UC Santa Cruz
UC Santa Cruz Electronic Theses and Dissertations

Title
On Cooperation, Coordination, and Virtual Economies

Permalink
https://escholarship.org/uc/item/0sb0j8zr

Author
Baumer, Matthew Allen

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

ON COOPERATION, COORDINATION, AND VIRTUAL ECONOMIES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Matthew Allen Baumer

June 2016

The Dissertation of Matthew Allen
Baumer
is approved:

_____

Professor Daniel Friedman, Chair

_____

Professor Nirvikar Singh

_____

Professor Johanna Francis

_____

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

On Cooperation, Coordination, and Virtual Economies

by

Matthew Allen Baumer

Virtual economies are growing as internet technology continues to advance. In **Aggregate Dynamics in a Large Virtual Economy: Prices and Real Activity in Team Fortress 2**[1], we analyze a large and complete set of transaction data from the Team Fortress 2 (TF2) virtual economy, which was designed to allow for decentralized barter as the sole exchange institution. A small subset of goods emerges endogenously to act as media of exchange. Taking one of these money goods as numeraire, we generate daily prices for thousands of goods. We then generate macroeconomic indicators, including nominal growth and inflation. We find evidence of a particular sort of nominal rigidity related to the circulation of multiple types of currency goods, and also find some localized asset price bubbles associated with announcements by the game designers.

Continuing work with this complete set of transactions in **Emergence of Networks and Market Institutions in a Large Virtual Economy**[2], we construct trader and goods networks, and track them over time using metrics such as node strength, assortativity, betweenness and closeness. The trading platform was designed to make barter exchange as attractive as possible; money was not part of the design and all players were

---

[1]Coauthored with Curtis Kephart
[2]Coauthored with Curtis Kephard and Daniel Friedman

ix

created equal. Yet, within weeks, several specific goods emerged as media of exchange, and various specialized traders appeared that facilitated exchange. Eventually trade was predominantly money-mediated and market-makers played a major role. Our results illustrate how network analysis can capture the spontaneous emergence of economic institutions by applying common conceptions of centrality to goods and traders.

The closing work is **Minimum effort coordination in continuous time - An experimental analysis with changing payoff structures**[3]. The minimum effort game is among the most studied coordination games because there exists no social dilemma yet coordination failure has been commonly observed. We extend the minimum effort game to continuous to test whether that is sufficient to induce high levels of coordination. We find that continuous time is not always enough to induce efficient coordination in spite of decreased signaling costs and extensive information about the decisions of other players. We finally extend our model to examine the effects of "gradualism" on coordination and find that they are still less efficient compared to treatments with a mild penalty parameter.

---

[3]Coauthored with Thomas Campbell and Maren Tonn

To my parents,

without whose support I would never have

been able to accomplish this

Acknowledgments

Part I

Part 1: Virtual Economies

Chapter 1

Aggregate Dynamics in a Large Virtual Economy: Prices and Real Activity in Team Fortress 2

Written with coauthor Curtis Kephart.

1.1 Introduction

The object of our study is the virtual economy of Team Fortress 2 (TF2) developed and overseen by Valve Software. This economy and others like it hold great potential for researchers: millions of users engaging in billions of economic transactions involving thousands of different types of goods; the game designers are near-omnipotent social planners able to create and destroy goods and implement policy at will; and they gather essentially complete micro data that enables precise construction of macro

variables. Our goal in this work is to develop a methodology of taking data in the form of a bilateral barter exchange history that is tractable and can be used to define familiar notions from macroeconomics. We develop price time series for individual goods, and using those we defined a price index, a measure of aggregate wealth, and measures of inflation. In this way, we hope to describe an environment that is perhaps alien to many economists in such a way that the unfamiliar is familiar and that will inspire the study of virtual economies by future researchers.

The TF2 economy has some features that are unusual, even for a virtual economy. There is no explicit currency good, and trading occurs exclusively through decentralized barter. Goods are homogeneous and of known quality (i.e. there are no "lemons" as in Akerlof (1970)). Items are also durable and do not depreciate due to "wear and tear" in the way that a physical item would. Another issue likely important is that items are indivisible and can only be exchanged in discrete quantities (e.g. it is impossible to trade half of a common currency item, the treasure key, as keys are not capable of being split). There is also a significant amount of activity that is due to a very small number of very active individuals, which we will refer to as "high net worth individuals" (HNWIs). These quirks will be leveraged in future papers to discuss the issue of the spontaneous emergence of money, the emergence of trade intermediaries, and information brokerage services by applying concepts from network theory to map the interactions between different types of user.

Our approach advances ideas presented in Castronova et al. (2009) and Castronova (2008) by implementing more rigorous economic indicators of aggregate eco-

nomic behavior in a large virtual economy. But there are also some crucial differences in our work: Castronova studies Everquest II, a economy with explicit currency (gold pieces) and in-game posted-price markets available to the users, thus trade in that environment would not be considered barter or decentralized in any sense. Our work also more directly adheres to methodology commonly used in modern empirical economic techniques.

Everquest II and TF2 are far from the only such examples of large virtual worlds with economic activity: "Second Life" is an entire virtual world, complete with in-game real estate, stores, jobs, and of course other people. "World of Warcraft" has players fight monsters and each other with the hope of saving the realm from the great evil that threatens it and has players engaging in money-mediated trade with each other to facilitate this end.

