. In this latter case the network does not percolate and, in the absence of a giant component, no collective behaviour is possible. In this case, the only solution to Eq. (3.14) is $\overline{h} \approx 0$.

Suppose for simplicity that $f\langle k \rangle$ is somewhat larger than unity (say 5 or more). Then, it follows that $|\tanh(\frac{\overline{h}}{h^*})| \approx 1$ and Eq. (3.14) has two possible solutions: $\overline{h} \approx \pm fh^* \langle k \rangle$.[†] Now we can plug these solutions in Eq. (3.12), which yields a second self-consistent equation:

$$\langle k \rangle = \frac{Nr}{r + N\exp\left(\mp\phi\langle k \rangle\right)}, \qquad \text{where } \phi \equiv \alpha f h^*. \tag{3.15}$$

[†]For $f\langle k \rangle > 1$ but not so large, the qualitative discussion below remains valid, up to prefactors of order unity.

The positive trust self-consistent solutions

Let us focus first on the case where a *positive average* trustworthiness appears, corresponding to the minus sign in the exponential in Eq. (3.15). Assume first that $\phi\langle k\rangle \gg \log N$. Then, the second term in the denominator is completely negligible and $\langle k\rangle \approx N$, which obeys the above hypothesis provided $\phi \equiv \alpha f h^* > \log N/N$, which we will assume in the following. This corresponds to a self-sustained "euphoric state" where the network is full and confidence at its peak. This solution always exists unless ϕ is vanishingly small: in the absence of the detrimental effects studied below, a dense network should appear due to the positive feedback term that favours link formation when confidence rises.

A second, sparse but percolating (i.e. with a giant component) solution can also exist. To see that this is the case, assume now that $z = \mathcal{O}(1)$. Then, Eq. (3.15) leads to $\langle k\rangle = zr$, where

$$z = \exp\left(\phi z r\right). \tag{3.16}$$

This self-consistent equation depends on the product ϕr:

- When $\phi r > e = 2.71..$, there is no solution to this equation. Only the dense network solution described above exists.

- When $\phi r < e = 2.71..$, on the other hand, there are 2 solutions $z_<$ and $z_>$, one stable corresponding to a sparse, but trustful network, and a dynamically unstable one, which is nevertheless interesting since the associated value for $\langle k\rangle^* = z_> \cdot r$ is the critical value above which a sparse network is unstable and flows towards the fully connected solution above. Said differently, if the spontaneous fluctuations around the stable solution $\langle k\rangle = z_< r$ are not strong enough to reach $\langle k\rangle^*$ with appreciable probability, the sparse network will appear dynamically stable. This is indeed the case when ϕ is

small enough.

A negative trust self-consistent solution

An important question at this point is whether this model also allows for the existence of sustained negative average trustworthiness values $\overline{h} < 0$, i.e. a connected, but suspicious society. This would correspond to the positive sign in the exponential in Eq. (3.15). In this case, the solution for large N is:

$$y\exp(\phi y) = r, \qquad y = \langle k \rangle. \tag{3.17}$$

When $\phi r \ll 1$, the solution of Eq. (3.17) is $\langle k \rangle \approx r$, therefore when $r > 1$ the solution with negative $\overline{h}$ is indeed self-consistent. Hence a self-sustained state of distrust in a sparse network (but with a giant component) is possible when a) ϕ is small enough (i.e. distrust is not too detrimental to link formation) and b) r sufficiently large (i.e. agents meet often enough so that links are created even if the two parties are mutually suspicious). This corresponds, pictorially, to a "wary" society in which distrustful relationships are the norm.

On the other hand, if $\phi r \gg 1$, we have

$$\langle k \rangle \approx \frac{1}{\phi}\left[\log(\phi r) + \mathcal{O}(\log\log(\phi r))\right]. \tag{3.18}$$

Equation (3.18) shows that as ϕ grows, $\langle k \rangle$ decreases until the giant component disappears (when $\langle k \rangle < 1$) and the solution with $\overline{h} < 0$ is no longer viable. For large r, this occurs for a certain value $\phi_c \sim \log r + O(\log\log r)$. We have checked numerically that this "wary society" phase indeed exists in our model and is not an artifact of the mean-field approximation.

Summary

Summarizing, for $\phi = \mathcal{O}(1)$ and $r > 1$ there are *three* viable solutions, one corresponding to very dense networks and positive self-sustained collective trust, and the two other to sparse networks (but still percolating, $\langle k \rangle > 1$), one with positive and one with negative self-sustained trust. These latter two solutions however disappear as ϕ increases, beyond $\sim e/r$ for the former and $\sim \log r$ for the latter.

The above analysis assumed that the average intrinsic trustworthiness is $m = 0$. When $m > 0$, the self consistent equation becomes:

$$\overline{h} = m + fh^* \langle k \rangle \tanh\left(\frac{\overline{h}}{h^*}\right). \tag{3.19}$$

Clearly, this equation now selects the dense, positive confidence solution as soon as αm is not vanishingly small. This is the situation we have considered in our simulations.

3.4.3 Coupling with speed of trust degradation

We now study the influence of the panic parameter d on the trustworthiness in Eq. (3.1), i.e. the positive feedback effect that may trigger a link breaking avalanche when an increase of perceived risk takes place. We set the homophily term β to zero for the time being and look into the general case in the next section.

As a warm-up exercise, let us compute the evolution of $\langle k \rangle_t = \sum_k kP(k,t)$ from Eq. (3.10). Multiplying by k and summing over k yields

$$\frac{d}{dt}\langle k \rangle_t = \frac{2}{(1+z)N^2}\left[zrN - (zr+1)\langle k \rangle_t\right]. \tag{3.20}$$

At equilibrium, with $\frac{d}{dt}\langle k\rangle_t = 0$, we trivially recover the result in Eq. (3.12):

$$\langle k\rangle_{eq} = \frac{Nzr}{N+zr}.$$

For small deviations from equilibrium, $\langle k\rangle_t$ is described by an Ornstein-Uhlenbeck process that can be fully characterized from the knowledge of the variance of k.

Now, in our model with feedback we assume that all events contributing to lowering the degree of the nodes will lead to a decrease of trustworthiness. Restricted to events lowering the degree, this contribution can be written as

$$\frac{d}{dt}\langle(\Delta k)_-\rangle = \frac{2}{(1+z)N^2}\Sigma_k kP(k,t) = \frac{2\langle k\rangle_t}{(1+z)N^2}. \tag{3.21}$$

After $T = \frac{N^2}{2}$ time steps, which is the average time it takes to attempt to change the status of each link once, the total contribution to degree decrease is

$$\langle(\Delta k)_-\rangle \approx \frac{\langle k\rangle}{1+z}. \tag{3.22}$$

Again in a mean-field spirit, the resulting expression for z is

$$z = \exp\left(\alpha\overline{h} - d\langle(\Delta k)_-\rangle\right), \tag{3.23}$$

meaning that the stronger the activity that decreases connectivity, the smaller the value of z and hence the larger the probability of breaking further links. There is also a second contribution to $\min(0, \delta h)$ arising from the time fluctuations of $\overline{h}$ itself, but it is much smaller in the equilibrium region we are focusing on.

Hence, we find a set of self-consistent equations valid when $f\langle k\rangle \gg 1$ and $\overline{h} > 0$:

$$z = \exp\left[\langle k\rangle\left(\phi - 2d\frac{1}{1+z}\right)\right] \tag{3.24}$$

$$\langle k\rangle = \frac{Nzr}{N+zr}. \tag{3.25}$$

Let us study the possible solutions to Eq. (3.24) and Eq. (3.25). Suppose first that $N \ll rz$. In this case, we have from Eq. (3.25) that $\langle k\rangle \approx N$. The self-consistent Eq. (3.24) then leads to

$$z \approx \exp\left(\phi N\right),$$

which is indeed such that $N \ll rz$ provided that $\phi \gg \frac{\log(N)}{N}$. This solution corresponds to such dense a network that the downwards degree fluctuations cannot destabilize it, at least locally.

However, there might coexist a second solution, even for values of ϕ where it would not exist for $d = 0$. Suppose now that $z = \mathcal{O}(1)$ and $\langle k\rangle \approx rz$, which we assume to be larger than 1 to allow for non-zero collective trust $\overline{h} > 0$ to exist and be locally stable. The self-consistent equation now reads

$$z = \exp\left(\phi rz - 2d\frac{rz}{1+z}\right). \tag{3.26}$$

It is clear that there is no solution to Eq. (3.26) when d is small and $\phi r > e$. However, there is a critical value of d, denoted by d^*, above which Eq. (3.26) has two solutions: $z_< < 1$, which is stable at least for d not too large, and $z_> > 1$, which is unstable. This is illustrated in Fig. 3.6. As d increases further, $z_<$ becomes smaller and smaller and at one point becomes itself unstable, leading to limit cycle dynamics. This small $z_<$ solution however corresponds to a completely disconnected network.

The existence of a second, sparse solution for large enough d corresponds well to our nu-

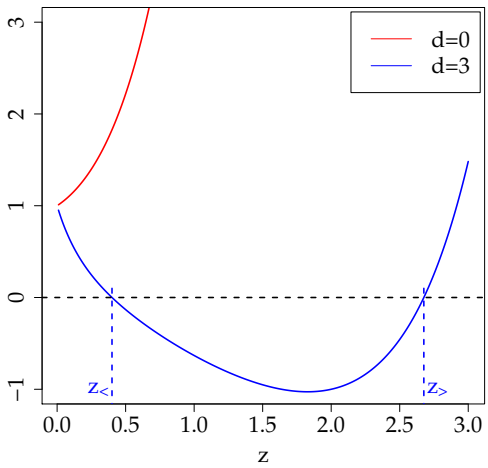


Figure 3.6: Graphical representation of Eq. (3.26) for $d = 0$ and $d = 3$ and $\phi = 2$, $r = 1$. For $d = 0$ (red line) the function $f(z) := \exp\left(\phi r z - 2d\frac{rz}{1+z}\right) - z$ has no zeros, which means that $z = \mathcal{O}(1)$ and $\langle k \rangle \approx rz$ are not compatible conditions in this case. However, for large enough d, for example $d = 3$ (blue line), two solutions emerge: $z_< < 1$ and $z_> > 1$. The former is dynamically stable, while the latter is unstable and characterizes the critical transition path from the dense solution towards the sparse solution.

merical observations: the network attempts to connect but trustworthiness is small and cannot grow because it is killed by spontaneous negative fluctuations.

The intermediate, unstable solution $z_>$ is also interesting as it again characterizes the critical transition path from the dense solution towards the sparse solution (and vice versa). For large d, one finds $z_> \approx 2d/\phi$, corresponding to a characteristic average degree $k_> \approx 2dr/\phi$. When $k_>$ is much smaller than N, the dense solution has an exponentially small (in N) probability of spontaneous destabilisation. However, as $k_>$ increases towards N, fluctuation induced crash events become more and more frequent, as shown in Fig. 3.5.

3.4.4 Homophily

We finally turn to the influence of homophily, i.e. the β term in the definition of z in Eq. (3.4). Here we assume, as in [46], that the network is at all times an Erdös-Rényi network with a

time dependent density of links $q_t = \langle k \rangle_t / N$. We also assume, as above, that the network is well-formed, with $f\langle k \rangle_t$ somewhat larger than unity so that one can assume that for most nodes, the following approximation holds:

$$h_i \approx h_{i,0} + fh^* k_i. \tag{3.27}$$

Once again, two cases should be considered. One corresponds to dense networks in which $\langle k \rangle \sim N$. In this case, fluctuations of node degree are at most of order $\sqrt{N}$. In fact, the homophily term leads to cliques of connected nodes with a relatively homogeneous degree, so we expect these fluctuations to be much smaller than $\sqrt{N}$. Therefore, one can estimate z as

$$\log z \approx \alpha N - \beta\sqrt{N}, \tag{3.28}$$

which shows that unless α is very small, the highly connected phase is not destabilized by homophily.

In the case of sparse but percolating networks with $\langle k \rangle > 1$, the dispersion of trustworthiness that prevents links from forming has two distinct origins. One is the intrinsic heterogeneity of the nodes, measured by the root mean square σ of the fields $h_{i,0}$. The second is the degree heterogeneity which, for an Erdös-Rényi network with $q_t = \mathcal{O}(N^{-1})$, is given by $\sqrt{Nq_t} = \sqrt{\langle k \rangle_t}$. Using $\langle k \rangle_t = zr$, valid in the sparse phase, one finally ends up with the following schematic estimate of the homophily term:

$$\beta|h_i - h_j| \longrightarrow \beta\sqrt{c\sigma^2 + c'zr}, \tag{3.29}$$

where c, c' are numerical constants of order unity. This leads to a new self-consistent equation

for the link activity z in the sparse phase:

$$z = \exp\left(\phi r z - \beta\sqrt{c\sigma^2 + c' z r}\right). \tag{3.30}$$

It is graphically clear that this equation behaves much in the same way as Eq. (3.26): for small β and $\phi r > e$, no solution exists except for dense networks. But as β increases, two non-trivial solutions, $z_<$ and $z_>$ appear, corresponding to a sparse solution that is not able to connect because of the strong repulsion between different nodes. This corresponds to the sparse phase observed in the phase diagram of the model for large β, see Fig. 3.2.

3.5 Conclusion

We have introduced, in the spirit of [46, 88], a highly stylized model for the asymmetric build-up and collapse of collective trust in a network where the links and the trustworthiness of the nodes dynamically co-evolve. The basic assumption of our model is that whereas trustworthiness begets trustworthiness (meaning that a higher level of trustworthiness is more favourable to link formation), trustworthiness heterogeneities, both across nodes and in time, are detrimental to the network. In particular, panic also begets panic, in the sense that sudden drops of trust may lead to link breaking (or "sell-offs" in the context of financial markets) that further decreases trustworthiness. We have shown, using both numerical simulations and mean-field analytic arguments, that there are extended regions of parameter space where two equilibrium states coexist: one corresponds to a favourable, well connected network with a high level of confidence, and the second is an unfavourable, poorly connected and low-confidence state. In these coexistence regions, sudden spontaneous jumps between the two states can occur. These transitions are not induced by any major catastrophe that would replace a favourable equilibrium by an unfavourable one, but rather by random fluctuations

that trigger the switch between two *already existing equilibria*. When the system becomes large, however, these jumps become less and less frequent, unless an external parameter is changed – corresponding, for example, to a measure of the overall economic activity that sets the average trustworthiness level. For large systems, the phenomenon of spontaneous crises is replaced by the notion of strong *history dependence*: whether the system is found in one state or in the other essentially depends on initial conditions: ergodicity is dynamically broken.

Our stylized model only aims at this stage to provide a generic (but certainly oversimplified) conceptual framework to understand how financial markets, or the economy as a whole, can shift so rapidly from a relatively efficient state to chaos, when nothing "material" has changed at all, when *our minds are no less inventive, our goods and services no less needed than they were last week*, as noted by President Obama. Our model illustrates Keynes remark: *a conventional valuation which is established as the outcome of the mass psychology of a large number of ignorant individuals is liable to change violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference* [76]. A theoretical challenge is of course to take our framework seriously and think about how such a model could be calibrated with real data, for example using interbank loan networks (see e.g. [16]), CDS data or survey results as in [55]. An obvious goal would be to obtain early warning signals for potential trust collapse and crises [107] that could, in some cases, look like precursor avalanches or "crackling noise" (see [103], and for a recent review on this theme, [19]).

4

Do investors trade too much? A laboratory experiment.

The contents of this chapter are the result of joint work with Domenico Massaro, Jean-Philippe Bouchaud, Damien Challet and Cars Hommes [36].

4.1 Introduction and aim of the study

Financial bubbles and crises are potent reminders of how far investors' behaviour may deviate from perfect rationality. Many behaviour biases are now well documented, such as propensity for trend following (or extrapolative expectations [59]), overconfidence, herding and panic, disposition bias (cutting gains too early and losses too late), home bias, etc [14]. One of the best established biases is the tendency to trade too much. Because of transaction costs (fees and price impact), the more active traders usually perform worse on average [94]. This is usually related to overconfidence or "animal spirits", i.e., as Keynes put it, *a spontaneous urge to action rather than inaction.* For example, somewhat expectedly, men tend to trade much more than women [13].

Another reason is that it is usually difficult to separate wheat from chaff, i.e. signal from noise. Mistaking noise for signal inevitably leads to excess trading. In fact, finding the optimal frequency at which a noisy signal should be traded in the presence of transaction costs and market impact is a highly non-trivial problem in general.

Artificial market experiments now have a rather long history. They have repeatedly demonstrated that bubbles arise deceptively easily through trading even when the traders know the fundamental price of an asset [106]. More recently, the authors of Ref. [70] found a variety of transient behaviours ranging from rapid convergence of the price to its fundamental value, to oscillations and bubbles. These patterns can be explained by adaptive learning behaviour based on simple price extrapolation rules (trend following, mean reverting/anchoring, etc.), rather than rational expectations.

The present study belongs to the above tradition of artificial markets, where subjects are given clear information about what they should do to optimize their profits, and then left to trade a noisy signal on their own device. Although the rationally optimal strategy is clearly to

buy and hold the (fictitious) asset, we observe that our subjects in fact start trading wantonly, which is both individually and collectively detrimental. When the experiment is immediately repeated with the same subjects, we see a significant improvement of the collective performance, which is however still substantially lower than the (optimal) buy-and-hold strategy. The clear-cut conclusion of our experiment is that our investors do indeed trade too much, but that they improve to some extent when given a second chance.

Because we have run many sessions of our experiment, each with a rather large number of agents, several statistical properties of the price time series and of the behaviour of our economic agents could be studied. For example, although our subjects are physically separated and cannot communicate, we have seen that a significant amount of synchronisation takes place in the decision process, that can therefore only be mediated by the price trajectory itself.