Even the NYSE has made its operations completely digital. Traders physically standing on the trading floor on Wall Street are in fact conducting all of their business through computer servers located in Mahwah, New Jersey. Stock traders are now similarly employed in the business of exchanging zeroes and ones in a computer database, albeit with higher stakes and a much greater degree of sophistication. The NYSE and its affiliated traders have had almost 200 years to develop their institutions; what commerce in virtual economies will look like once it matures is an open question.

1.2 Research Questions

Q1: What is the trend in real growth per-capita and how can we explain this trend?

Our primary goal is a basic macroeconomic characterization of this large virtual economy. We will examine the dynamics of real growth and explain what are the economic causes of dynamics by performing a decomposition of nominal growth into its constituent components: growth of the price level, real growth, and population growth.

Q2: How does aggregate price level respond to macro-level shocks to a component of the money supply?

An appealing consequence of the complete nature of our dataset is the ability for us to pinpoint precisely what might be causing, for example, bouts of inflation or deflation. An example of an exogenous shocks that we examine is a holiday promotional events that led to an increase in the rate of creation of "scrap metals", which effectively is a shock to the rate of growth of one component of money supply. But there are multiple monetary goods, so we seek to answer what happens when only one of the moneys is exposed to a shock.

Q3: How do markets for individual items respond to micro-level shocks?

A quirk of this environment is that there are numerous unexpected events that can be taken as exogenous by market participants. For example, a number of cosmetic items were suddenly "retired", meaning they were removed from the store and new items of these types could no longer be acquired, fixing their number in the economy.

We might expect this intervention to increase prices – essentially a negative shock to supply – but it is also possible that market participants' speculation "overshoots" the new (post-announcement) fundamental value.

1.3 Environment and Data

Team Fortress 2 is a competitive multiplayer first-person-shooter game which has two teams of typically 6 to 10 combatants vying for supremacy. Winning could result from (depending on the game mode) killing enough of your opponents (but don't worry, death is only temporary!), capturing a briefcase full of valuable intelligence from the heart of your opponent's base and sneaking it back to your home base, or successfully pushing a cart full of explosives to your opponents base to blow them up. One round of the game typically takes ten to twenty minutes. Each player in a game chooses their character class from nine different options such as quick and agile Scout, the pro-social Medic, or the deceptive Spy and try to do their best to help their team achieve glorious victory.

1.3.1 Economic Environment

TF2 debuted in 2007 and initially followed the standard video game business model: players pay for a copy and can play to their hearts' content. Then, in 2008, a new dimension was added to the game: an item system which allowed users to collect virtual goods which would customize the look and play style of their characters. As people played, they would randomly receive item drops (and some special items could result
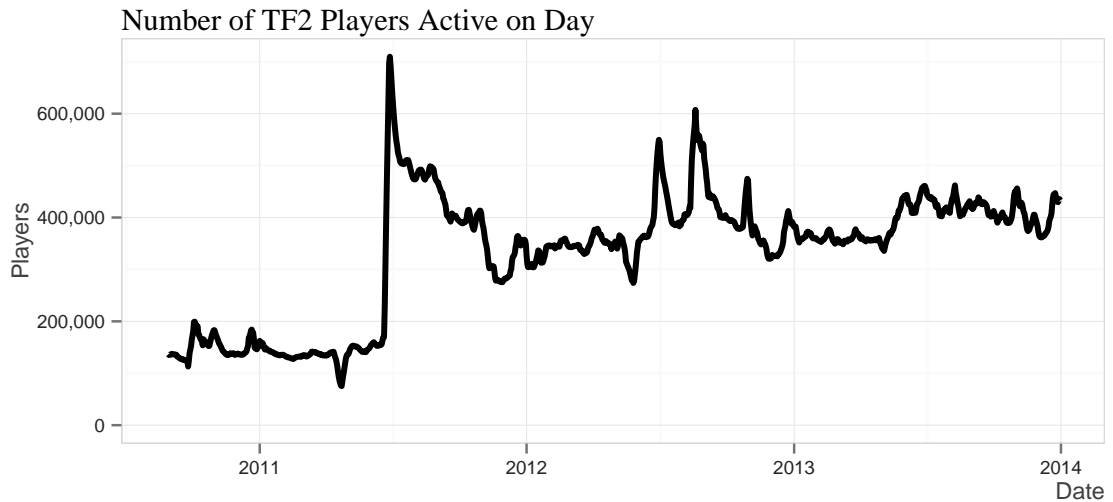
Figure 1.1: Daily TF2 Players

Note: Count of unique players logged into TF2 over day. Seven day moving average applied.

from completing a list of in-game achievements), but there was no way to exchange items with each other. A September 2010 update introduced two institutions which continue on until today: a barter platform to exchange items with other players, and a virtual store where items could be directly purchased from Valve using a credit card. Figure 1.1 displays the number of active players on a daily basis. In November 2013 alone, there were more than 2.1 million different users that spent some amount of time playing TF2. At the end of July 2011, TF2 went "free-to-play" (F2P), removing the requirement to purchase a game license before people are allowed to play, at which point the game generated revenue only by selling in-game items on the official store.