This resonates with what happens in real financial markets, where price changes themselves appear to be interpreted as news, leading to self-reflexivity and potentially unstable feedback loops. In fact, our experimental setting was such that panic and crashes were possible but these did not happen. In fact, synchronisation appeared to be stronger for buy trades than for sell trades, and no cascades or "fire sales" effects could be detected. Although this came as a sort of disappointment (though not related to any shadenfreude on our part), we believe that we understand why our market turned out to be hyperactive and volatile, but stable. This suggests more experiments, with different parameter values, to see whether panic-prone markets can be artificially simulated.

One particularly interesting feature of our artificial market is that we model price impact, i.e. the fact that the very action of agents modifies the price trajectory. This is now believed to be a crucial aspect of real financial markets, which may lead to feedback loops and market instabilities [20, 24, 31, 82, 83]. What is particularly interesting in our experiment is that excess trading significantly impacts the price trajectory and is strongly detrimental to the wealth of

our economic agents. In other words, unwarranted individual decisions can lead to a substantial loss of collective welfare, when mediated by the mechanics of financial markets. Although our setting is highly stylized, this suggests that a similar mechanism may be at play in the real economy as well. In fact, taking into account that in real financial markets, contrary to what happens in our experiment, participants do not have access to clear information and to a rationally optimal strategy, one would expect this detrimental mechanism to be even more pronounced in real markets than in our simplified and controlled experimental set-up.

The outline of this chapter is as follows. We first start in Sect. 4.2 by describing the experimental motivation and set-up. The precise instructions given to our subjects are detailed in Appendix B.1. Sect. 4.3 explains the rational benchmark to which we want to compare the experimental results. We show in particular that rational agents should favour, in the present situation, a buy-and-hold strategy. We then summarize our main results in Sect. 4.4, which includes a refined statistical analysis of the behaviour of agents in 4.4.1. Section 4.5 offers our conclusions, open questions, and future experiments.

4.2 Experiment motivation and set-up

4.2.1 Qualitative description

The basic idea of our experiment is to propose to subjects a simple investment "game" in which they can use the cash they are given to invest in a fictitious asset that will – they are told – increase in value at an average rate of $m = 2\%$ per round, however with some fluctuations of root mean square (RMS) $s = 10\%$ per round. The game may stop randomly at each time step with probability $p = 0.01$. In other words, the game is expected to last around 100 time steps. If an amount w_0 is invested in that asset at time $t = 0$, the wealth of the inactive

investor will accrue to

$$w_T = w_0 \exp\left[mT + \xi s\sqrt{T}\right] \tag{4.1}$$

at time T, where ξ is a noise term with zero mean and unit variance. The numerical value of the term in the exponential is therefore equal to $2 \pm \xi$ for $T = 100$, leading to a most probable profit of $e^2 - 1 \approx 640\%$, and a probability to end up in the red of only approximately 3.5%. As we shall show below, the fully rational decision is to buy and hold the asset until the game ends; for the students participating in the experiment, the most probable gain would represent roughly EUR 160, a very significant reward for spending two hours in the lab. In other words, the financial motivation to "do the right thing" is voluntarily strong.*

In order to make the experiment more interesting, and trading even more unfavourable, the asset price trajectory is made to react to the subjects decisions, in a way that mimics *market impact* in real financial markets. The precise implementation is given below, but the idea is that while a buying trade pushes the price up, a selling trade pushes the price down. It is made perfectly clear to the subjects that this impact will amount for them as a cost, since the transaction price of their trade will be the impacted price. This should therefore be a strong incentive not to trade, since the corresponding transaction cost is approximately 3%, i.e. comparable to the average profit per round.

However, market impact introduces an interesting twist, since subjects now have to guess if the observed price fluctuations are due to the "natural" fluctuations (of RMS s) they are warned about at the beginning of the game, or if they are due to the action of their fellow subjects. This was meant to provide a potentially destabilizing channel, where mild sell-offs could spiral into panic and crashes.

*Let us note in passing that if subjects had behaved rationally, our budget for this experiment would have been exhausted after a few sessions, and we would have been forced to interrupt the experiment!

4.2.2 Experimental set-up

The interested reader should refer to Appendix B.1, in which we reproduce the detailed instructions given to the subjects at the experiment. Each session of the experiment is done with around 30 volunteers, most of them students at the University of Amsterdam (see [7, 10, 12, 68] for details on this and earlier similar experiments). Each subject is randomly assigned a computer in the laboratory. Once the experiment starts, his screen displays a plot and table of the past asset prices, together with the time evolution of his own wealth, measured in a fictitious currency unit, and set to 100 at the start of the experiment. The exchange rate is 100 units of this fictitious currency to EUR 25.

During each session of the experiment, subjects have to decide every 30 seconds if they want to hold cash or shares in the next period. If they have cash at period t, they can decide either to use it all to buy shares or to stay out of the market at period $t+1$. Conversely, if they have shares at period t, they have to choose between selling them all for cash or staying in the market at period $t+1$. Fractional orders are not allowed, therefore each player is either in the market or out of the market at all times.

Physical barriers in the laboratory guarantee that there is no communication between subjects during the experiment. In addition, each set of subjects sits 2 experiment sessions of random duration, so that learning mechanisms can be investigated. The duration of each session is not pre-defined: at each time step the probability that the experiment stops is $p = 0.01$. This indefinite horizon helps to mitigate behaviour bias towards the end of the session, as explained in [62]. For each set of subjects, the combined duration of the two sessions was approximately 1 hour. In fact, we in fact hand-picked exponential-distributed end-times such that the session would not stop too early.† At the end of the experiment, each subject rolls a

†There were in fact 3 experiment sessions per set of subjects. However, the first session, whose sole purpose was to get the subjects at ease with the software in use, was very short and these results were discarded in the

dice to determine which of the two sessions will be used to calculate his take-home profits.

Regarding the initial conditions, we explored two different possibilities. We first studied the case where each player starts the experiment holding 1 share worth 100 units of currency. We realized that this lead to a strong "animal spirit" bias: preferring action to inaction, our subjects started selling very early on for no good reason at all. While interesting in itself, we rather decided to give each player 100 units of currency in cash in the beginning of the experiment. This was done for most sessions and the results we will report below are all initialized by giving our agents cash rather than shares.

In order to prevent fast-thinking subjects from getting bored and losing focus, we set up a parallel experiment in which each subject was asked to predict the nominal price of the asset in the next period after submitting his decision. At the end of the experiment, in order to comply with the minimum average earnings stipulated by CREED[‡] and to measure risk aversion, subjects chose a number of lotteries from a given set and played one of their choices for free (selected at random). Further details about these auxiliary tasks can be found in Appendix B.1.

The crucial aspect of our experimental setting is the price update rule. As mentioned above, we want our market to reproduce price impact, a crucial aspect of financial markets: as an agent submits a (large) buying or selling order at time t, the price p_{t+1} at which the transaction is going to be fully executed is (a) not known to him at time t and (b) adversely impacted by the very order that is executed. The price dynamics of Eq. (4.9) is easily modified to include price impact (see Appendix B.1) and now reads

$$p_{t+1} = p_t \cdot \exp\left(m + s\eta_t + I_t\right), \tag{4.2}$$

analysis.

[‡]The Center for Research in Experimental Economics and Political Decision, Faculty of Economics and Econometrics, University of Amsterdam

where $m = 2\%$, η_t is a noise term drawn from a Student's t-distribution with 3 degrees of freedom and unit variance, as commonly observed in financial markets [22][57].[§] Instead of choosing a new realization of η_t for each run, we keep the very same noise time series in all the experimental sessions, so as to be able to aggregate the results of different sessions.

In Eq. (4.2), s is a constant that sets the actual amplitude of the noisy contribution to the evolution of the price (i.e. the price volatility) and is chosen to be $s = 10\%$. These numbers correspond roughly to the average return and the volatility of a stock index over a quarter. Therefore, one time step in our experiment roughly corresponds to three months in a real market, and 100 steps to 25 years. The last term I_t is the price impact caused by all the orders submitted at time t, which we model as:

$$I_t = \frac{N_t}{N} \frac{B_t - S_t}{B_t + S_t}, \tag{4.3}$$

where N_t is the number of subjects who submitted an order at time t and N is the total number of subjects in a given session (i.e. the "depth" of our market). B_t and S_t, in currency units, are the total amount of buying orders and selling orders, respectively. Note that for a single buying (or selling) order, the impact is given by $1/N$, i.e. around 3% for a market with 30 participants (and less if the market involved more participants, as is reasonable). On the other hand, if all the agents decide to buy (or sell) simultaneously, a price increase (decrease) of approximately 171% (65%) would ensue.

It is clear that if the subjects decided never to trade, one would have $I(t) \equiv 0$ and the price at the end of the session would be given by Eq. (4.1), which would lead to a most probable 640% increase in wealth. As we explain in the next section, the rational decision (individually

[§]Mathematically, the average of the exponential of a Student distribution is infinite, because of rare, but extreme values of η_t. In order not to have to deal with this spurious problem, we impose a cut-off beyond $|\eta| = 10$, with no material influence on the following discussion.

and collectively) would be to buy and hold shares until the end of the game. What we observe instead is that our agents cannot help trading in and out of the market. The trading activity is in fact so high in the first sessions that they barely break even, earning a meagre 0.75% on average. Remarkably, all groups learn to some extent and trade much less in the second sessions, leading to a much better average earning of 92%.[¶]

4.3 Rational Benchmark

We now devote our attention to the adequate rational benchmark for this experiment, which is a standard of comparison broadly used in Economics [91, 102, 105].

If the participants in our experiment were fully rational and assumed others to be rational as well, then all the agents would act in the same way if faced with the same information. In this scenario, there is no heterogeneity and we can limit our study to a single representative agent.

It is worth remarking at this point that this experiment is a *de facto* risk-free opportunity, in the sense that the subjects are paid at the end of the experiment if their net profit is positive but do not owe any amount if their net profit ends up negative.

4.3.1 Risk-neutral agents

Let the wealth of this agent be w_t at time t. If we assume that the session ends at $t = t_F$, purely rational agents have two possible strategies. The first one is to stay out of the market for $t > 1$ and hold cash until $t = t_F$ which yields an expected final wealth $\mathbb{E}\left(w_{t_F}\right) \approx w_0 \cdot \exp(-1)$ if the subjects start with shares and collectively pull out of the market (i.e. $I_1 = -1$) and $\mathbb{E}\left(w_{t_F}\right) = w_0$ if they start with cash.

[¶] In practice the average payouts were bigger because participants could not incur in losses, therefore negative contributions did not play a role in the actual average payout.

The second strategy consists in being fully invested in the market at $t = 1$ and hold the shares until $t = t_F$; this yields $\mathbb{E}(w_{t_F}) \approx w_0 \cdot (1 + m + s^2/2)^{t_F}$ regardless of the initial condition. The average outcome of the first strategy is at best zero profit, while the second strategy provides large average profits when the experiment session lasts for a long time.

Taking into account that these are the only two possible scenarios with a risk neutral population, it is obvious that the rational benchmark should be to enter or stay in the market at $t = 1$ and hold the shares until the end of the experiment session. The longer the duration of the experiment, the larger the expected profits.

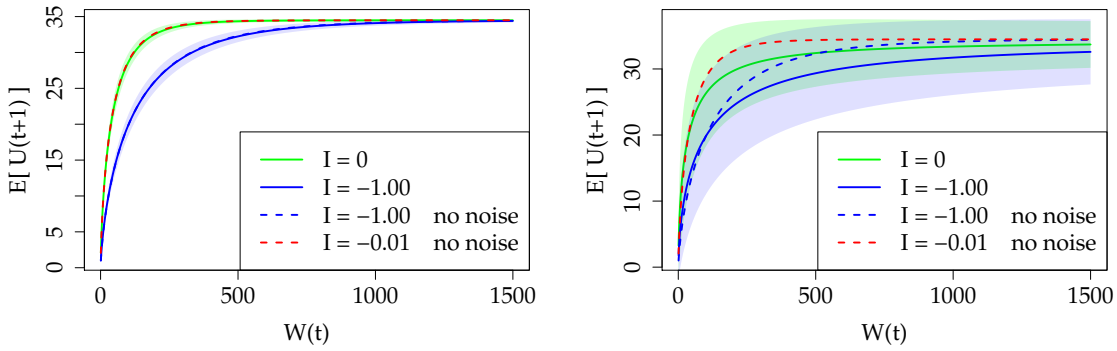
4.3.2 Myopic risk-aversion

Holt *et al.* showed in [67] that risk aversion in individual traders increases with their level of wealth and proposed the utility function

$$U(w) = \frac{1 - \exp\left(-\alpha w^{1-r}\right)}{\alpha}, \tag{4.4}$$

where α and r are positive parameters. [64, 65] are other notable studies concerning the measurement of risk aversion in a laboratory with human subjects.

Based on the fact that the utility function in Eq. (4.4) is concave, one could be tempted to consider that for high enough volatility s, myopic rational traders, i.e. traders who only look one time step ahead when making their decisions, would choose collectively to step out of the market as soon as their wealth reaches a certain level. However, they anticipate that if they all decided to sell, the price of the asset would be negatively affected by the impact factor I_t on top of the unknown noise term $s\eta_t$. Consequently, it is clear that at any given time step, it is always a better option to stay in the market and be affected by the noise factor alone, instead of selling and being affected by both the noise factor and the negative impact factor.



(a) RMS $s = 0.1$ **(b)** RMS $s = 1.0$

Figure 4.1: Average utility function of Eq. (4.4) for different scenarios with bounded rationality, volatility $s = 0.1$ (left) and $s = 1.0$ (right) and varying level of wealth $W(t)$. $I = 0$ refers to the scenario where all players stay in the market. The colored regions correspond to 1-σ around the average utility; this gives insight on the possible behaviour of risk-averse players.

We can however use Eq. (4.4) to study a few cases of bounded rationality within myopic optimization. We can think of players who disregard the rules of the game and act upon their instinct, i.e. just forget the noise term and sell their shares when their level of wealth reaches a critical level W^*. Another possibility is that one player tries to outsmart the others, believing that he will be able to sell his shares while everyone else stays in the market. We therefore compare, for a number of levels of wealth, the expected utility given by Eq. (4.4) in the following anticipated scenarios: all players hold to their shares, all players sell their shares taking into account their impact and the noise, all players sell their shares but forget about the noise term, one player only sells his shares and forgets about the noise term. These scenarios are compared in Fig. 4.1. The magnitude of the noise s defines most of the differences.

In Fig. 4.1 the solid lines represent the predictions for a population consisting of risk-neutral agents only. We confirm in this graphic representation what we argued above, i.e. these solid lines never intersect and it is always better to stay in the market (green solid line) than to step out of it collectively (blue solid line). However, when we consider bounded rationality and myopic risk-averse traders, the conclusion may be different, in particular when

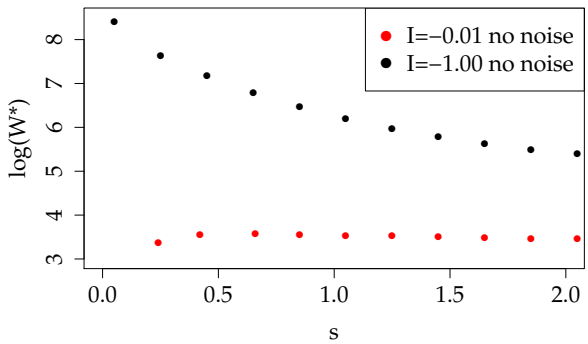


Figure 4.2: Critical threshold values W^*, as a function of s, beyond which the single-player-out (in red) and all-out strategies (in black) are triggered (corresponding, respectively, to the intersection point of the green/red dotted line and green/blue dotted line in the previous figure).

the agents forget about the random fluctuations of the market, which impact the price at each time step.

We analyse the expected utility of an agent who forgets about the random fluctuations of the market when he is the only one that sells (red dashed line) and when everyone sells (blue dashed line).

In these scenarios, there will be a value of wealth $W(t)$ above which, in the mind of these boundedly rational agents, it pays off to sell whatever they are holding. This point depends on the agent thinking either that he will be the only one selling or that everyone will, as well as on the magnitude of the random fluctuations or noise s. These values are critical thresholds which we represent as a function of s in Fig. 4.2 through numerical simulations. As we would expect from the concavity of the utility function in Eq. (4.4), the value W^* beyond which boundedly rational agents of this sort would sell decreases as a function of s when they consider the possibility of everyone selling at the same time.

In summary, we showed that rational agents performing myopic optimization would collectively buy and hold shares until the end of the experiment, while the assumption of bounded rationality allow for us to predict a critical value W^* beyond which traders would get nervous and sell. This does not, however, account for variability amongst the agents, which most cer-

tainly impacts the observed events in real markets and in our experiment.

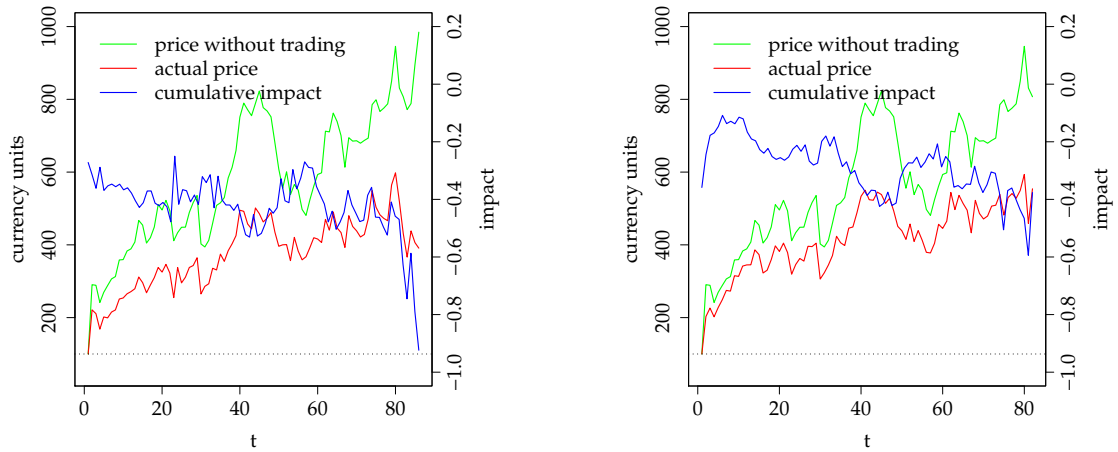
4.4 Results

4.4.1 Trading activity

Wealth and activity

We observe that the behaviour of subjects in second sessions is consistently different from their behaviour in first sessions, which is a result of learning. The Welch two-sample t-test applied to the final wealth and average activity rate of each subject in first and second runs statistically confirms this difference (p-value $\approx 2^{-16}$). Therefore, we merge all the first sessions into one data set and all the second sessions into another data set. These aggregated data sets lead to the price time series illustrated in Fig. 4.3, and immediately reveal that trading significantly weakens the upwards price trend: the average slope is divided by a factor of approximately 2 in the first sessions and 1.5 in the second sessions. The realized prices in Fig. 4.3 are proxies for the maximum earnings a subject would achieve if he used the buy-and-hold strategy within his group. However, the realized price log returns remain highly correlated with the "bare" price time series: the correlation coefficient is 0.85 in first sessions and 0.89 in second sessions. The higher values of the slope and of the correlation coefficient for the second sessions are due to a lower trading activity.

The positions of the traders – in or out of the market – are mostly intermittent, which implies excessive trading (Fig. 4.4). However, when the same subjects play for a second time, we see that some of them actually learn the optimal strategy, which translates into "green corridors" in Fig. 4.4b. Therefore, the distribution of average trading activity changes when the same set of people play the game for the second time, as shown in Fig. 4.5a and in Fig. 4.5b. The number of people keeping trading to a minimum increases significantly in second ses-



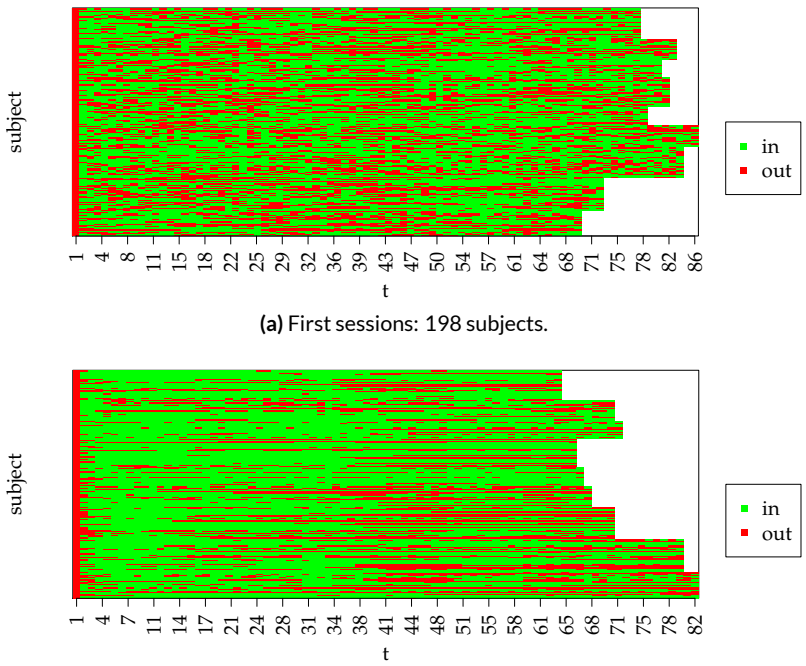
(a) First sessions, 198-18 subjects. The correlation between realized price log returns and price log returns in the absence of trading is 0.85.

(b) Second sessions, 201-23 subjects. The correlation between realized price log returns and price log returns in the absence of trading is 0.89.

Figure 4.3: Average or typical price time series for first sessions and second sessions.

sions, where only a few "hopeless outliers" keep trading activity above 40%, i.e. they changed their market positions in more than 40% of the periods. In both cases, the final wealth of the agents is strongly anti-correlated with average trading activity, which is expected since trading is costly. In fact, if a trader decides to buy shares at period t and to sell them at period $t+1$, he will, on average, end up with less cash than he started because of his own contribution to price impact. It is worth remarking that there is no price impact when the shares are liquidated at the end of each experimental session. This way, for a large number of periods, the average final wealth of a risk-neutral population would be approximately the same if the agents were given shares instead of cash in the beginning of the session.

The average wealth of the subjects in first and second sessions is shown in Fig. 4.6 as a function of time. We must not forget that the durations of the sessions were all different because of the indefinite time horizon. Thus, the number of data points used in the averaging process is not the same for each time t but a decreasing step function of t. This can be easily understood through Fig. 4.4a and Fig. 4.4b. In particular, this means that the last point in



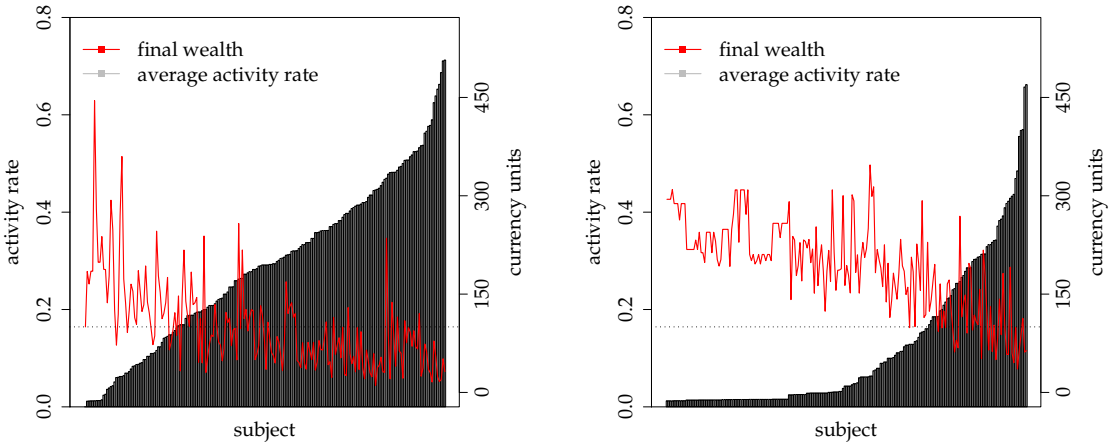
(a) First sessions: 198 subjects.

(b) Second sessions: 201 subjects.

Figure 4.4: Positions – in the market (green) or out of the market (red) – in first sessions (above) and in second sessions (below). It is worth noting that not all the sessions lasted the same number of periods, hence the white space in both figures for large t. In fact, for each case, only one session – the longest – lasted until the maximum time t displayed, $t = 86$ for first sessions (above) and $t = 82$ for second sessions (below).

Fig. 4.6a is not the average final wealth observed in first runs because all of the runs but one stopped before that point. The same applies to the last point in Fig. 4.6b.

In what concerns the average components of wealth over time, there are also differences between first sessions (Fig. 4.7a) and second sessions (Fig. 4.7b). In first sessions, where overall trading activity is high (Fig. 4.5a), the average wealth does not follow the upward trend one would expect in a set-up with "guaranteed" average growth of 2% per period. In fact, players trade so much that they keep eroding their wealth when they sell and affording fewer and fewer number of shares when they buy. This results in very low earnings at the end of the session, as we had shown before in Fig. 4.6a. On the other hand, in second sessions, where overall trading activity is much lower (Fig. 4.5b), the average wealth does increase with time.



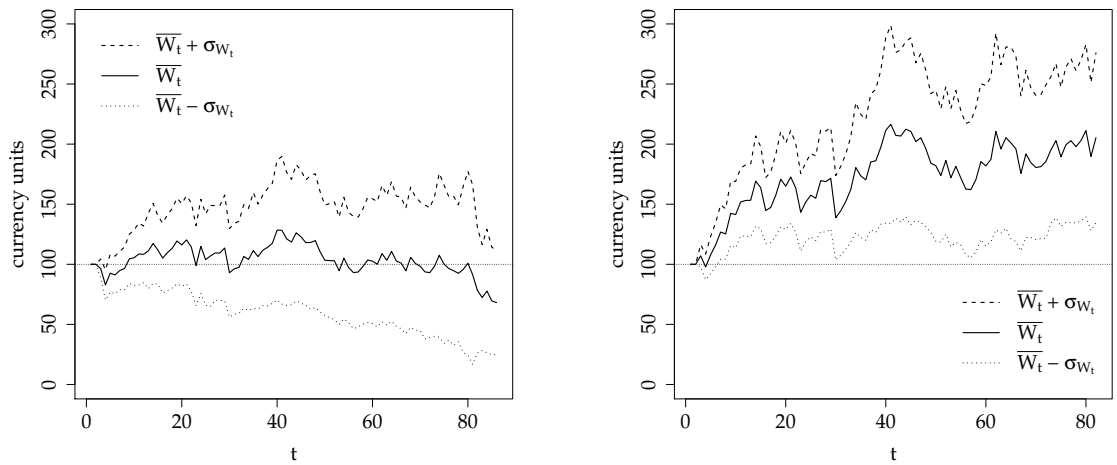
(a) First sessions: 198 subjects and 69-85 periods. The average final wealth was 100.75 units of currency, the average activity rate was 29% and the correlation between the two quantities was −0.62%.

(b) Second sessions: 201 subjects and 63-81 periods. The average final wealth was 191.97 units of currency, the average activity rate was 12% and the correlation between the two quantities was −0.73%.

Figure 4.5: Activity rate and final wealth in first sessions (left) and in second sessions (right).

Although the number of shares owned eventually decreases, this is due not only to excessive trading but also to the fact that some subjects hope for a similar noise pattern in the price time series, therefore they cash in their earnings before the end of the experiment and stay out of the market from that point onwards. This is particularly visible in Fig. 4.7b, when a surge in price triggers selling orders over several periods which result in a higher average amount of cash and, naturally, in a lower average number of shares. Although this is also visible in the middle of the time series in first sessions (Fig. 4.7a), the difference between the two cases is that most of the resulting cash is eventually reinvested in the first sessions, while in the second sessions this does not happen: cash holdings consistently increase after the initial investment phase (Fig. 4.7b).

As we showed in Sect. 4.3 the rational strategy in our experimental set-up would be to collectively buy-and-hold and reap the benefits from the baseline average return of 2% per period in the absence of trading. We see that the behaviour of the agents is very far from this



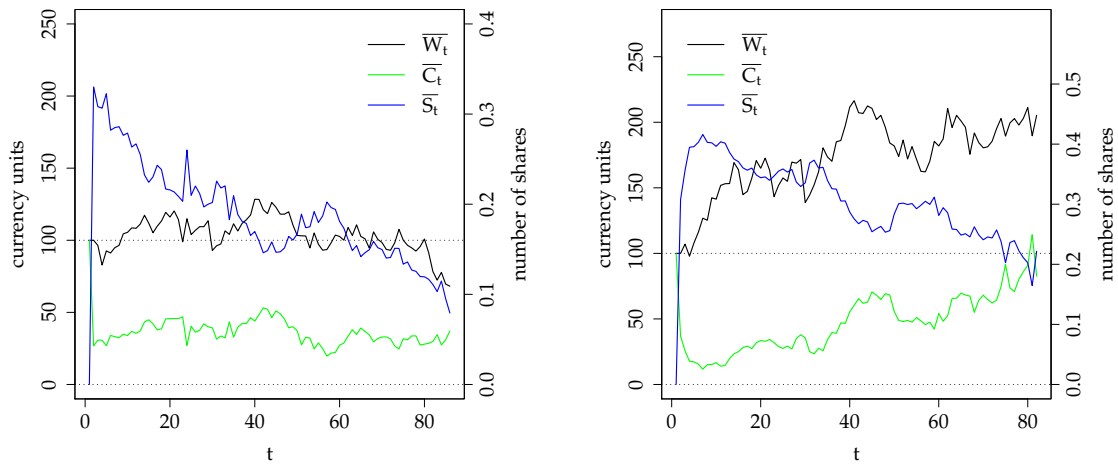
(a) First sessions, 198-18 subjects. **(b)** Second sessions, 201-23 subjects.

Figure 4.6: Average or typical wealth for first sessions and second sessions. The subjects fared much better in second sessions, in which their average final wealth was approximately twice their initial endowment, than in first sessions, in which the vast majority did not even break even.

benchmark, even in the second sessions, in spite of a significant decrease in activity. The average performance in second runs is indeed still far from what it would have been if everyone in the room used the optimal buy-and-hold strategy, i.e. despite the learning there is still excess trading activity which translates into collective pain – since even the virtuous agents are adversely impacted by the selling activity of wanton agents.

Collective trading modes – activity correlations, panic & euphoria

Our subjects surely trade too much, but can we describe in more detail why they trade and how correlated their activity is? In fact, our initial intuition – which turned out to be quite far from what actually happened — was that the agents would not trade at the beginning of the game, letting the price rise from its initial value of 100 to quite high values, say 400, before starting to worry that others might start selling, pushing the price back down and potentially inducing a panic chain reaction. This would have translated into either a major crash, or perhaps smaller downward corrections, but in any case a significant *skewness* in the



(a) First sessions, 198-18 subjects. **(b)** Second sessions, 201-23 subjects.

Figure 4.7: Average wealth $\overline{W_t}$ (black), average cash $\overline{C_t}$ (green) and average number of shares $\overline{S_t}$ (blue) in first sessions (left) and in second sessions (right). While $\overline{W_t}$ is consistently eroded over time in first sessions due to excess trading activity, it has a clear upward trend in second sessions, which can be explained by learning through experience.

distribution of returns – absent in principle from the bare price series which is constructed to be perfectly symmetrical, since the noise ξ in Eq. (4.2) is symmetric. In fact, as we will see below, the empirical skewness of the particular realization of the noise turns out to be negative, so the reference point that we shall be comparing to must be shifted.

We have therefore measured the relative skewness of the distribution of price changes, over different time intervals $\tau = 1, 2, \ldots$. The idea is that a panic spiral would lead to a negative skewness that becomes larger and larger when measured on larger time intervals, before going back down to zero after the typical correlation time of the domino effect. This is called the "leverage effect" in financial markets, and is observed in particular on stock indices where the negative skewness indeed grows as the time scale increases, before decreasing again, albeit very slowly [21].

In order to reduce the measurement noise, it is convenient to measure the negative skewness using two low-moment quantities. One is $P(r > m) - 1/2$, i.e. the probability that the price variation on time scale τ exceeds its average value. If this quantity is positive,

it means that large negative returns are more probable than large positive returns, so as to compensate the number of positive returns larger than the mean. Another often used quantity is the mean m_t of the returns minus the median, normalised by the RMS of the returns on the same time scale. Again, if the median exceeds the mean, the distribution is negatively skewed. See [21, 99] for further details about these estimators of skewness. Both quantities were found to give the same qualitative results, thus we chose to average these two definitions of skewness and plot them as a function of τ, averaged over all first and second sessions.

The result is shown in Fig. 4.8. The blue dots correspond exactly to the time series of bare prices because there is only one (collective) trade in the buy-and-hold strategy, right at the first period, which we discard from the computation. Although the bare returns were constructed using a Student's t-distribution with 3 degrees of freedom, which by definition is not skewed, we see in Fig. 4.8 that the bare prices do not have zero skewness. This illustrates the role of the noise, which gives way to different values of skewness for bare prices depending on the number of periods of the session. We observe in Fig. 4.8 that the realized skewness of trade impacted returns is typically larger (i.e. less negative) than bare returns, but without any significant time dependence. This suggests that buying orders tend to be more synchronized than selling orders, specially in the first sessions, but that neither buying nor selling orders induce further buy/sell orders. In short, there is no destabilising feedback loop in the present setting, which explains why we never observed any crash in our experiments; if anything, buy orders tend to be more collective than sell orders.

In order to detect more precisely the synchronisation of our agents, we construct the activity correlation matrix A, defined as follows:

$$\mathcal{A}_{ij} = \frac{1}{T}\sum_t \vartheta_i(t)\vartheta_j(t) - \frac{1}{T}\sum_t \vartheta_i(t) \times \frac{1}{T}\sum_t \vartheta_j(t), \tag{4.5}$$

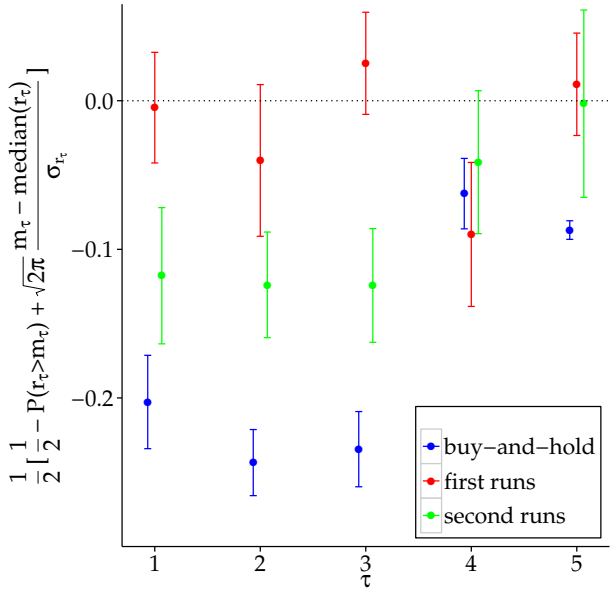


Figure 4.8: Average skewness of price log returns as a function of time τ in first sessions (red) and second sessions (green), together with the skewness of log returns in the buy-and-hold strategy, i.e. in the absence of trading (blue).

where $\vartheta_i(t)$ is the activity of agent i at time t, $\vartheta_i(t) = 0$ if he is inactive, $\vartheta_i(t) = \pm 1$ if he buys or sells.