An item in the context of the Valve marketplace is any virtual good that can be stored in a player's inventory (henceforth referred to as a "backpack") and be traded. These may include TF2 items, installation licenses for other games on Valve's digital

7

distribution platform called Steam, and items from games other than TF2 on the Steam platform. Backpacks have finite space, but the capacity is large enough (300 item slots) that most users are unlikely feel this constraint. As well, there are "backpack expanders" that can be purchased from Valve for $.99 which loosen this constraint by granting an additional 100 item slots.

The process of successfully completing a trade is as follows: Find a trading partner through communication channels that can be internal or external to TF2, add them to your contact list, request a trade session, arrange an exchange in that session which makes both parties happy, and then execute the trade after multiple layers of confirmation. This is a quite inconvenient system for the market participants, but it represents an opportunity for inquiring economists to study actual human behavior in an environment in which we are theoretically well versed. It is important to point out that the economy by construction was designed to support only barter.

Our sample consists of a full log of all transactions occurring between 9 August 2011 and 31 May 2013, a 661 day interval. There were more than 70 million barter transactions, which averages out to more than 100,000 trades per day or over one trade per second. This is the primary source of the data set which we will use to do the following analysis. Across these 70 million individual transactions, over 300 million virtual items changed hands. There were 4,267,832 unique traders participating in the barter market, with the median trader conducting 4 exchanges, and with approximately one third of traders exchanging ten or more items over the sample period. Some traders participated in a large number of trades; the top ten accounts by trade count each
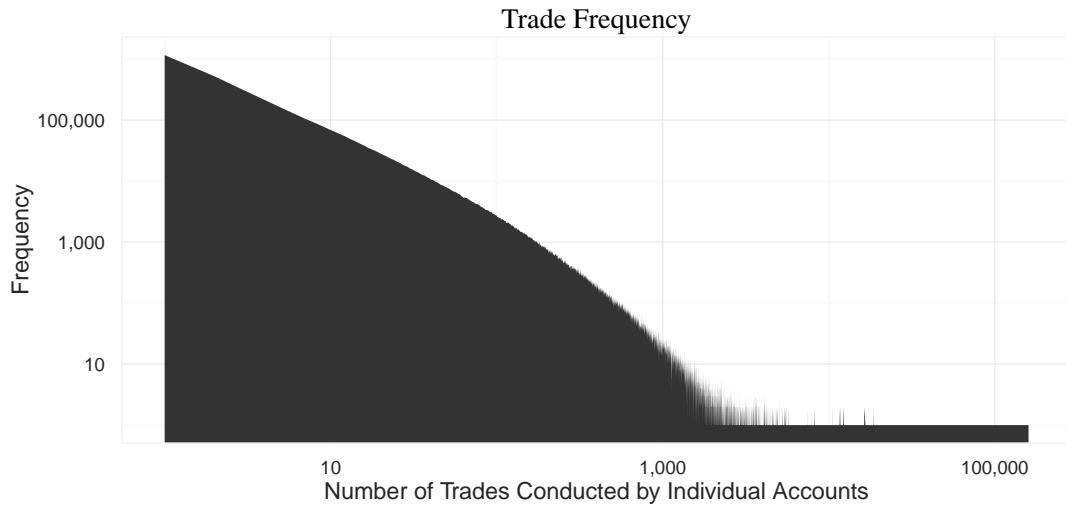
Figure 1.2: Histogram of User Trade Frequency

conducted over 150,000 barter transactions.[1]

The Team Fortress 2 trading environment represents the largest dataset of a barter exchange market that we are aware of. This is all the more remarkable since barter markets today tend to emerge in environments which feature weak institutions and consequently have meager record-keeping.

Items in TF2 have various types. There are consumables that are used in conjunction with other items (e.g. a can of paint that can be used on a cosmetic that changes the item's color palette, or a name tag that lets the player choose a custom name for their item) and durables which can be used for as long as the owner wishes and do not undergo any sort of depreciation as a result of use. All durables have associated

---

[1]User Privacy: In order to protect the privacy of individuals involved in the TF2 Economy, user identities were were anonymized, timestamps masked, and any data containing unique user identifiers was held on Valve Corporation machines. Though the researchers were given access to the full log of market transactions, all other company supplied metrics removed users who marked their Steam backpacks to private.

class restrictions; some durables can be equipped by any class and others can only be equipped by one or a few classes.

In addition, each individual item is designated one of a number of different "qualities", which serve primarily to signal scarcity and characteristics of the item. These include:

- Unique: counterintuitively, the most common item quality

- Unusual: adds a custom effect to the item such as flames erupting from the item's surface and are overall the rarest and most sought after quality

- Strange: track various statistics for the player when worn

There are a few other qualities of items, but they are generally simple variations of unique.

Players can gain items from a number of different sources: random drops from playing (although there is a cap of how many items can be received per time period from this source), direct purchase from the "Valve store" using real cash, special promotions (e.g. holidays, as a reward for completing some achievement, or as an incentive for buying another game), trading with other players, by opening crates which require a key which is then consumed along with the crate, and through a crafting system introduced in December of 2009.

From observation of the set of items most commonly used as a unit of account on independent community-created trading posts, there is evidence that the widely accepted commodity currencies include three denominations of "metals", as well as "keys", "Bill's

Hats", and "Earbuds". The three different types of metals in order of increasing value are scrap, reclaimed, and refined metal. There exists an in-game system that allows conversion of one denomination into another in either direction at the rate of 3 lower valued to 1 of the next higher valued. For example, anyone can convert 3 scrap metals into 1 reclaimed, then combine that reclaimed with 2 more reclaimed to create a refined, then break that refined back into 3 reclaimed. There is no cost associated with these conversions beyond the time it takes to perform the necessary mouse clicks.

Metals result from scrapping (deleting) weapons from your backpack and are used in combination with other metals and items to create new items via predefined recipes. Keys only originate from store purchases and may be used to open crates that contain new items with various probabilities. Crates are analogous to raffle tickets; if you pay the cost of one key to open a crate, you will most likely get an item worth somewhat less than the key but there is a small chance to get a very valuable item worth much more.

Metals and keys are created and consumed regularly. Bill's Hats and Earbuds, in contrast, entered the market as promotional items given away in the past and can no longer be found or purchased directly from Valve. Their supply is bounded by the current number in existence and slowly shrinks due to people quitting the game or deleting them.

Once a player is in possession of an item, they will not lose it unless they either trade, delete, or consume the item in the case of consumables. At the end of 2012, the ability for players to sell items directly to other players for Valve store credit in an

official centralized posted-price marketplace was added. This store credit is denominated in the player's local currency and is redeemable for TF2 items purchased from the Valve store as well as the purchase of licenses for other games from Valve's digital distribution platform called "Steam".

This demonstrates an important distinction between this economy and the physical world; in order to produce a good there are raw materials that necessarily must be consumed due to conservation of mass. But the production of an additional good in this virtual economy requires no more than a additional line saved to a database. There is still technically an upper bound on how many items can exist, but for practical purposes this horizon is infinite and the marginal cost of production of these goods is zero for Valve.

Another distinction between this environment and physical economies comes from the nature of consumption. Most real goods are actually consumed at some rate and once they are used up, are no longer usable again. This does not happen in TF2. Most consumption is of goods which are perfectly durable (with the exception of tools, but tools either result in or modify durables). We can then think of the size of this economy as being the aggregate value of the stock of durables and tools.

1.3.2 Data

Much of our data takes the form of logs documenting barter transactions of virtual items between two users. These are lists of transactions linked to users and the individual items associated with the trade. These data were supplied to us via

a half terabyte sized relational database from which we generated observations in the form of Table 1.1. Each row in the transactions log represents the movement of a single item and is associated with a unique trade identifier, two unique player identifiers (one for the sender of the item and one for the recipient), a unique item-level identifier which no two items share (AssetID), and an identifier for the specific item *type* which identical items would share with each other (EconAssetClass). For example, if a player possesses two unique quality "Bill's Hats" that are otherwise identical, they would share an EconAssetClass but also each will be associated with unique AssetID that represents the specific individual item. When an item is traded its old AssetID is removed from the originating user's inventory and a new one is created for the user receiving it. Thus, we can track both individual items as well as individual classes of items, defined as items which share a type and quality which makes them functionally identical.

Table 1.1: Example data snippet

TradeID	PartyA	PartyB	Time	AppID	AssetID	NewAssetID	Origin	EconAssetClass
1	1203	1876	1351926000	440	38818	41361	1	100
2	4256	172	1351927010	440	39425	41362	0	194921
2	4256	172	1351927010	440	41359	41363	1	158535
3	993	8384	1351928320	440	41339	41364	0	207

We present as an example Table 1.1. By looking at trade IDs, we can classify each individual trade into categories such as simple monetary trades or simple barter trades, as will be discussed in detail later. Party A and B allow us to track the trading behavior of individual traders and the AssetID and NewAssetID let us track the movements of individual items as they pass from user to user. Origin indicates which user

is the recipient of the item transfer and EconAssetClass is the identifier which lets us determine the specific item type that was traded. In this fabricated example, the first trade was a one-way exchange where a player with ID number 1876 gave an item to another player with ID number 1203 and received nothing in return. The item that was given away was of type 100. The next trade involved the player 4256 giving an item of type 194921 to player 172 and receiving an item of type 158535 in exchange.

1.4 Estimating Prices from Barter Data

Our approach to generating prices for individual items is to define one good among the emergent currencies to be our numeraire, calculate spot exchange rates between the other currencies and our numeraire, and convert goods exchanged for those alternative currencies into the corresponding value in terms of the numeraire. This approach gives us price estimates which allow for direct value comparisons between all items. We also generate statistics for each item including daily turnover, number of trades, and stocks.

The question of how to define which goods are used as "currencies" and which are not is not a trivial one, but this discussion is not something we shall delve into in this paper.[2] Since the different metals can be converted costlessly into each other in either direction at the rates mentioned previously, we convert all price observations involving

---

[2]In the following chapter, we will rigorously identify goods that appear to be the most "money-like" based on their characteristics in the data, but for now we will simply take money goods for granted and assign currency status to those items which are used as a unit of account in the major community-run pricing resources.

	Basket on Side 1	**Basket on Side 2**

FX - Money for Money — All Money Goods of any type or mix ↔ All Money Goods of any type or mix

SM - Simple Monetary

    SM.Keys - Simple Monetary with Keys — All Key(s) ↔ Non Money Good(s) of one type

    SM.Mix - Simple Monetary with Mix — All Money Goods of any type or mix ↔ Non Money Good(s) of one type

SB - Simple Barter — Non Money Good(s) of one type ↔ Non Money Good(s) of one type

OW - One-Way — Any Item(s) ↔ Empty

EE - Everything Else

    EE.SM - Simple Monetary in EE — Money Goods of any type or mix | Non Money Good(s) of one type ↔ All Money Goods of any type or mix

EE - All other trades

Figure 1.3: List of Trade Classification Types.

metal into the equivalent value in terms of refined metal.