For each session, we diagonalize A and study the three largest eigenvalues, corresponding to the more important principal components of the subjects' activity. In order to detect synchronisation, where a substantial fraction of agents tend to act in exactly the same way across the experiment, we compute the dot products of these three eigenvectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ and the uniform vector $\vec{e} = (1, 1, \ldots, 1)/\sqrt{N}$. Then, we take the absolute value and average the maximum over all runs. It may happen that the "synchronized" mode does not correspond to the largest eigenvalue of A, while still being amongst the most important ones. The resulting values are represented in Fig. 4.9 for the first and second sessions. The dashed lines depict the cases where agents would act completely at random, which would lead to a value of this overlap at approximately 0.35. We see that the experimental results are clearly larger than the benchmark case: approximately 0.57 for the first sessions and approximately 0.5 for

the second sessions (compared to a maximum value of 1 for a fully collective activity mode).

This method above allows us to make a quantitative analysis and statement about overall synchronization in our experiment, be it through selling or buying orders, and we find that there is indeed significant synchronisation. We now look into the separate cases of synchronization for selling orders and for buying orders. In order to do this, we construct an activity correlation matrix as in Eq. (4.5) but change the definition of $\vartheta_i(t)$ accordingly. This way, when we study the synchronization concerning only buying orders, as in Fig. 4.10a, we define $\vartheta_i(t) = 0$ if agent i is inactive or sells and $\vartheta_i(t) = 1$ if he buys. Likewise, in the case where we look into synchronization over selling orders only, as in Fig. 4.10b, we set $\vartheta_i(t) = 0$ if agent i is inactive or buys and $\vartheta_i(t) = -1$ if he sells. We see in Fig. 4.10a and in Fig. 4.10b that splitting the data set as explained does yield similar results: the experiment results are larger than the benchmark for each case, being the difference more marked in first sessions than in second sessions. Again, the synchronisation of buy orders appears to be, according to this metric, stronger than that of sell orders.

We therefore conclude that even if our subjects cannot directly communicate with one another, there is a significant synchronisation of their activity, in particular during the first sessions and, as the skewness of the distribution reveals, for the buying activity. The mechanism for this synchronisation can only come from the common source of information that the subjects all observe, namely the price time series itself.

Furthermore, we observe an asymmetry if we repeat the above method conditionally on the sign of previous returns, in the sense that synchronisation is stronger for buying orders conditional on negative previous returns and for selling orders conditional on positive previous returns. This indicates mean reversion, which is in line with our findings in subsection 4.4.3.

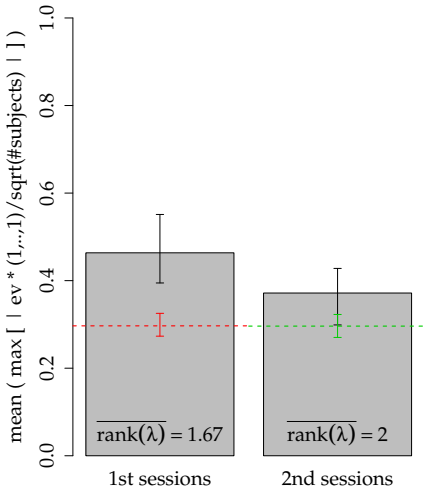
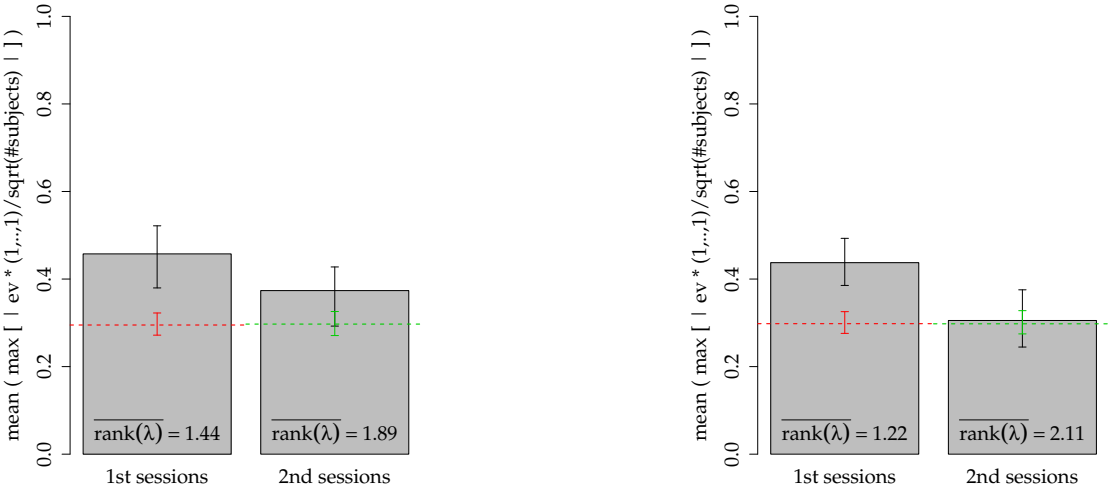


Figure 4.9: Average maximum absolute value of the three dot products between the eigenvectors corresponding to the three largest eigenvalues, and the unit vector. All orders are considered except for the first time step, in which we expect a natural bias towards synchronization.



(a) Buying orders. **(b)** Selling orders.

Figure 4.10: Average maximum absolute value of the sum of the components of the eigenvectors corresponding to the three largest eigenvalues. On the left we restrict ourselves to buying orders, while the result for selling orders is shown on the right. The first time step is excluded from the data set because we expect a natural bias towards synchronization.

Clustering

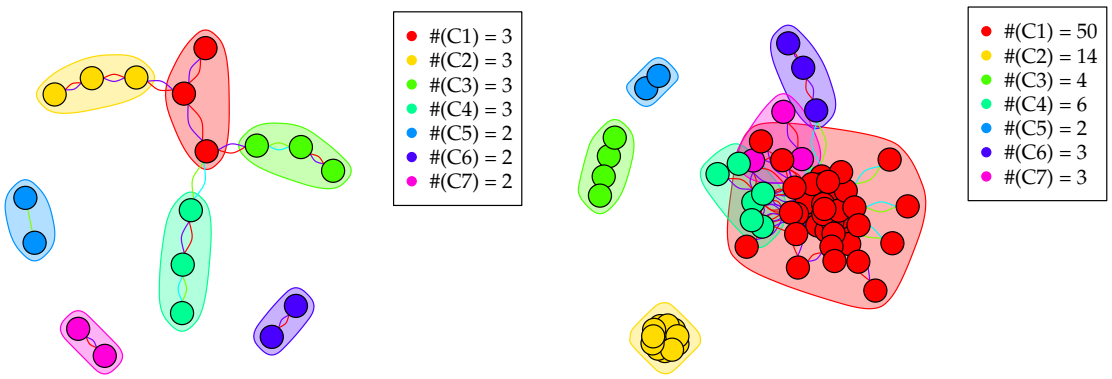
Fig. 4.5 shows us, once again, that the average final wealth in first sessions is much smaller than in second sessions, which is tied to the higher average trading activity of the subjects when they play the game for the first time. In second sessions, we observe a number of subjects who kept trading activity very low, increasing their chances of a positive payout at the end of the experiment. As we discussed in Sect. 4.4.1, this is an indication that the subjects learn. In any case, there are always traders who keep trading at very high rates and lose money in the process.

However, Fig. 4.5 does not provide insight about common patterns in the behaviour of the subjects. We know from Fig. 4.4 that at least in second runs a number of subjects use the buy-and-hold strategy or similar, which corresponds to the green horizontal "corridors" in the figure. Therefore, we apply clustering techniques and find groups of similar traders in the data sets. Afterwards, we look into the trading activity and trading performance in each cluster.

As in [110], we applied false discovery rate (FDR) methods to validate links between subjects to the data set with composite data from first sessions and to the data set with composite data from second sessions. The variable used to establish links (i.e. similarity) between subjects was their position – in or out of the market – over time for each subject. The FDR rejection threshold was 1%.

The number of subjects in each cluster is summarized in Tab. 4.1 for first sessions and second sessions. The clusters identified in the second sessions are much larger than those identified in first sessions, which is expected because the number of "intermittent" players in the game was lower in second sessions.

We see in Fig. 4.12 that clusters with a lower average trading activity tend to have a higher



(a) First sessions **(b)** Second sessions

Figure 4.11: Clusters for first and second sessions using the FDR algorithm of [110][109] with a threshold of 1% applied to positions – in or out of the market – over time.

Cluster	1	2	3	4	5	6	7
N	3	3	3	3	2	2	2

(a) First sessions

Cluster	1	2	3	4	5	6	7
N	50	14	4	6	2	3	3

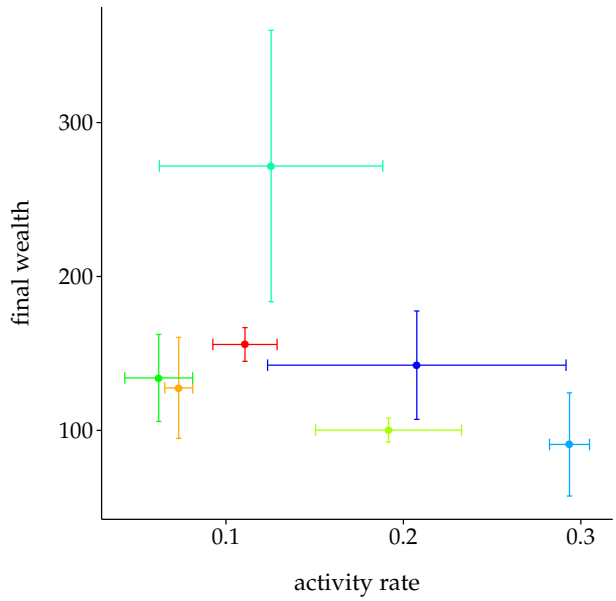
(b) Second sessions

Table 4.1: Number of subjects in each cluster identified through an FDR algorithm with a threshold of 1% applied to positions – in or out of the market – over time.

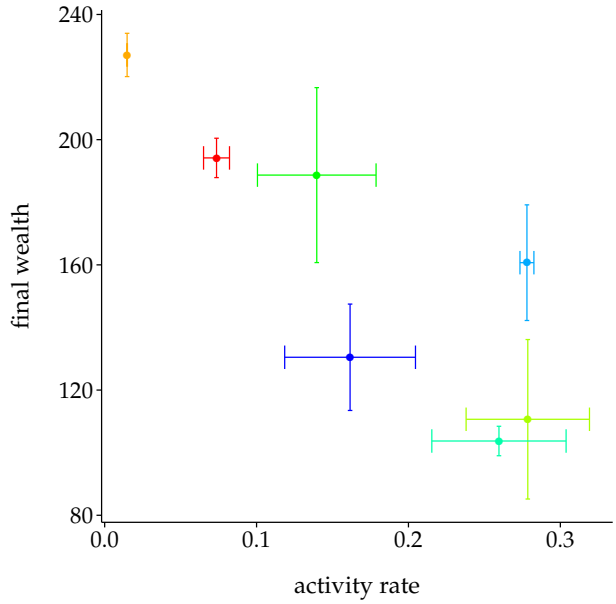
average final wealth. A notable example is the cluster number 2 in Fig. 4.12b, which includes 14 subjects (see Fig. 4.11b) who kept trading to a minimum in second runs and maximized their returns. Conversely, the cluster number 6 in Fig. 4.12a consists of 2 traders (see Fig. 4.11a) with very high average trading activity and, as a consequence, low final wealth.

4.4.2 Lottery and risk aversion

In order to assess the risk aversion of the subjects, we presented them with a choice of lotteries as in [67]. The subjects had to choose, for 10 pairs of lotteries, if they preferred the risky or the safe bet. For each pair of lotteries, the average pay-off was not the same for both options available (risky or safe). This way, a risk-neutral agent would simply calculate the average pay-off for each lottery in the pair and choose the one yielding the highest average pay-off.



(a) First sessions



(b) Second sessions

Figure 4.12: Average trading activity and final wealth each cluster represented in Fig. 4.11.

However, a risk-averse agent would also factor in the probability of walking away with almost nothing – one of the possible outcomes in the riskiest lotteries.

In fact, we observe that, for each pair of lotteries, the agents did not choose the risky option whenever its expected pay-off was larger than that of the safe option. This shows that the subjects were indeed risk-averse. If we order the lottery pairs according to the expected pay-off of the risky lottery, we can rank the agents according to their risk aversion based on the pay-off level beyond which they choose the risky bet instead of the safe bet. The higher this level, the more risk-averse is the agent.

Besides, we also wanted to assess whether or not the risk aversion of the subjects depended on the absolute values involved, as in [67], where the authors find a strongly concave utility function. Therefore, we repeated the above with the exact same lottery choices but with all the rewards multiplied by 5.

We discarded from our data set the cases in which the subjects chose the safe lottery after having previously chosen the risky option for a lower pay-off advantage.

The interested reader is referred to Appendix B.1, in which we included all the instructions handed to the subjects during the experiment, in particular the aforementioned lottery choices.

In Fig. (4.13) and (4.14) we show the distribution of subjects according to their risk aversion, both for low and high rewards. The blue vertical line shows where a risk neutral subject would be, based on the expected pay-off differences alone. The fact that the majority of the subjects – 76% for the low-pay-off and 89% for the high pay-off lotteries – have a risk aversion larger than 4 indicates that the participants in our experiment were risk-averse. Moreover, if we compare the two curves, red and green, we see that the subjects tend to safer choices when the lottery pay-off is higher, which indicates that the risk aversion of our population not only depends but increases with the pay-off level. This is corroborated by the Welch two-sample

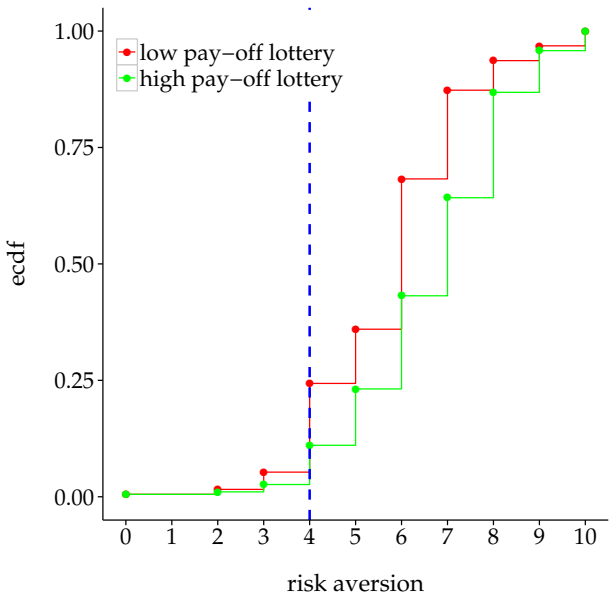


Figure 4.13: Risk aversion cumulative distribution function. The x-axis is the pay-off level beyond which agents choose the risky bet. The higher this level, the more risk-averse is the agent.

t-test, which states that the mean risk aversion values for low pay-off and high pay-off are statistically different with a p-value of $1.5e^{-06}$. These results are perfectly in line with what Holt and associates found in [67].

Finally, we see in Fig. (4.15) that the activity rate decreases and the average final wealth increases with the level of risk aversion, both in second runs (green and blue) and in first runs (red and orange). This is confirmed in Fig. (4.16), which shows the corresponding linear regressions (using weights $\sqrt{N}/\sigma$) with 95% intervals of confidence.

Utility function calibration

As in [67], we use the lottery choices of the subjects to calibrate the utility function described in Eq. (4.4). In short, we apply a maximum-likelihood method to find the parameters which maximize the probability that the observed lottery choices are dictated by Eq. (4.4).

In this spirit, the first step we take is to model the probability P_i^R that a subject chooses the

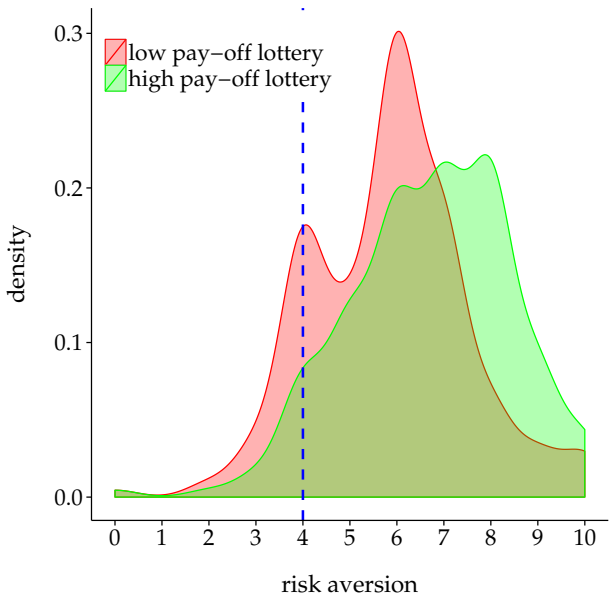


Figure 4.14: Risk aversion probability density estimation, normal kernel.