From all of the goods used as commodity currencies, we choose keys to be the numeraire. Keys were selected because they appear to have the most stable value, likely due to the fact that their supply is allowed to expand as well as contract and the price is anchored to the dollar since keys can only be produced in the economy through direct purchases from the Valve store at a price of $2.49 per key. The other potential currency goods either were introduced later on (Bill's Hats and Earbuds) or displayed rapid expansion of supply (faster than the growth of population) causing instability in estimated prices.

We define a **simple monetary** (SM) trade observation as an exchange involving a single non-currency item type and any basket of commodity currency items. In order to use SM trades to estimate prices that are comparable to each other, prices need to be measured in a common unit, which we refer to as "synthetic keys". A synthetic key price is the equivalent key-value of a good perhaps exchanged for non-key money(s). We calculate daily exchange rates between different types of money items by looking at the subset of trades that are **money for money** (FX), which are defined as trades which have only money goods on both sides. See Figure 1.3 for a complete classification of all possible trade types.

By looking at these FX trades, we generate daily inter-money exchange rates as follows. Define Q_{it}^{KM} as the quantity of *keys* traded for *metals* on date t in transaction i, where transaction i is among the subset of trades involving only metal on one side and only keys on the other. Q_{it}^{MK} is likewise the quantity of *metals* (expressed in terms of refined) traded for *keys* in the same exchange. A single metal/key exchange rate observation is thus,

$$R_{i,t}^{KM} = \frac{Q_{i,t}^{KM}}{Q_{i,t}^{MK}}$$

The daily spot metal/key exchange rate is then the median of all i exchange rate observations on date t (weighted by the number of keys in each observation), allowing us to value any quantity of metal in terms of the going rate for keys at that moment. By a similar process we derive daily synthetic key values of Bill's Hats and Earbuds.

16

Over 910,000 transactions inform our FX sample (approximately 1,300 a day), of which 700,468 are metal-for-keys exchanges, 107,651 are Bill's Hat for some combination of metals and keys, and 104,566 involve Earbuds for some metal-key combination. Spot Bill's Hat and Earbuds exchange rates are based on trades involving keys and/or metals, converting metals into synthetic keys at the day's metal/key exchange rate.

Using these spot exchange rates to express all SM trades in terms of keys, a simple monetary price observation is as follows:

$$P_{it}^{SM} = \frac{V_{2,it}^{S.Key} - V_{1,it}^{S.Key}}{Q_{1,it}^{SM}}$$

Where $V_{2,it}^{S.Key}$ is the value, in terms of synthetic keys, of the all-money side of a SM trade, $V_{1,it}^{S.Key}$ is the synthetic key value of any currency goods on the side of the trade that involves a non-money item, and $Q_{1,it}^{SM}$ is the quantity of the non-money good involved in the SM trade. $V_{1,it}^{S.Key}$ can be thought of as a cashier making change when a larger than necessary denomination of currency is used to make a purchase.

Over 9 million trades provide SM price observations, or an average of approximately 14,000 per day. We aggregate our sample of asynchronous price observations on time period and EconAssetClass (item type) to generate price time series for each individual variety of item. Figure 1.4 demonstrates our price time series for an arbitrarily chosen item, namely a stylish sombrero hat called "Old Guadalajara" which is wearable only by the Pyro character class. Notice that there are discrete bands above and below the price trend line; this is a consequence of the indivisibility of the currency goods.

17
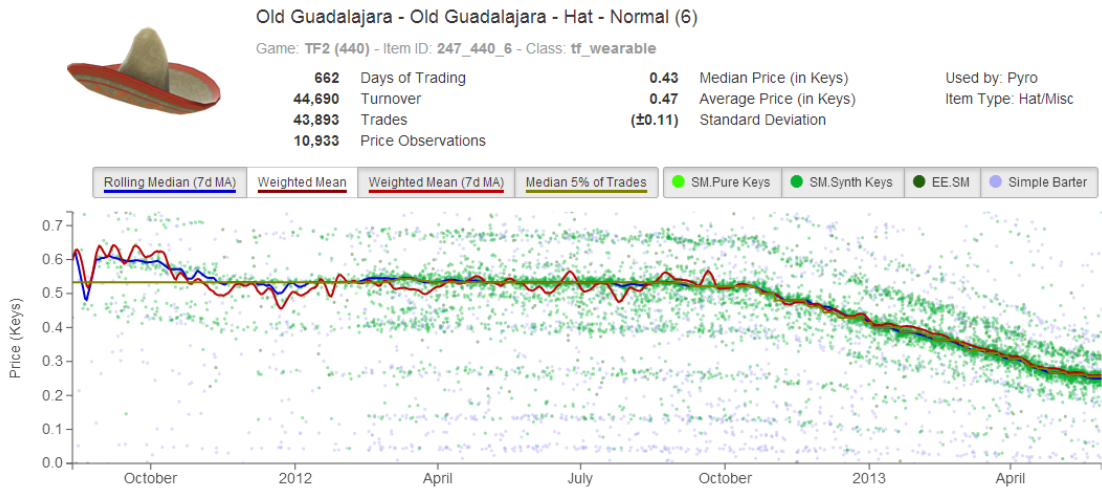
Figure 1.4: Price time series and meta data.