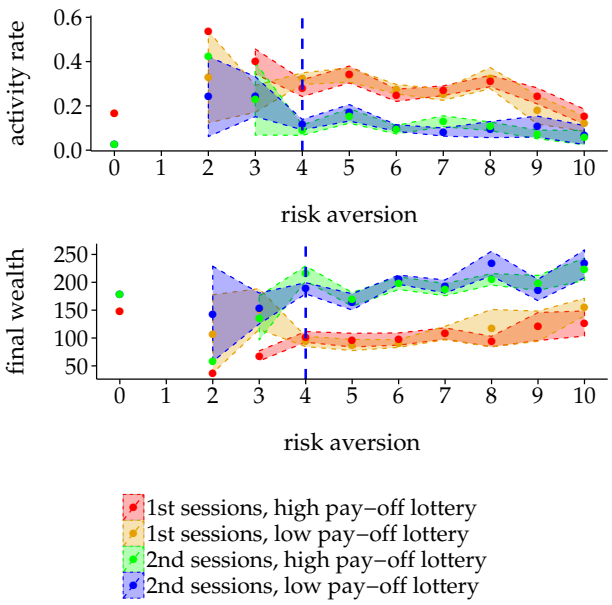


Figure 4.15: Average final wealth and activity rate for each level of risk aversion, for first and second runs.

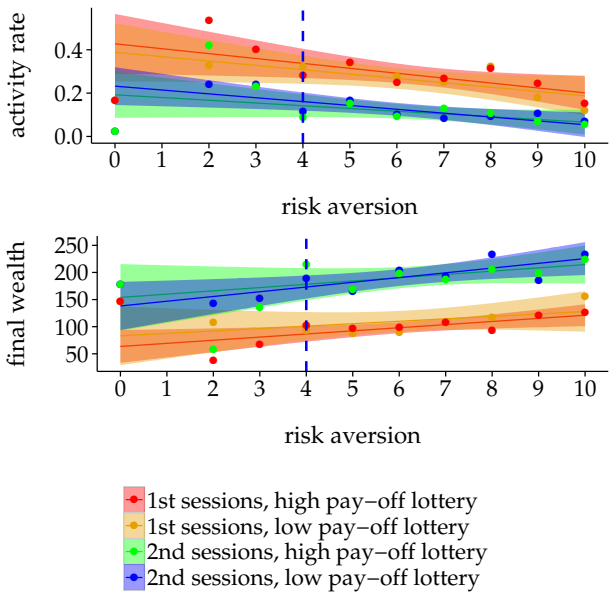


Figure 4.16: Linear regressions with 95% intervals of confidence of the average final wealth and activity rate on risk aversion, for first and second runs.

risky lottery out of the pair i of lotteries. As in [67], we define

$$P_i^R := \frac{E[U_i^R]^{\frac{1}{\mu}}}{E[U_i^R]^{\frac{1}{\mu}} + E[U_i^S]^{\frac{1}{\mu}}}, \tag{4.6}$$

where $E[U_i^R] = \sum_{k=1}^{2} p_k U_k^R$ is the expected utility of the risky lottery and $E[U_i^S]$ is the expected utility of the safe lottery in the pair i. Each lottery has two possible outcomes $k = 1, 2$, each with probability p_k and utility U_k given by Eq. (4.4), parameterized by two numbers α and r. The parameter μ is a real number which allows one to consider a range of scenarios between equiprobable choices ($\mu = +\infty$) and utility maximization ($\mu \to 0$). This corresponds to the so called "logit rule" (see [19] or [6] for further details).

Secondly, we define the likelihood function

$$\mathcal{L}(\beta, y) = \prod_i (P_i^R)_i^y \cdot (1 - P_i^R)^{1-y_i}, \tag{4.7}$$

	average	confidence interval
α	0.106	$[0.085, 0.130]$
r	0.345	$[0.263, 0.443]$
μ	0.114	$[0.101, 0.133]$

Table 4.2: Parameters of Eq. (4.6) obtained via maximum-likelihood estimation and correspondent 95% confidence intervals. We applied the Nelder-Mead algorithm to maximize Eq. (4.8) and used the bias-corrected and accelerated (BCa) bootstrap method for the 95% confidence intervals.

where y_i are the observed choices for each lottery pair, i.e. $y_i = 0$ if the subject chose the safe lottery and $y_i = 1$ if he chose the risky lottery from pair i. In addition, $\beta = [r, \alpha, \mu]$ includes all the model parameters in Eq. (4.6).

This way, we have the log-likelihood function

$$\log[\mathcal{L}(\beta, y)] = \sum_i y_i \log(P_i^R) + (1 - y_i)\log(1 - P_i^R). \tag{4.8}$$

Finally, the last step is to find the model parameters that maximize Eq. (4.8). We apply the Nelder-Mead algorithm [92] and use the bias-corrected and accelerated (BCa) bootstrap method [45] for 95% confidence intervals. The results are summarized in Tab. 4.2. Quite remarkably, the values of the a-dimensional parameters r and μ are found to be very close to those reported in [67] for their lottery experiments ($\mu = 0.13$, $r = 0.27$). In particular, $r > 0$ implies increasing relative risk aversion and decreasing absolute risk aversion.

4.4.3 Price predictions

The fact that the subjects input their price predictions allows us to have a glimpse of their frame of mind. In both runs, the subjects did not input anything in about 7% of the time, as price prediction was not a mandatory activity; in the following, we restrict our analysis to the subjects that did report their predictions.

The discussion focuses on the predicted log returns, i.e. from subject i's price predictions

$\widehat{p}_i(t+1)$, we compute the predicted log returns $\widehat{r}_i(t+1) = \log[\widehat{p}_i(t+1)/p(t)]$, for all subjects. The average expectation in the first run is -0.01 and 0.02 in the second run. The percentage of positive predicted returns is 54% in the first run, and 58% in the second run. Figure 4.17 illustrates the full empirical distribution functions of expectations for both runs. The starting point for each the positive and negative distributions represents the fraction of positive and negative expecations, thus the higher jump at $r = 0$ for the second run reflects the increase in the fraction of positive expected returns.

We then check how the distributions of positive and negative expected returns are related to the Student noise term given in input. This gives a first clue about the type of extrapolation from past returns that the agents use: if they use linear extrapolation rules, the power-law tails of the Student distributions will be found again in the distributions of return expectations. On the other hand, fear or greed may lead to non-linear extrapolations and thus will modify either the tail exponent of these distributions, or even the nature of the tails.

The most obvious finding is that the actions of the agents increase the volatility of the baseline signal (in dashed lines) as the empirical distribution functions are above the baseline signal for both runs. The amplification of the noise for positive expectations are almost the same in the two runs. The scale of negative expectations is however markedly different in the two runs: it was much larger during the first run. Robust power-law tail fitting that determines the most likely starting point of a power-law r_{min} and that yields the exponent α yields parameters reports in Table 4.3. Quite remarkably, the parameters of the positive and negative tails are simply swapped between the two runs: thus not only the scale of negative expectations changes, but the nature of largest positive and negative expectations also changes. The fitted tail exponent is not far from 3, the one of the Student noise.

The fact that the subjects have fat tailed predictions means that they form their predictions by learning past returns, which do contain fat tails because of the Student's t-distributed

	r_{min}	α
run 1 $r > 0$	0.21	3.5
run 2 $r > 0$	0.10	2.7
run 1 $r < 0$	0.13	2.6
run 2 $r > 0$	0.18	3.5

Table 4.3: Fits of the power-law part of return expectations; r_{min} denotes the most likely starting point of the power-law.

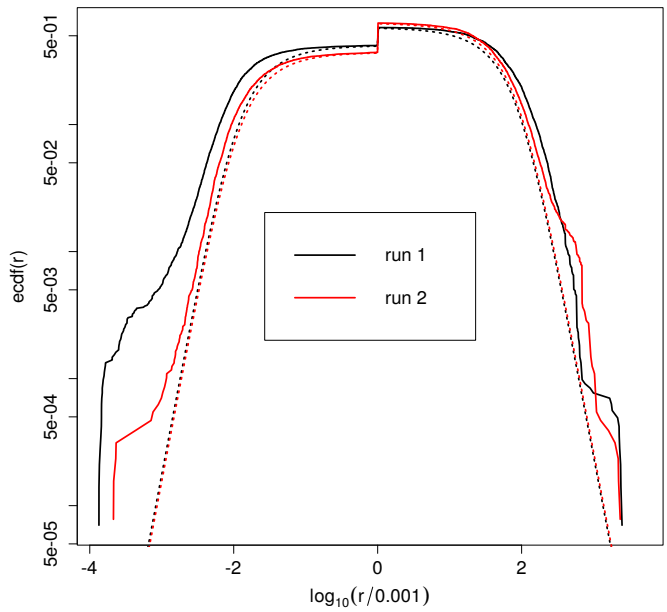


Figure 4.17: Reciprocal empirical cumulative distribution functions of negative (left) and positive (right) expected price returns during the first and second runs (black and red lines, respectively). The baseline return distribution is plotted in dashed lines. The jump at $r = 0$ indicates the respective fraction of positive and negative predictions.

noise. One thus expects some relationship between predicted returns and past returns. This is in line with the best established fact about real investors, which is the contrarian nature of their trades: their net investment over a given period is anti-correlated with past price returns [26, 60, 72, 75]. In addition, previous experiments [70] have demonstrated that four simple classes of linear predictors using past returns are usually enough to reproduce the observed price dynamics.

This suggests to fit the return predictions with a simple linear relationship:

$$\widehat{r}_i(t+1) = \omega_0 + \omega_1 r(t), \tag{4.9}$$

where ω_0 and ω_1 are two real fitting parameters.

Section 4.4.3 discusses ω_0 while Sect. 4.4.3 is devoted to ω_1.

Average predictions (ω_0)

Trades only tell about the consequence of the state of mind (i.e., the price expectation) of traders when they are active. But traders (both in real life and in experiments) are in fact inactive most of the time. As a consequence, trades alone are unlikely to be able to explain why traders are inactive. Since we have both trades and subject price expectations, we are able to give a consistent picture of activity and inactivity as a consequence of price return expectations.

The key is to compute return expectations conditionally on the actions of the subjects. There are four possible actions: buying, selling, holding shares and holding cash. Figure 4.18 reports the conditional distributions of price return predictions, for both runs. The results are qualitatively the same for both runs: the conditional distributions are clearly separated; the main difference between the two runs is that the variance of expectations among the pop-

ω_0, 1nd run	Sell	Hold cash	Hold shares
Buy	1.5e-8	3.9e-3	1.2e-5
Sell		3.8e-5	8.2e-5
Hold cash			1.5e-1

Table 4.4: p-values of Mann-Whitney tests between the average return predictions of all state pairs for first runs.

ω_0, 2nd run	Sell	Hold cash	Hold shares
Buy	2.1e-7	1.5e-4	1.0e-7
Sell		6.1e-7	5.9e-5
Hold cash			6.6e-6

Table 4.5: p-values of Mann-Whitney tests between the average return predictions of all state pairs for second runs.

ulation is much reduced during the second run.

Let us break down the results for each possible action.

1. When the subjects hold assets, their expectations are in line with the baseline return of 2%.

2. When the subjects hold cash, their expectations are significantly lower (essentially zero).

3. When the subjects make a transaction, however, their expectations of the next returns are anti-correlated with their actions, i.e., they buy when they expect a negative price return and vice versa.

Thus, the actions of the subjects are fully consistent with their expectations: they do not invest when they do not perceive it as worthwhile and they keep their shares when they have a positive expectation of future gains. Interpreting their expectations when they trade is less straightforward, however. The fact that their actions are anti-correlated with their expectations means that they try to exploit a perceived price anomaly and thus either minimize their price impact, obtain a rebate, or both.

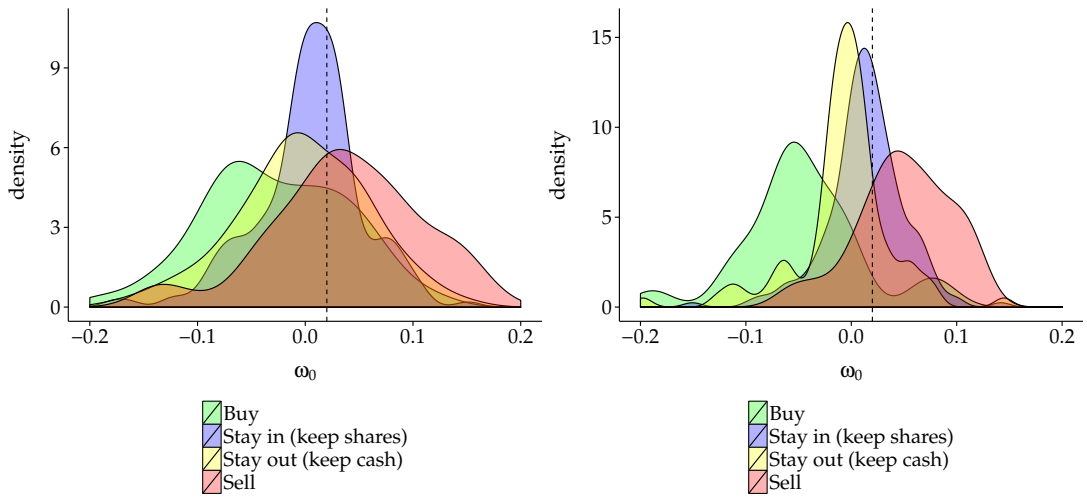


Figure 4.18: Conditional densities of average return predictions ω_0 during the first runs (left plot) and second runs (right plot) for the four types of decisions.

ω_1, 1st run	Sell	Hold cash	Hold shares
Buy	9.6e-2	2.3e-5	8.8e-10
Sell		3.7e-2	1.7e-4
Hold cash			2.9e-2

Table 4.6: p-values of Mann-Whitney tests between the linear coefficient w_1 between all state pairs for first runs.

Predictions and past price returns (ω_1)

Coefficient ω_1 encodes the linear extrapolation of the past return on future returns. Figure 4.19 shows that during the first run, this coefficient was negative when the agents did not act and zero when they did trade. The second run is different: the coefficients do not seem to depend much on the state, the only clear difference is betwen holding cash and holding shares.

The lack of influence of this coefficient is confirmed when one measures the average predicted return conditional on the action of the subjects, which gives results very close to ω_0.

We could not find any dependence of $\omega_{0,1}$ on the lottery results.

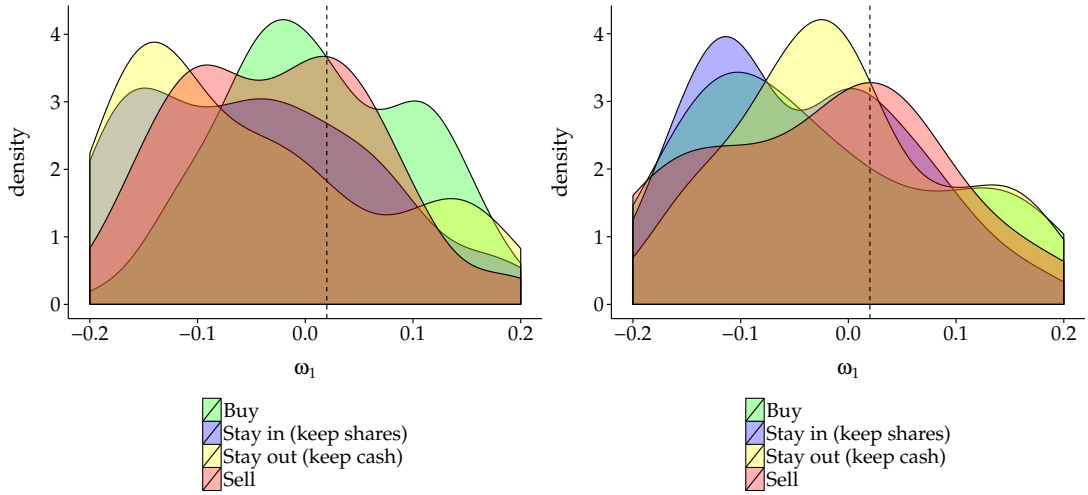


Figure 4.19: Conditional densities of average return predictions ω_1 during the first runs (left plot) and second runs (right plot) for the four types of decisions.

ω_1, 2nd run	Sell	Hold cash	Hold shares
Buy	6.2e-1	7.6e-1	4.1e-2
Sell		5.6e-1	8.6e-2
Hold cash			2.9e-4

Table 4.7: p-values of Mann-Whitney tests between the linear coefficient w_1 between all state pairs for second runs.

4.5 Conclusion

We presented the results of a trading experiment in which the pricing function favours early investment in a risky asset and no posterior trading. As explained in Sect. 4.2 and in Sect. 4.3, the subjects would make an almost certain gain of over 600% if all of them used the risk-neutral, rational strategy, i.e. if they all bought shares in the first period and held them until the end of the experiment.

However, market impact as defined in Eq. (4.3) acts *de facto* as a transaction cost which erodes the earnings of the traders. When the subjects participate in the experiment for the first time, their trading activity is so high that their profits average to almost zero! They are however found to fare much better when they repeat the experiment as they earn 92% on average – which is still much below the performance of the simple risk-neutral rational strategy outlined above. We therefore find that *investors do trade too much*, even in an environment where trading is clearly detrimental and buy-and-hold is an almost certain winning strategy (at variance with real markets where there is nothing like a guaranteed average return of 2% per period).

At each period, we also ask the subjects to predict the next price of the asset. This provides us with additional information about our controlled artificial market, i.e. we have access not only to the decisions of each trader but also to their expectations. It is important to emphasize that this information would not be available in broker data. In fact, knowing the expectation behind each decision of each trader – including the decisions to do nothing – is one of the advantages of human experiments compared to empirical analysis of real data. Using this information, we confirm in Sect. 4.4.3 that the traders in our experiment have a contrarian nature, which, together with the pattern of excessive trading, is one of the known features of individual traders in real financial markets, as discussed in [40] and [26]. We find it very

interesting that we observe these known features in such a controlled environment.

As in [67], we also use lotteries at the end of the experiment to confirm that the subjects are not risk-neutral but risk-averse. In addition, we observe that their relative risk aversion increases with the pay-off level in a way that is quantitatively similar to the results reported in [67]. When put in correspondance with the results of the trading experiment, we observe that the activity rate decreases and the average final wealth increases with the level of risk aversion of the subjects.

In addition, we use false discovery rate algorithms to show that the traders clustered in terms of positions – in or out of the market – over time, despite the fact that they do not communicate with each other during the experiment, which implies that the similarities in trading patterns arise because certain traders react in the same way to the common price time series. Unsurprisingly, the clusters identified in second sessions are much larger than those in first sessions, while clusters corresponding to traders with a lower average trading activity typically have a higher average final wealth.

We find through the distribution of predictions of price log returns that the subjects are not rational but still act in a way which is fully consistent with their expectations. This happens because their expectations depend on whether they are invested or not, as well as on their trading actions.

Contrary to what happens in real financial markets, we have not observed any "leverage effect" (i.e. an increase of volatility after down moves). Although there is a clear detrimental collective behaviour in all the sessions of this experiment, we do not witness any big crash or avalanche of selling orders that would result from a panic mode. As we discuss in Sect. 4.4.1, the coordination amongst traders was actually slightly stronger when buying than when selling, resulting in a positive skewness of the returns. In order to induce crashes, the duration of the experiment could be expanded while the time available for each decision could be re-

duced. The former would increase the probability of a sudden event by increasing the number of trials and the amounts at stake, while the latter could contribute to higher stress levels amongst subjects and increased sensitivity to price movements. However, we believe that a more efficient path to generate panic would be to reduce the "normal" volatility level while increasing the amplitude of "jumps" in the bare return time series. Within the current setting, large fluctuations do not seem surprising enough to trigger panic among our participants. Another idea, perhaps close to what happens in financial markets, would be to increase the impact of sell orders and reduce the impact of buy orders when the price in high, mimicking the fact that buyers are rarer when the price is high (on this point, see the recent results of Donier et al. on the Bitcoin [42].)

Other natural extensions of this experiment would include the possibility of fraction orders and hedging, as well as short selling. It is however essential to keep in mind that any increase in experimental design complexity should by followed by an increase in the number of collected experimental points. Otherwise, it will be difficult if not impossible to disentangle cause and effect from stochastic noise in the experimental data.

5

Conclusion and future work

This thesis explores two distinct yet related subjects pertaining to the larger topic of collective destabilising phenomena in socio-economic systems.

In chapter 3 the focus is on a theoretical model of dynamics of trust in networked societies inspired by the aftermath of the 2008 crisis, while chapter 4 discusses the results of a trading experiment which aims at a better understanding of how real people behave in a realistic trading set-up.

The basic assumption of the model of dynamics of trust in networked societies is that link

creation is boosted when trustworthiness is high and negatively impacted by trustworthiness heterogeneity across nodes. As shown in chapter 3, there are regions of the parameter space where two equilibria coexist: a sparse low-confidence network and a dense network where trustworthiness abounds. It is worth remarking that the transitions between states are triggered by small fluctuations in trustworthiness over time. In the case of a real network, the actual parameters fitting the observed data could help understand and mitigate the mechanisms through which the jumps between different equilibra occur. In the particular case of an interbank market network, the calibration method could use CDS data or survey results as a proxy for trustworthiness and interbank loans as the links between nodes. These data can be found in databases of central banks across the world and supervisory authorities could benefit from early warning signals for potential trust collapses and crisis, which would be a natural extension and application of the work presented in chapter 3.

In the trading experiments explained in chapter 4 the most important conclusion is that individual investors trade excessively and do not typically follow the risk-neutral, rational strategy, even in an incredibly advantageous set-up that guarantees them an average return of 2% per period, for approximately 100 periods, in the absence of trading. One should stress that no real financial market or product offers its participants a guaranteed 2% average return per period over 100 consecutive periods. Unfortunately for the subjects, their actions turn a potential 600% average profit into a meagre 0.75% average profit. In addition, the data collected on price predictions indicates that the traders are contrarian, which is in line with the best established fact about real investors. The application of false discovery rate algorithms reveals that the traders cluster in terms of positions – in or out of the market – over time because they react similarly to the same input (prices). As expected, two quantities which vary significantly across clusters are the average final wealth and the average trading activity, which are negatively correlated. Contrary to real financial markets, we do not witness any

"leverage effect". Moreover, we do not observe any big crash or avalanche of selling orders as a consequence of panic. As suggested in chapter 4, lowering the volatility of external fluctuations and imposing less frequent but more severe downward shocks could trigger the cascade events which we did not observe in this set of trading experiments. In addition, another future avenue to explore would consist in making the price impact function depend on the price of the asset, which is close to what happens in real financial markets. Finally, one could also make the experiment more realistic by allowing for fractional orders, hedging and short selling, which would however require a higher number of collected experimental points per session.

A

Sudden trust collapse in networked societies

A.1 Model specifications

The adjacency matrix at time t is denoted by $J_{ij,t}$, while the trustworthiness of node i at time t is given by $h_{i,t}$. N is the total number of nodes in the network and $k_{i,t}$ is the degree of node i at time t, i.e., $k_{i,t} = \sum_j J_{ij,t}$.

At each time step t, the links between nodes are updated first. Then, the new trustworthiness of each node is computed.

Therefore, the evolution of the system at each time step happens in two distinct steps as follows.

1. Create, destroy, or leave $sN,\ s \in]0,1]$, links untouched:

$$P(J_{ij,t} = 0 | J_{ij,t-1} = 1) \quad =: \quad \Pi_{ij}^{+} = \frac{1}{1 + z_{ij,t-1}} \tag{A.1}$$

$$P(J_{ij,t} = 1 | J_{ij,t-1} = 0) \quad =: \quad \Pi_{ij}^{-} = \frac{r}{N} \frac{z_{ij,t-1}}{1 + z_{ij,t-1}}\,, \tag{A.2}$$

where

$$z_{ij,t} = \exp\left[\alpha \overline{h_t} + \alpha'(h_{i,t} + h_{j,t} - 2\overline{h_t}) - \beta |h_{i,t} - h_{j,t}|\right] \text{ and } r \in \mathbb{R}^{+}. \tag{A.3}$$

2. Update the trustworthiness values h_i:

$$h_{i,t} = h_{i,0} + f k_{i,t} \tanh\left(\frac{1}{c k_{i,t}} \widetilde{P_{i,t}}\right) + d \cdot \min\left(0,\ \delta h_{i,t}\right), \tag{A.4}$$

where:

$$\delta h_{i,t} = h_{i,t-1} - h_{i,t-2} + f\left[k_{i,t} \tanh\left(\tfrac{1}{f k_{i,t}} \widetilde{P_{i,t}}\right) - k_{i,t-1} \tanh\left(\tfrac{1}{f k_{i,t-1}} P_{i,t-1}\right)\right],$$

$$\widetilde{P_{i,t}} = \textstyle\sum_j J_{ij,t} h_{j,t-1},$$

$$P_{i,t} = \textstyle\sum_j J_{ij,t} h_{j,t},$$

and $f, d \in \mathbb{R}^{+}$.

Regarding the first step, it is worth remarking that $\lim_{N\to\infty} \Pi_{ij}^{+} \cdot N^2 \propto N$, which implies that the number of new links per node remains finite even for large N. $z_{ij,t}$ is a measure of the propensity of nodes i and j to link or remain linked at time t, which we assume to increase with $\overline{h_t}$ and $h_{i,t} + h_{j,t} - 2\overline{h_t}$. On the other hand, we consider that $z_{ij,t}$ is bigger if $|h_{i,t} - h_{j,t}|$ is smaller, i.e., that the likelihood of node i linking with node j increases with the similarity of their perceived trustworthiness in the community (homophily).

The term $\min(0,\ \delta h_{i,t})$, with its intrinsic asymmetry, is a proxy for the panic sentiment mentioned in the main text. Besides, $\widetilde{P_{i,t}}$ is the tentative cumulative trustworthiness of the peers of node i at time t, while $P_{i,t}$ is the actual value.

We can view the parameter s as a mere refresh rate in the algorithm but we can also interpret it as a measure of overall communication intensity between nodes.

A.2 Panic factor d and stability

Let us consider the case where node i ends up without any links at time $t_L + 1$. Moreover, let us assume that $z_{ij,t}$ is small enough for us to neglect new links involving node i as per Eq. (A.1). For the sake of simplicity, let us define $\tau := t - t_L$. In this notation, node i has at least one link at $\tau = 0$ and becomes disconnected from the rest of the network at $\tau = 1$. Moreover, let us define $h_n := h_{i,\ \ +n}$ and $h_{\text{init}} := h_{i,0}$.

Then, we have from Eq. (A.4) that

$$h_n = h_{\text{init}} + d(h_{n-1} - h_{n-2}), \quad n \geq 2. \tag{A.5}$$

In this scenario, Eq. (A.5) defines the fate of node i, as it determines whether its trustworthiness $h_{i,t}$ enters an infinite downfall or not.

Equation (A.5) can be re-written as

$$\mu_n = \Delta\mu_{n-2} + \nu, \tag{A.6}$$

where:

$$\Delta = d \begin{bmatrix} d-1 & d \\ 1 & -1 \end{bmatrix} \tag{A.7}$$

$$\mu_n = \begin{bmatrix} h_n & h_{n-1} \end{bmatrix}^T \tag{A.8}$$

$$\nu_n = h_{\text{init}} \begin{bmatrix} d+1 & 1 \end{bmatrix}^T. \tag{A.9}$$

After some computations, Eq. (A.6) becomes

$$\mu_{2n+1} = \Delta^n(\mu_1 - v) + v \tag{A.10}$$

$$\mu_{2n+2} = \Delta^n(\mu_2 - v) + v, \tag{A.11}$$

where $v = h_{\text{init}}[\ 1 \quad 1\ ]^T$.

We can simplify Eq.(A.11) and Eq. (A.11) further to obtain:

$$h_{2n+1} = \frac{d}{2q}(h_1 - 2h_0 + h_{\text{init}})\left(\lambda_1^n + \lambda_2^n\right) + \frac{1}{2}(h_1 - h_{\text{init}})\left(\lambda_1^n - \lambda_2^n\right) + h_{\text{init}} \tag{A.12}$$

$$h_{2n+2} = \frac{d}{2q}(h_1 - 2h_1 + h_{\text{init}})\left(\lambda_1^n + \lambda_2^n\right) + \frac{1}{2}(h_2 - h_{\text{init}})\left(\lambda_1^n - \lambda_2^n\right) + h_{\text{init}}, \tag{A.13}$$

where $\lambda_1 = \frac{1}{2}d(d-2+q)$ and $\lambda_2 = \frac{1}{2}d(d-2-q)$, with $q = \sqrt{d^2 - 4d}$ are the eigenvalues of Δ in Eq. (A.6).

Therefore, under the assumptions we made in the beginning of this section, there are the following possibilities regarding the fate of node i:

1. If $d > 4$, $\lambda_1, \lambda_2 \in \mathbb{R}$ and $|\lambda_1| = \lambda_1 > 4 > 1$. Thus, the system is unstable and $h_{i,t}$ will tend to infinitely large negative values after node i becomes disconnected from the network. Consequently, the probability of a new link involving node i tends to 0 exponentially quickly. Moreover, $\lim_{d\to\infty} \lambda_1 = \infty$ and $\lim_{d\to\infty} \lambda_2 = 1^+$.

2. If $0 < d < 1$, $\lambda_1 = \lambda_2 = d$. Therefore $|\lambda_1| < 1$ and $|\lambda_2| < 1$. Thus, the system is stable and $h_{i,t}$ will eventually return to values close to $h_{i,0}$, which allow for the creation of links between node i and the rest of the network.

3. If $1 < d < 4$, the evolution of $h_{i,t}$ would be unstable and unbounded for $\tau > 1$ in the absence of the asymmetry in the panic factor defined in Eq. (A.4). However, this asymmetry condition gives rise to a situation in which $h_{i,t}$ eventually returns to a point close to $h_{i,0}$, where link formation is possible. This happens when $\delta h_{i,t}$ becomes non-negative, which implies $d \cdot \min(0, \delta h_{i,t}) = 0$.

The eigenvalues λ_1 and λ_2 corresponding to the cases above are represented in Fig. A.1.

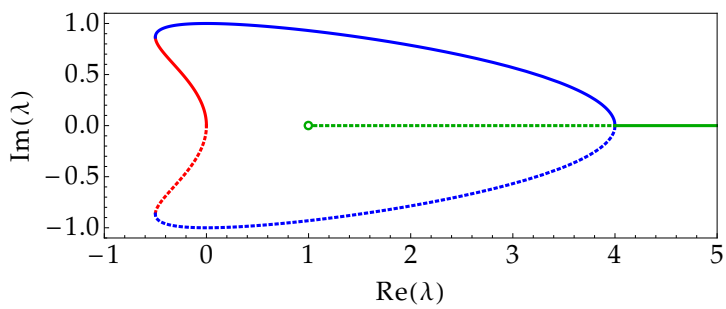


Figure A.1: Parametric representation of λ_1 and λ_2 in the complex plane. The unstable case ($d > 4$) is depicted in green, while the stable regime ($0 < d < 1$) is in red. The case with $1 < d < 4$, in which there is instability but $h_{i,t}$ is bounded, is in blue.

B

Detrimental collective effects in trading experiments

B.1 Written laboratory materials

B.1.1 Experimental instructions and quiz

In the following pages, the reader can find the exact instructions presented to the subjects in the beginning of the experiment.

Experimental Instructions

1. Overview

This is an experiment on economic decision making. If you follow the instructions carefully and make good decisions you may earn a considerable amount of money that will be paid to you in cash at the end of the experiment. The whole experiment is computerized, therefore you do not have to submit the paper on your desk. Instead, you can use it to make notes. There is a calculator on your desk. If necessary, you can use it during the experiment. Please do not talk with others for the duration of the experiment. If you have a question please raise your hand and one of the experimenters will answer your question in private.

Today you will participate in one or more market "sequences", each consisting of a number of trading periods. There are two objects of interest in this experiment, *shares* and *cash*, the latter denominated in *francs*. In each period you can trade shares in a market with a computer program, called *market maker*, and the currency used in the market is francs. You pay francs when you buy shares and receive francs when you sell shares. In each period you will have the opportunity to participate in trading or take no action in the market. Details about how this is done are discussed below in section 3. All trading will be in terms of francs. The cash payment to you at the end of the experiment will be in euros. The conversion rate is 4 francs to 1 euro.

2. Sequences of trading periods

As mentioned, today's experiment consists of one or more sequences, with each sequence consisting of an uncertain number of periods. Each period lasts 20 seconds. In each period you have to decide if you want to buy shares, sell shares, or hold either your francs or shares, i.e., take no action in the market. The amounts of your shares and francs will be shown on your computer screen. At the beginning of each period a computer program will spin a virtual roulette wheel visualized on your screen colored in blue for a proportion of 99 percent and colored in pink for the remaining 1 percent. If the black pointer ends up in the blue region, the sequence will continue with a new, 20-seconds period. If, instead, the black pointer ends up in the pink region, the sequence will end and your franc balance and the amount of shares for the sequence will be final (see Figure 1). Thus, at the start of each period, there is a 1 percent chance that the period will be the last one played in the sequence, and a 99 percent chance that the sequence will continue with at least one more period.

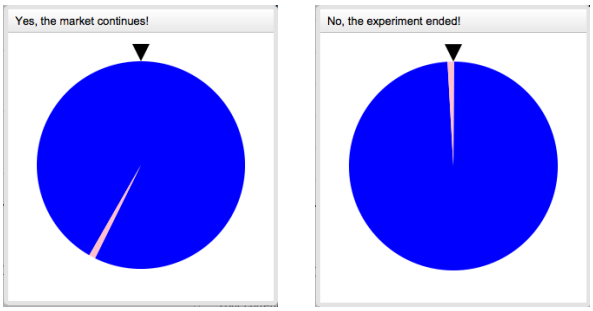


Figure 1. Left panel: experiment continues to next period. Right panel: experiment ends.

If less than 50 minutes have passed since the start of the first sequence, a new sequence will begin. You will start the new sequence just as you started the first sequence. If more than 50 minutes have passed since the beginning of the first sequence then the current sequence will be the last sequence played, meaning that the next time the roulette spin ends up with the black pointer in the pink region the sequence will end and the experiment will be over.

If, by chance, the final sequence has not ended by the three-hour period for which you have been recruited, we will gradually increase the pink region in the roulette wheel until the chance that the sequence will continue equals 0.

3. Market and trading rules

The market works as follows. At the beginning of the experiment, each participant will be given an initial endowment of 100 francs. In each period shares can be traded with the market maker. In particular, in each period each participant is allowed to hold either only shares or only francs. Therefore, during each period you may choose to

- Sell <u>all</u> your shares to the market maker in exchange for francs, if you are holding shares;
- Buy shares from the market maker by investing <u>all</u> your francs, if you are holding francs;
- Hold either your shares or francs, and take no action in the market.

Therefore, you cannot buy additional shares if you are already holding shares. Vice versa, you cannot sell shares if you are holding francs.

Trading within each period *t* occurs according to the following mechanism. First, at the beginning of each period *t*, the computer program spins the roulette wheel. Depending on the outcome of the roulette spin, we can distinguish two cases.

<u>CASE 1.</u> Roulette spin ends up in the blue region: market continues to next period

The price per unit of share in period *t*, denoted by P_t, is announced by the market maker and visualized on your computer screen. You can then decide whether to sell shares (if you are holding shares), buy shares (if you are holding francs), or hold either your shares or francs, and take no action in the market. In each period *t*, if you decide to sell or buy, your transaction will take place in the next period at price P_{t+1}.