Note: A typical individual item price time series. Scatter points reflect individual transactions and their implied valuation. Multicolored lines reflect various temporal aggregate methods deriving daily prices.

Prior to October 2012, the first of these bands are .1-.15 keys away from each other, which would correspond to the value of one reclaimed metal at contemporary market exchange rates.

An additional 8.5 million trades offer Simple Barter (SB) item value observations as well – trades that involve only two non-money items. However, we only use SM price observations and did not incorporate SB prices because they appear to have a more complicated valuation method than SM trades. It appears that when traders meet, if the buyer of the specific item does not or can not pay in currency items, they must pay a premium with their non-money items, meaning the trade won't be balanced in terms of value. This would simply introduce mean-zero noise to valuations if we assume that all items are equally sought after by barter traders. But if some items were relatively more sought after than other for barter exchange, there would be some item-specific

18

fixed term that would need to be controlled for. We therefore choose to exclude SB observations from our price estimations as we determined that the number of SM trades is sufficiently large that our estimation process will be precise.

Our temporal aggregation approach assumes that each item at every moment possesses an underlying "fundamental market valuation" based on its characteristics and relevant market conditions. We then take each individual price observation as a noisy signal for that item's contemporary fundamental value. That is, we assume SM price observations are drawn from their true values, plus some error process. It is worth mentioning that some items appear to have reasonably complex profiles, such as bimodality in price, which we take as further evidence of the economic significance of currency indivisibility.

To estimate the price of a given item on a given day, we start with a seven day window centered on that day and collect all observed SM transactions involving that item. We then clean out observations beyond the 1st and 9th price deciles as there are outliers which, for thinly traded items, can lead to a large amount of volatility. To estimate prices using a rolling average, we then apply a weighting function to these price observations based on temporal distance from the day in question and widen the time window beyond one day if necessary.[3]

A distinguishing characteristic of this environment is the constant addition of new types of items that players can buy or find. This methodology involves taking observed transactions around a given day and using those to estimate spot prices. This

---

[3]See Appendix 4.1.1 for more details regarding determination of appropriately wide time windows.

approach is not ideal for pricing items soon after their introduction because there will be relatively few observations. To mitigate this issue, we also develop a hedonic pricing model that imputes prices of items based on observable characteristics and supplement price estimates directly as above with estimates from this hedonic model for use in our price index. This hedonic model will be discussed further in the next section.

1.5   Methods

1.5.1   Market Capitalization

We now turn to characterizing the size and growth rate of the TF2 virtual economy. Due to the relative lack of production, GDP is not an appropriate measure for this. We instead calculate the "market capitalization" which we are defining as the total key-value of aggregate item stocks held by active players, where a player is designated "active" if they have played within 90 days. To calculate this, we take the level of existing stocks of each item in each time period and multiply them by the prevailing price in that time period, then sum over all items. We will denote aggregate nominal wealth in period t as W_t and is defined as

$$W_t = \sum_{i=1}^{N_t} p_{i,t} S_{i,t}$$

where at time t there are N_t total different goods, $p_{i,t}$ is the price of good i and $S_{i,t}$ is its outstanding stock. One quirk of this economy is that a large majority of existing

goods do not undergo any sort of depreciation. This means that value is constantly being created but relatively rarely being destroyed; compare this to, for example, the value created by a pig farmer. He creates value by raising pigs and selling the pork, but this value ceases to exist once the pork is eaten. Keys and metals are consumed in a similar fashion to this pork, but it is rare for other economically significant items to disappear. But this begs the question: even though item stocks should be increasing over time, is the real wealth of the average individual agent increasing along with it?

The nominal growth of all active players' inventory holdings can be written

$$W_{t+1} = G_t W_t \tag{1.1}$$

where W_t represents the nominal wealth and G_t represents the growth rate of nominal wealth in period t. G_t is the product of three components, population growth G_t^P, per-capita real growth G_t^R, and growth of prices (i.e. inflation) G_t^I. Thus, we can take logs of equation 1.1 to find (where lower cases denote log levels):

$$\Delta w_t = w_{t+1} - w_t = g_t^P + g_t^R + g_t^I \tag{1.2}$$

To better understand the causes of shifts in nominal aggregate wealth, we will take advantage of this decomposition but before we can do this we will need measures for each of these components.

1.5.2 Törnqvist Price Index

The simplest starting point for a basic price index is a Laspeyres index which uses a quantity basket fixed to a base year and estimates price increases by allowing prices to adjust in each time period:

$$P_t^{Laspeyres} = \frac{\sum_i^N p_{it}q_{i0}}{\sum_i^N p_{i0}q_{i0}}$$

However, there is a particular problem with direct implementation of a basic Laspeyres index: New items are constantly being introduced. If we choose a base period early in our timeline, we will leave out all of the items which were introduced later on which are likely to be economically important. But if we choose a base period late in our timeline, since there are some items which did not exist early in the sample, we can have no prices for items in early periods. And, indeed, this is a significant issue for our environment. At the beginning of our data set, there are about 630 different item types traded, and at the end there are over 1600. The common alternative to a basic Laspeyres index is a Paasche index. Paasche indices suffer from a closely related issue; they take the quantity index from the current year in the denominator rather than quantities from the base year. But we can have no prices in the base time period for items which were introduced later on since we have no observed trades of goods that did not exist. Our strategy for solving this problem is twofold. First, we use a modified Törnqvist index rather than Laspeyres or Paasche. Second we use a hedonic model to estimate what prices for goods would have been just before their introduction.