Therefore, if you decide to **sell shares** in period *t*, you will receive an amount of francs in period *t+1* given by your amount of shares in period *t* times the price of shares in period *t+1*, which is:

Francs in period *t+1* = amount of shares in <u>period *t*</u> × price of shares in <u>period *t+1*</u>

If you decide to **buy shares** in period *t*, you will receive an amount of shares in period *t+1* given by your amount of francs in period *t* divided by the price of shares in period *t+1*, which is:

Shares in period *t+1* = amount of francs in <u>period *t*</u> / price of shares in <u>period *t+1*</u>

If you decide to **hold either your shares or your francs** in period *t*, the amount of shares or francs that you own will simply carry over to the next period *t+1*.

Notice that shares need not to be bought or sold in integer units. For example, suppose that in period *t* you own 127.65 francs and you decide to buy shares. Suppose then that the unit price of shares announced in period *t+1* turns out to be 172.50. This means that you will receive 127.65/172.50 = 0.74 shares in period *t+1*.

CASE 2. Roulette spin ends up in the pink region: experiment ends

The market maker announces the price per unit of share in period *t*, denoted by P_t, and all the buying and selling orders placed in the previous period *t-1* are executed at price P_t but it will not be possible to make any other decision to buy, sell or hold either shares or francs. Your final market earnings will then be computed as explained in section 4.

The timing of trading is summarized in Figure 2

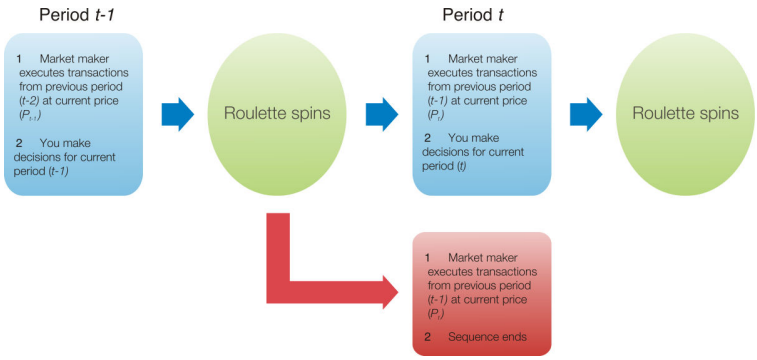


Figure 2. Timing of trading

How the price is determined

At the end of each period *t*, the market maker will collect the buy and sell orders and use them to determine the price P_{t+1} for the next period *t+1*. The percentage change between the price in period *t+1* and the price in period *t*, also called *return*, is approximately given by the sum of the following components

a) A constant positive term equal to 2%
b) A "trade impact factor" which depends upon the difference between the total amounts of buy and sell orders from the participants in the market. In particular:

- The higher the amount of buy orders in one period, the higher the price in the next period. Therefore, each buy order has a positive impact on price. In particular, if you and all other participants in the market decide to buy shares at the same time, the trade impact factor will be +100%.
- The higher the amount of sell orders in one period, the lower the price in the next period. Therefore, each sell order has a negative impact on price. In particular, if you and all other participants in the market decide to sell shares at the same time, the trade impact factor will be -100%.

Therefore the trade impact factor can take a maximum value of +100% and a minimum value of −100%. If all participants in the market decide to hold their shares or francs, the trade impact factor is zero.

c) Price shocks that can take positive and negative values with the same probability.

Given that the percentage change of prices from one period to the other is approximately given by the sum of the terms listed above (a + b + c), in case of no trading activity by any participant, price grows on average by approximately 2% every period. Figure 3 reports examples of typical price patterns in markets without any trading activity at any period, i.e. if all participants held their shares until the end of the experiment. All numbers in Figure 3 are provided only to give EXAMPLES, they SUGGEST NOTHING about the duration and price realizations of the experiment you are about to start.

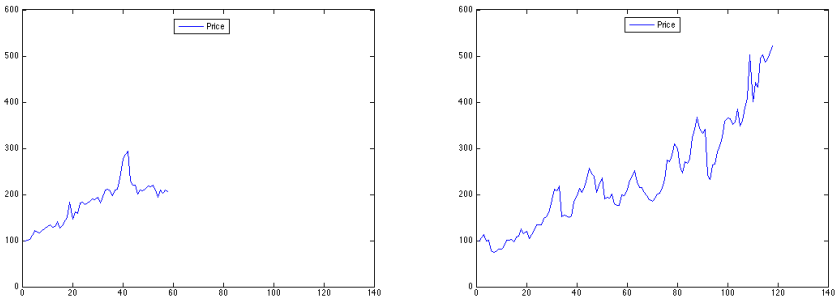


Figure 3. Examples of price patterns without any trading at any period

4. Your earnings

Net market earnings

When a sequence is terminated, that is whenever the roulette spin ends up in pink region, your end-of-sequence balance will be computed.

- If you are holding francs when the sequence is terminated, your amount of francs will determine your end-of-sequence balance.
- If you are holding shares when the sequence is terminated, the market value (in francs) of your shares, given by the amount of your shares times their price in the end of the sequence, will determine your end-of-sequence balance.

Your net market earnings will then be given by your end-of-sequence balance minus 100 francs, corresponding to the initial endowment that you received at the beginning of the experiment:

Net market earnings = end-of-sequence balance – initial endowment.

Therefore your earnings from participating in the market will be given by your end-of-sequence balance in francs minus your initial 100 francs endowment.

Forecast earnings

In addition to the money that you can earn from participating in the market, you can earn money by accurately forecasting, in each period *t*, the future price of shares in period *t+1*. You will earn a forecast prize of 0.10 Euro per period if your forecast of the shares' price is within the interval

[0.95 × realized price, 1.05 × realized price].

For example, if the realized price in period t is $p_t = 100$, you will earn the forecast prize if your forecast for p_t is within the interval [95, 105]. If, for example, the realized price in period t is $p_t = 200$ you will earn the forecast prize if your forecast for p_t is within the interval [190, 210].

Total earnings

Your total earnings for participating in today's experiment will equal the net market earnings that you have at the end of the sequence plus any money that you receive for the forecast task. If your net market earnings are negative, or smaller than the show up fee of 7 Euro, then your earnings from participating in the market will be zero and you will only receive the show up fee of 7 Euro

Earnings from trade = max(net market earnings, show up fee)

plus the money you earned for the forecasting task.

If you participate in more sequences, one of them will be randomly selected and your earnings will equal the total earnings in the selected sequence. As mentioned, the cash payment to you at the end of the experiments will be in Euro. The conversion rate is 4 francs to 1 Euro.

5. The computer screen

Below is a sample screen for a fictitious player 1 at the start of period 1.

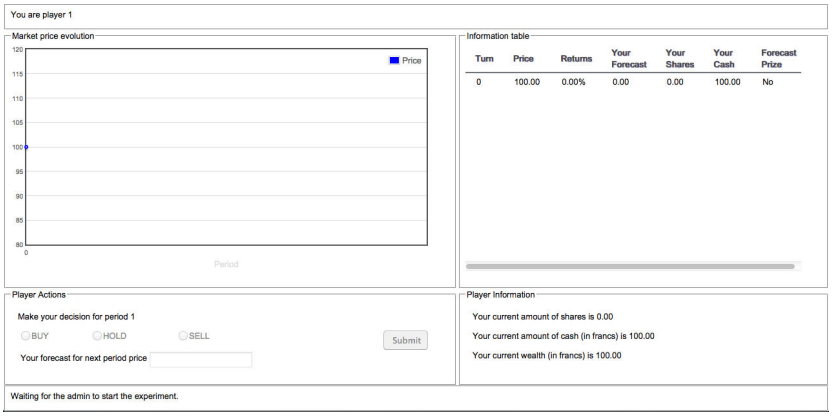


Figure 1. Example of screenshot

In every period, after the roulette spin, each player must perform two tasks:

- Decide whether to buy shares, hold either shares or francs, or sell shares by clicking on the corresponding radio button, i.e., BUY, HOLD, SELL. As explained in section 3, you can decide to buy shares only if in that period you are holding francs, and to sell shares only if in that period you are holding shares. Therefore, in every period, the only active buttons on your screen will be the buttons corresponding to your available actions;
- Enter a forecast of the shares' price in the next period. Please use the dot symbol to separate decimals (example: 10.32).

The box for "Player Actions" is located in the *bottom-left* corner of the screen. After making your choices, you have to submit your decisions by clicking on the "Submit" button.

The box in the *bottom-right* corner of the screen named "Player Information" reports the following information:

- The amount of shares you own in the current period
- The amount of francs you own in the current period
- Your wealth (in francs) in the current period, given by your francs (if you are holding francs) or your shares times the current price (if you are holding shares)

The rest of the screen allows you to track results from previous periods. The graph in the "Market price evolution" box on the *upper-left* corner of the screen reports a graphical representation of the shares' price over time. The table contained in the "Information table" box on the *upper-right* corner of the screen displays additional information about the results in the experiment and it is supplemental to the graph in the left part of the screen. The *first column* of the table shows the time period. The most recent period is always at the top. The *second* and the *third columns* show

respectively the price of shares and the returns, which represent the percentage change in shares price between the current period and the previous. A positive return, say in period 10, means that the price increased from period 9 to period 10, while a negative return in period 10 means that the price decreased from period 9 to period 10. The *fourth column* reports your forecasts (made in the previous period) of the price in the current period. For example, in period 6 the number at the top of the fourth column will report the forecast you entered in period 5 for the price of the shares at period 6. The *fifth* and *sixth columns* report respectively your amounts of shares and francs. Finally, the *seventh column* of the table shows whether in each period you earned the forecast prize or not.

The status bar at the bottom of the screen contains information about the status of the experiment and monitors how much time, out of the 20 seconds constituting the duration of each period, you have to take your decisions. If time is up before you make your choices, the computer program will select HOLD as default action, i.e., you will hold either your shares or francs, and use your previous period forecast.

6. Final Quiz

It is important that you understand these instructions. Before continuing with the experiment, we ask that you consider the following scenarios and provide answers to the questions asked in the spaces provided. The numbers used in the quiz are merely illustrative; the actual numbers in the experiment may be quite different. You may find it useful to consult the instructions to answer some of these questions.

Question 1: Suppose that a sequence has reached period 25. What is the chance that this sequence will continue with another period, namely period 26? _____. Would your answer be any different if we replaced 25 with 7 and 26 with 8?

- Yes
- No

Question 2: Suppose that the sequence has reached period 12, the price announced for that period is $P_{12} = 15$ and you own 5.3 shares. If you decide to sell your shares, how many shares are you allowed to sell?_____. Suppose that you indeed decide to sell in period 12 and that the price announced in period 13 is $P_{13} = 10$. How many francs will you receive?_____.

Question 3: Suppose that the sequence has reached period 5, the price announced for that period is $P_5 = 8$ and you own 10 francs. If you decide to buy your shares, how many francs are you allowed to invest?_____. Suppose that you indeed decide to buy in period 5 and that the price announced in period 6 is $P_6 = 20$. How many shares will you receive?_____.

Question 4: Suppose that at the beginning of period t a price $P_t = 200$ is announced and then the roulette spin ends up in the pink region. Suppose that you have 1.5 shares. What will be your net market earnings (in francs) at the end of the sequence?_____.

Question 5: Suppose that when the experiment ends, i.e., the roulette spin ends up in the pink region, you have 100 francs. What will be your net market earnings (in francs) at the end of the sequence?_____.

Question 6: Suppose that in period t you and all other participants decide to hold either your francs or shares and take no action in the market. Do you expect next period's price P_{t+1} to

- Increase
- Decrease

[Turn sheet for final question]

Question 7: Consider the following scenarios

a) In period *t* you and all other participants have shares. <u>All of you</u> decide to sell.
b) In period *t* you and all other participants have shares. <u>Only you</u> decide to sell.

Regarding scenario a), do you expect next period's price P_{t+1} to

– Increase
– Decrease

How big will the "trade impact factor" (see section "How the price is determined") be in scenario a)?_____.

Regarding scenario b), do you expect the " trade impact factor" to be:

- Bigger than in scenario a)
- Smaller than in scenario a)

Moreover, do you expect:

- Bigger P_{t+1} in scenario a) than in scenario b)
- Smaller P_{t+1} in scenario a) than in scenario b)

B.1.2 Lottery

You will now face an additional task that will give you the chance to earn extra money, which will be added to what you already earned in today's experiment.

You will face two sequences (sequence 1 and sequence 2) of 10 decisions each. Details about the decisions that we ask you to make are described in the following sheets.

After you make decisions for both sequence 1 and sequence 2, you will randomly select one of the two sequences by picking a ball from a jar containing balls numbered 1 and 2, and your choice will determine which sequence will be used to determine your payoff. Obviously Sequence 1 and sequence 2 have the same chance of being chosen.

When you have completed all your decisions, and you are satisfied with those decisions please raise your hand and you will be called for payment.

You may now read the instructions on the following sheets.

SEQUENCE 1

You will face a sequence of 10 decisions. Each decision is a paired choice between two options labeled "Option A" and "Option B". For each decision you must choose either Option A or Option B. After making your choice, please record it on the attached record sheet under the appropriate headings.

The sequence of 10 decisions you will face are as follows:

Decision	Option A	Option B
1	Receive €4.00 1 in 10 chances OR Receive €3.20 9 in 10 chances	Receive €7.70 1 in 10 chances OR Receive €0.20 9 in 10 chances
2	Receive €4.00 2 in 10 chances OR Receive €3.20 8 in 10 chances	Receive €7.70 2 in 10 chances OR Receive €0.20 8 in 10 chances
3	Receive €4.00 3 in 10 chances OR Receive €3.20 7 in 10 chances	Receive €7.70 3 in 10 chances OR Receive €0.20 7 in 10 chances
4	Receive €4.00 4 in 10 chances OR Receive €3.20 6 in 10 chances	Receive €7.70 4 in 10 chances OR Receive €0.20 6 in 10 chances
5	Receive €4.00 5 in 10 chances OR Receive €3.20 5 in 10 chances	Receive €7.70 5 in 10 chances OR Receive €0.20 5 in 10 chances
6	Receive €4.00 6 in 10 chances OR Receive €3.20 4 in 10 chances	Receive €7.70 6 in 10 chances OR Receive €0.20 4 in 10 chances
7	Receive €4.00 7 in 10 chances OR Receive €3.20 3 in 10 chances	Receive €7.70 7 in 10 chances OR Receive €0.20 3 in 10 chances
8	Receive €4.00 8 in 10 chances OR Receive €3.20 2 in 10 chances	Receive €7.70 8 in 10 chances OR Receive €0.20 2 in 10 chances
9	Receive €4.00 9 in 10 chances OR Receive €3.20 1 in 10 chances	Receive €7.70 9 in 10 chances OR Receive €0.20 1 in 10 chances
10	Receive €4.00 10 in 10 chances OR Receive €3.20 0 in 10 chances	Receive €7.70 10 in 10 chances OR Receive €0.20 0 in 10 chances

After you have made all 10 decisions, you will be called in a separate room for payment and we will throw a ten-sided die (the faces are numbered from 1 to 10, and the "0" face of the die will serve as 10) twice, once to select one of the ten decisions to be used, and a second time to determine what your payoff is for the option you chose, A or B, for the particular decision selected. Even though you will make ten decisions, only ONE of these will end up affecting your earnings, but you will not know in advance which decision will be used. Obviously, each decision has an equal chance of being used to determine your earnings.

Consider Decision 1. If you choose Option A, then you receive €4.00 if the throw of the ten-sided die is 1, while you receive €3.20 if the throw is 2-10. If you choose Option B, then you receive €7.70 if the throw of the ten-sided die is 1, while you receive €0.20 if the throw is 2-10. The other decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between €4.00 or €7.70.

Please circle your choice for each of the 10 decisions on your record sheet. Notice that you may choose Option A for some decisions and Option B for others.

SEQUENCE 2

You will face a sequence of 10 decisions. Each decision is a paired choice between two options labeled "Option A" and "Option B". For each decision you must choose either Option A or Option B. After making your choice, please record it on the attached record sheet under the appropriate headings.

The sequence of 10 decisions you will face are as follows:

Decision	Option A	Option B
1	Receive €20.00 1 in 10 chances OR Receive €16.00 9 in 10 chances	Receive €38.50 1 in 10 chances OR Receive €1.00 9 in 10 chances
2	Receive €20.00 2 in 10 chances OR Receive €16.00 8 in 10 chances	Receive €38.50 2 in 10 chances OR Receive €1.00 8 in 10 chances
3	Receive €20.00 3 in 10 chances OR Receive €16.00 7 in 10 chances	Receive €38.50 3 in 10 chances OR Receive €1.00 7 in 10 chances
4	Receive €20.00 4 in 10 chances OR Receive €16.00 6 in 10 chances	Receive €38.50 4 in 10 chances OR Receive €1.00 6 in 10 chances
5	Receive €20.00 5 in 10 chances OR Receive €16.00 5 in 10 chances	Receive €38.50 5 in 10 chances OR Receive €1.00 5 in 10 chances
6	Receive €20.00 6 in 10 chances OR Receive €16.00 4 in 10 chances	Receive €38.50 6 in 10 chances OR Receive €1.00 4 in 10 chances
7	Receive €20.00 7 in 10 chances OR Receive €16.00 3 in 10 chances	Receive €38.50 7 in 10 chances OR Receive €1.00 3 in 10 chances
8	Receive €20.00 8 in 10 chances OR Receive €16.00 2 in 10 chances	Receive €38.50 8 in 10 chances OR Receive €1.00 2 in 10 chances
9	Receive €20.00 9 in 10 chances OR Receive €16.00 1 in 10 chances	Receive €38.50 9 in 10 chances OR Receive €1.00 1 in 10 chances
10	Receive €20.00 10 in 10 chances OR Receive €16.00 0 in 10 chances	Receive €38.50 10 in 10 chances OR Receive €1.00 0 in 10 chances

After you have made all 10 decisions, you will be called in a separate room for payment and we will throw a ten-sided die (the faces are numbered from 1 to 10, and the "0" face of the die will serve as 10) twice, once to select one of the ten decisions to be used, and a second time to determine what your payoff is for the option you chose, A or B, for the particular decision selected. Even though you will make ten decisions, only ONE of these will end up affecting your earnings, but you will not know in advance which decision will be used. Obviously, each decision has an equal chance of being used to determine your earnings.

Consider Decision 1. If you choose Option A, then you receive €20.00 if the throw of the ten-sided die is 1, while you receive €16.00 if the throw is 2-10. If you choose Option B, then you receive €38.50 if the throw of the ten-sided die is 1, while you receive €1.00 if the throw is 2-10. The other decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between €20.00 or €38.50.

Please circle your choice for each of the 10 decisions on your record sheet. Notice that you may choose Option A for some decisions and Option B for others.

RECORD SHEET FOR SEQUENCE 1

Decision 1	Circle Option Choice A B
Decision 2	Circle Option Choice A B
Decision 3	Circle Option Choice A B
Decision 4	Circle Option Choice A B
Decision 5	Circle Option Choice A B
Decision 6	Circle Option Choice A B
Decision 7	Circle Option Choice A B
Decision 8	Circle Option Choice A B
Decision 9	Circle Option Choice A B
Decision 10	Circle Option Choice A B

LAB COMPUTER ID:

RECORD SHEET FOR SEQUENCE 2

Decision 1	Circle Option Choice A B
Decision 2	Circle Option Choice A B
Decision 3	Circle Option Choice A B
Decision 4	Circle Option Choice A B
Decision 5	Circle Option Choice A B
Decision 6	Circle Option Choice A B
Decision 7	Circle Option Choice A B
Decision 8	Circle Option Choice A B
Decision 9	Circle Option Choice A B
Decision 10	Circle Option Choice A B

LAB COMPUTER ID:

References

[1] Albert, R. & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1), 47.

[2] Amini, H., Cont, R., & Minca, A. (2013). Resilience to contagion in financial networks. *Mathematical Finance*.

[3] Amman, H. M., Tesfatsion, L., Judd, K. L., Kendrick, D. A., & Rust, J. (2006). *Handbook of computational economics*, volume 2. Elsevier.

[4] Anand, K., Gai, P., & Marsili, M. (2010). The rise and fall of trust networks. *Progress in Artificial Economics*, 645, 77–88.

[5] Anand, K., Kirman, A., & Marsili, M. (2013). Epidemics of rules, rational negligence and market crashes. *The European Journal of Finance*, 19(5), 438–447.

[6] Anderson, S. P., De Palma, A., & Thisse, J. F. (1992). *Discrete choice theory of product differentiation*. MIT press.

[7] Anufriev, M. & Hommes, C. (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4), 35–64.

[8] Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *The American economic review*, (pp. 406–411).

[9] Assenza, T., Delli Gatti, D., & Grazzini, J. (2015). Emergent dynamics of a macroeconomic agent based model with capital and credit. *Journal of Economic Dynamics and Control*, 50, 5–28.

[10] Assenza, T., Heemeijer, P., Hommes, C., & Massaro, D. (2013). *Individual expectations and aggregate macro behavior*. Technical report, Tinbergen Institute Discussion Paper.

[11] Axelrod, R. (1997). The dissemination of culture a model with local convergence and global polarization. *Journal of conflict resolution*, 41(2), 203–226.

[12] Bao, T., Hommes, C., Sonnemans, J., & Tuinstra, J. (2012). Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control*, 36(8), 1101–1120.

[13] Barber, B. M. & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1), 261–292.

[14] Barberis, N. & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1, 1053–1128.

[15] Battiston, S., Delli Gatti, D., Gallegati, M., Greenwald, B., & Stiglitz, J. E. (2012a). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control*, 36(8), 1121–1141.

[16] Battiston, S., Puliga, M., Kaushik, R., Tasca, P., & Caldarelli, G. (2012b). Debtrank: too central to fail? financial networks, the fed and systemic risk. *Scientific reports*, 2, 541.

[17] Bianconi, G. & Marsili, M. (2004). Clogging and self-organized criticality in complex networks. *Physical Review E*, 70(3), 35105.

[18] Blanchard, O. (2014). Where danger lurks. *Finance & Development*, 51(3), 28–31.

[19] Bouchaud, J.-P. (2013). Crises and collective socio-economic phenomena: Simple models and challenges. *Journal of Statistical Physics*, 151(3-4), 567–606.

[20] Bouchaud, J.-P., Farmer, J. D., & Lillo, F. (2009). How markets slowly digest changes in supply and demand. (pp. 57 – 160).

[21] Bouchaud, J.-P., Matacz, A., & Potters, M. (2001). Leverage effect in financial markets: The retarded volatility model. *Physical Review Letters*, 87(22), 228701.

[22] Bouchaud, J. P. & Potters, M. (2003). *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge University Press.

[23] Caballero, R. J. & Simsek, A. (2009). Complexity and financial panics. *National Bureau of Economic Research*, (14997).

[24] Caccioli, F., Shrestha, M., Moore, C., & Farmer, J. D. (2014). Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance*, 46, 233–245.

[25] Camerer, C. F. (2004). Prospect theory in the wild: Evidence from the field. *Colin F. Camerer, George Loewenstein, and Matthew. Rabin, eds., Advances in Behavioral Economics*, (pp. 148–161).

[26] Challet, D. & de Lachapelle, D. M. (2013). A robust measure of investor contrarian behaviour. In *Econophysics of Systemic Risk and Network Dynamics* (pp. 105–118). Springer.

[27] Challet, D. & Zhang, Y.-C. (1998). On the minority game: Analytical and numerical studies. *Physica A: Statistical Mechanics and its Applications*, 256(3), 514–532.

[28] Charney, J. G., Fjörtoft, R., & Neumann, J. v. (1950). Numerical integration of the barotropic vorticity equation. *Tellus*, 2(4), 237–254.

[29] Colander, D., Goldberg, M., Haas, A., Juselius, K., Kirman, A., Lux, T., & Sloth, B. (2009). The financial crisis and the systemic failure of the economics profession. *Critical Review*, 21(2-3), 249–267.

[30] Coleman, J. (1958). Relational analysis: the study of social organizations with survey methods. *Human Organization*.

[31] Cont, R. & Wagalath, L. (2014). Fire sales forensics: measuring endogenous risk. *Mathematical Finance*.

[32] Contreras, M. G. A. & Fagiolo, G. (2014). Propagation of economic shocks in input-output networks: A cross-country analysis.

[33] Corsi, F., Marmi, S., & Lillo, F. (2013). When micro prudence increases macro risk: The destabilizing effects of financial innovation, leverage, and diversification. *SSRN Electronic Journal*.

[34] Cox, J. D. (2002). *Storm watchers: the turbulent history of weather prediction from Franklin's kite to El Niño*. John Wiley & Sons.

[35] da Gama Batista, J., Bouchaud, J.-P., & Challet, D. (2015a). Sudden trust collapse in networked societies. *The European Physical Journal B*, 88(3), 1–11.

[36] da Gama Batista, J., Massaro, D., Bouchaud, J.-P., Challet, D., & Hommes, C. (2015b). Do investors trade too much? a laboratory experiment. *arXiv preprint*.

[37] Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance*, 53(6), 1839–1885.

[38] Daniel, K., Hirshleifer, D., & Teoh, S. H. (2002). Investor psychology in capital markets: Evidence and policy implications. *Journal of monetary economics*, 49(1), 139–209.

[39] Dawid, H., Gemkow, S., Harting, P., Van der Hoog, S., & Neugart, M. (2014). Agent-based macroeconomic modeling and policy analysis: the Eurace@Unibi model.

[40] de Lachapelle, D. M. & Challet, D. (2010). Turnover, account value and diversification of real traders: evidence of collective portfolio optimizing behavior. *New Journal of Physics*, 12(7), 075039.

[41] Delli Gatti, D., Desiderio, S., Gaffeo, E., Cirillo, P., & Gallegati, M. (2011). *Macroeconomics from the Bottom-up*, volume 1. Springer Science & Business Media.

[42] Donier, J. & Bouchaud, J.-P. (2015). Why do markets crash? bitcoin data offers unprecedented insights.

[43] Dosi, G., Fagiolo, G., Napoletano, M., & Roventini, A. (2013). Income distribution, credit and fiscal policies in an agent-based keynesian model. *Journal of Economic Dynamics and Control*, 37(8), 1598–1625.

[44] Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6), 2059–2106.

[45] Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American statistical Association*, 82(397), 171–185.

[46] Ehrhardt, G., Marsili, M., & Vega-Redondo, F. (2006). Phenomenological models of socioeconomic network dynamics. *Physical Review E*, 74(3), 036106.

[47] Ehrhardt, G., Marsili, M., & Vega-Redondo, F. (2009). Homophily, conformity, and noise in the (co-) evolution of complex social networks. *Complexity and Spatial Networks*.

[48] Farmer, J. D. & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256), 685–686.

[49] Farmer, J. D., Gallegati, M., Hommes, C., Kirman, A., Ormerod, P., Cincotti, S., Sanchez, A., & Helbing, D. (2012). A complex systems approach to constructing better models for managing financial markets and the economy. *European Physical Journal-Special Topics*, 214(1), 295.

[50] Farmer, J. D. & Geanakoplos, J. (2009). The virtues and vices of equilibrium and the future of financial economics. *Complexity*, 14(3), 11–38.

[51] Gabbi, G., Iori, G., Jafarey, S., & Porter, J. (2015). Financial regulations and bank credit to the real economy. *Journal of Economic Dynamics and Control*, 50, 117–143.

[52] Gai, P., Haldane, A., & Kapadia, S. (2011). Complexity, concentration and contagion. *Journal of Monetary Economics*, 58(5), 453–470.

[53] Gai, P. & Kapadia, S. (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2120), 2401–2423.

[54] Gatti, D. D., Gallegati, M., Lillo, F., Iori, G., Hommes, C., Thurner, S., & Bouchaud, J.-P. (2011). CRISIS – complexity research initiative for systemic instabilities. http://www.crisis-economics.eu. Accessed: 2015-04-01.

[55] Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, 115(3), 811–846.

[56] Goetzmann, W. N. & Peles, N. (1997). Cognitive dissonance and mutual fund investors. *Journal of financial Research*, 20(2), 145–158.

[57] Gopikrishnan, P., Meyer, M., Amaral, L. N., & Stanley, H. E. (1998). Inverse cubic law for the distribution of stock price variations. *The European Physical Journal B-Condensed Matter and Complex Systems*, 3(2), 139–140.

[58] Graham, S., Parkinson, C., & Chahine, M. (2002). Weather forecasting through the ages. *NASA Earth Observatory*.

[59] Greenwood, R. & Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, 27(3), 714–746.

[60] Grinblatt, M. & Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of finland's unique data set. *Journal of Financial Economics*, 55(1), 43–67.

[61] Gualdi, S., Tarzia, M., Zamponi, F., & Bouchaud, J.-P. (2014). Tipping points in macroeconomic agent-based models. *Journal of Economic Dynamics and Control*, (pp.42).

[62] Hamman, J. R., Weber, R. A., & Woon, J. (2011). An experimental investigation of electoral delegation and the provision of public goods. *American Journal of Political Science*, 55(4), 738–752.

[63] Harmon, D., Stacey, B., Bar-Yam, Y., & Bar-Yam, Y. (2010). Networks of economic market interdependence and systemic risk. *arXiv preprint*, (pp.9).

[64] Harrison, G. W., List, J. A., & Towe, C. (2007). Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion. *Econometrica*, 75(2), 433–458.

[65] Harrison, G. W. & Rutström, E. E. (2008). Risk aversion in the laboratory. *Research in experimental economics*, 12, 41–196.

[66] Heise, S. & Kühn, R. (2012). Derivatives and credit contagion in interconnected networks. *The European Physical Journal B*, 85(4), 115.

[67] Holt, C. A. & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5), 1644–1655.

[68] Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control*, 35(1), 1–24.

[69] Hommes, C. (2013). *Behavioral rationality and heterogeneous expectations in complex economic systems*. Cambridge University Press.

[70] Hommes, C., Sonnemans, J., Tuinstra, J., & Van de Velden, H. (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies*, 18(3), 955–980.

[71] Iori, G. & Porter, J. (2012). Agent-based modelling for financial markets. *Handbook on Computational Economics and Finance*. Forthcoming.

[72] Jackson, A. (2003). The aggregate behaviour of individual investors. *Available at SSRN 536942*.

[73] Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational behavior and human decision processes*, 51(2), 296–312.

[74] Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The journal of economic perspectives*, (pp. 193–206).

[75] Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. *The Journal of Finance*, 63(1), 273–310.

[76] Keynes, J. M. (2006). *General Theory Of Employment , Interest And Money*. Atlantic Publishers & Dist.

[77] Kirman, A. (1989). The intrinsic limits of modern economic theory: the emperor has no clothes. *The Economic Journal*, (pp. 126–139).

[78] Kirman, A. (2010). *Complex economics: individual and collective rationality*. Routledge.

[79] Kirman, A. & Zimmermann, J.-B. (2001). *Economics with heterogeneous interacting agents*, volume 503. Springer Science & Business Media.

[80] Kirman, A. P. (1992). Whom or what does the representative individual represent? *The Journal of Economic Perspectives*, (pp. 117–136).

[81] LeBaron, B. (2006). Agent-based computational finance. *Handbook of computational economics*, 2, 1187–1233.

[82] Lillo, F. & Farmer, J. D. (2004). The long memory of the efficient market. *Studies in nonlinear dynamics & econometrics*, 8(3).

[83] Lillo, F., Farmer, J. D., & Mantegna, R. N. (2003). Econophysics: Master curve for price-impact function. *Nature*, 421(6919), 129–130.

[84] Lillo, F. & Pirino, D. (2015). The impact of systemic and illiquidity risk on financing with risky collateral. *Journal of Economic Dynamics and Control*, 50, 180–202.

[85] Loomes, G. & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, (pp. 805–824).

[86] Lorenz, J., Battiston, S., & Schweitzer, F. (2009). Systemic risk in a unifying framework for cascading processes on networks. *The European Physical Journal B*, 71(4), 441–460.

[87] Marsden, P. V. (1988). Homogeneity in confiding relations. *Social Networks*, 10(1), 57–76.

[88] Marsili, M., Vega-Redondo, F., & Slanina, F. (2004). The rise and fall of a networked society: a formal model. *Proceedings of the National Academy of Sciences of the United States of America*, 101(6), 1439–1442.

[89] McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444.

[90] Michard, Q. & Bouchaud, J.-P. (2005). Theory of collective opinion shifts: from smooth trends to abrupt swings. *The European Physical Journal B-Condensed Matter and Complex Systems*, 47(1), 151–159.

[91] Myerson, R. B. (1999). Nash equilibrium and the history of economic theory. *Journal of Economic Literature*, (pp. 1067–1082).

[92] Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *The computer journal*, 7(4), 308–313.

[93] Obama, B. (2009). President Barack Obama's inaugural address.

[94] Odean, T. (1999). Do investors trade too much? *The American Economic Review*, 89(5), 1279–1298.

[95] Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2014). Epidemic processes in complex networks.

[96] Peixoto, T. P. & Bornholdt, S. (2012). No need for conspiracy: Self-organized cartel formation in a modified trust game. *Physical Review Letters*, 108(21), 218702.

[97] Pin, P., Franz, S., & Marsili, M. (2008). Opportunity and choice in social networks. *Fondazione Eni Enrico Mattei Working Papers*.

[98] Putnam, R. (1995). *The case of missing social capital*. Technical report, Harvard University working paper.

[99] Reigneron, P.-A., Allez, R., & Bouchaud, J.-P. (2011). Principal regression analysis and the index leverage effect. *Physica A: Statistical Mechanics and its Applications*, 390(17), 3026–3035.

[100] Schelling, T. C. (1978). *Micromotives and macrobehavior*. WW Norton & Co Ltd.

[101] Schularick, M. & Taylor, A. M. (2009). *Credit booms gone bust: monetary policy, leverage cycles and financial crises, 1870–2008*. Technical report, National Bureau of Economic Research.

[102] Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy & Public Affairs*, (pp. 317–344).

[103] Sethna, J. P., Dahmen, K. A., & Myers, C. R. (2001). Crackling noise. *Nature*, 410(6825), 242–50.

[104] Sieczka, P., Sornette, D., & Holyst, J. A. (2011). The lehman brothers effect and bankruptcy cascades. *The European Physical Journal B*, 82(3-4), 257–269.

[105] Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, (pp. 99–118).

[106] Smith, V. L., Suchanek, G. L., & Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society*, (pp. 1119–1151).

[107] Squartini, T., van Lelyveld, I., Garlaschelli, D., & van Lelyveld, I. (2013). Early-warning signals of topological collapse in interbank networks. *Scientific reports*.

[108] Trichet, J.-C. (2010). Reflections on the nature of monetary policy non-standard measures and finance theory. In *ECB Central Banking Conference*, volume 18: ECB.

[109] Tumminello, M., Lillo, F., Piilo, J., & Mantegna, R. N. (2012). Identification of clusters of investors from their real trading activity in a financial market. *New Journal of Physics*, 14(1), 013041.

[110] Tumminello, M., Miccichè, S., Lillo, F., Piilo, J., & Mantegna, R. N. (2011). Statistically validated networks in bipartite complex systems. *PloS one*, 6(3), e17994.