Our modified Törnqvist index (Törnqvist, 1936) modeled after the way the US C-CPI-U handles its upper level price indices.[4] The Törnqvist index is superlative and built from Translog preference functions[5]. A Törnqvist price relative is as follows:

$$P_{t,t-1}^{\mathbb{T}} = \frac{P_t}{P_{t-1}} = \prod_{i=1}^{n} \left(\frac{p_{i,t}}{p_{i,t-1}} \right)^{\frac{1}{2}\left(\frac{p_{i,t-1}q_{i,t-1}}{V_{t-1}} + \frac{p_{i,t}q_{i,t}}{V_t} \right)}$$

where V_t is the total nominal value of all goods in the quantity basket in period t, thus $\frac{p_{i,t}q_{i,t}}{V_t}$ is the expenditure share of good i in period t. The quantity index we use to calculate was built by drawing a weekly sample of active players from the population and observing what those players were holding in their backpack. For a detailed description of our sampling methodology, please see Appendix 4.1.3.

The Törnqvist index helps to avoid the problem discussed above with the simple Laspeyres: since the base period for each calculation is the previous period, the number of new items introduced between base and current periods are minimized. As well, since the weights are value shares, new items being introduced simply decreases the weights of already existing items so the index does not increase due to increasing quantities of items. The chain Törnqvist price index from base period $t = 0$ to period T is thus:

$$\text{Chain } P_T^{\mathbb{T}} = \prod_{t=1}^{T} \left(P_{t,t-1}^{\mathbb{T}} \right)$$

One issue with our approach is due to the existence of items which are un-

---

[4]For more details, see Cage et al. (2003) and Bureau of Labor Statistics (2014) and ILO-IWGPS (2004).

[5]See chapter 18 of the **Export and Import Price Index Manual** (2009) released by the IMF for a detailed discussion of superlative indices
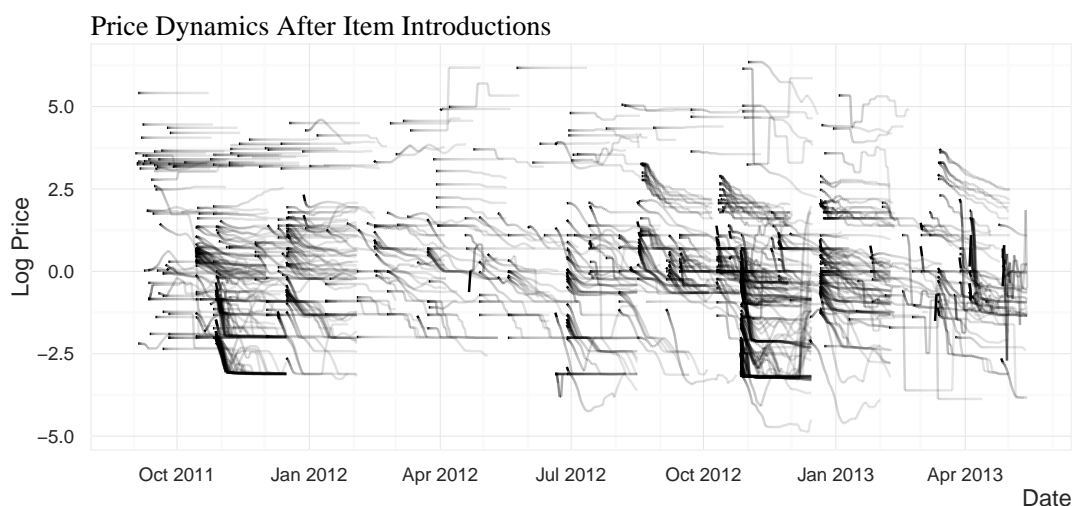
Figure 1.5: New item price time series.

Note: Each sparkline represents an item's price over its first fifty days.

tradable - that is we observe no prices – but which appear in our representative bundle. These items certainly have a non-zero value and they do enter and leave people's inventories, but we have no choice to exclude these from our index. This is the same way that national statistical offices handle non-priced services like family household services.

1.5.3   Hedonic Pricing Model

Another potential issue is that newly introduced items generally exhibit a commonality in price trajectories. Most new items start at a premium relative to similar items, and then steadily trade lower in price. Figure 1.5 displays the price dynamics of items starting with their introduction and tracing the time path of their log prices for the first fifty days thereafter. Log prices are used to shrink the visual distance between item time series, hopefully helping to focus on general price dynamics. Note that there

24

are clusters of new items around Halloween and the December holidays. Items with high starting prices (log price greater than 2.5, about 12 keys or more) appear to hold their value in most cases, but items with lower initial values nearly always trend downward.

The Törnqvist price relative discussed above ignores items for which price information is not present in adjacent periods, and thus the initial premium price on most new items is not captured by the existing methodology. Though this issue is likely mitigated by the fact that new items are infrequently traded and seen in relatively few inventories when first introduced–and so their weights would be quite low–the omission of item introductions likely biases our price index downwards.

We deal with the problem of new item introductions by implementing a hedonic pricing model (Diewert, 2003; Rosen, 1974) which estimates the prices of items based on that item's characteristics compared to the characteristics of other items with known prices. A similar hedonic price imputation approach is used by national statistical bureaus to estimate prices in conditions of changes in quality. We use the hedonic method as a best estimate of the initial values of each item based on the item's observable characteristics. This is accomplished by regressing dummies for each of these characteristics interacted with time dummies on each item's prices over time. For a given time period, this gives an estimated value for each characteristic an item can have. If we apply the assumption that an item's value is approximated by the sum of values of its parts, we can estimate the price of an arbitrary item given only its vector of characteristics. We then use these imputed prices as our best estimates for the value of items the day before they are introduced.

We impute unobserved prices via the following hedonic price model:

$$\ln(p_{it}) = \alpha + \delta_t D_t + \sum_{k=1}^{K}(\beta_{kt} \cdot x_{ik}) + \varepsilon_{it} \quad \text{for } t = 0, 1, ..., T \quad (1.3)$$

For the price p_{it} of item i in period t. D_t are fixed-effect time dummies (by week), x_{ik} is a dummy indicating whether or not item i possesses item characteristic k (such characteristics are time invariant), with error epsilon which has the standard assumption of being equal to zero in expectation. Thus δ_t is the parameter for the fixed effects of week t and β_{kt} is the parameter on characteristic k in week t.

The different characteristics x_{ik} we include in this model are item quality, class equipability as some items can be used only by certain classes and others can be used by any class, item equip slot such as weapon or hat, and finally a dummy indicating items held by a large proportion of active players which took a value of 1 if 3% or more of users held the item and applied to less than 25% of items. We believe that these characteristics sufficiently describe different items. We are limited by the fact that a certain degree of the differentiation between items is due to non-quantifiable aesthetics (e.g., two items can be identical with respect to the observables mentioned above, but one of them might have art design that is in some sense "more attractive" and thus would command a premium), but we believe that the number of different items is large enough that these will be sufficiently averaged out when we conduct our regression.
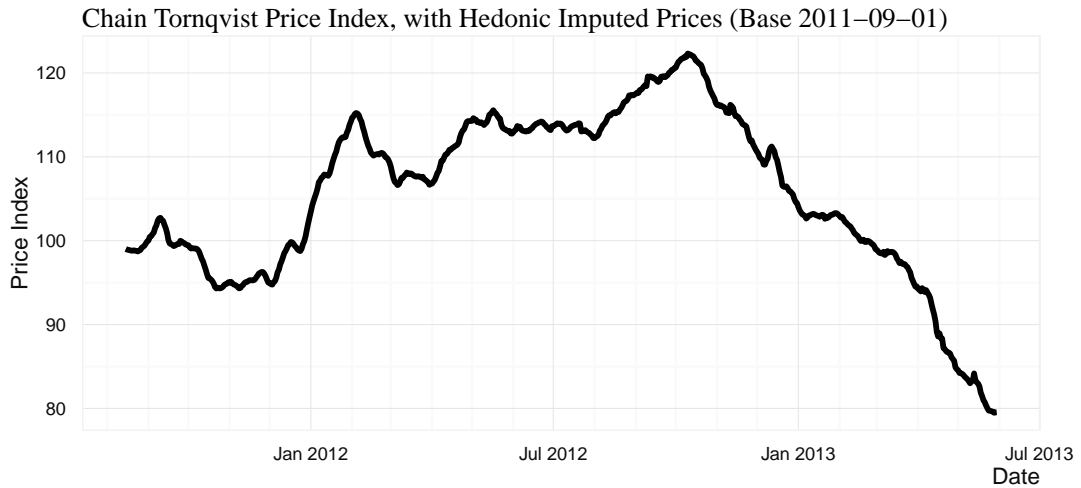
Figure 1.6: Aggregate Price Index using Chain Törnqvist with Hedonic Prices

1.6 Results

Our primary goal is the characterization of macroeconomic growth of this virtual barter environment. This requires the development of an aggregate price index and hedonic pricing models. Next, we present possible explanations for some of the observed macro-level behavior. We conclude with our analysis of the impact of micro-level shocks on individual items with evidence of an asset price bubble, the first bubble to be documented in a barter market as far as we are aware.

1.6.1 Aggregate Price Level

In Figure 1.6, we present the calculated chain Törnqvist price index. Overall, the price level based on representative backpack contents is relatively stable with slight deflation until approximately mid-December of 2011, when there is a surge of inflation that is possibly related to a Christmas event which brought an influx of new users into

27

the game and introduced holiday-themed items from new crates. This is followed by a dip towards the end of the first quarter of 2012 which proved to be temporary as prices return to their initial level and remain there for several months before seeing steady inflation until October 2012, where we see the most striking feature of our price index. Starting with the Halloween event of 2012, we see a sustained deflationary period. Our index returns to its initial level around March 2012 and keeps falling until the end of our sample.

1.6.2 Hedonic Model

The hedonic hypothesis postulates that for any given period, a good is a bundling of potentially time-varying price determining characteristics along with some possible aggregate price level effects that change from period to period.[6] Plotted in Figure 1.7 are the coefficients on the weekly fixed effect dummies D_t along with their first and second standard errors bands. These can be interpreted as an estimate for changes in the overall price level in a given week relative to the first week. Compare Figure 1.6 to Figure 1.7; with the exception of a peak in the first quarter of 2012 which does not appear in Figure 1.7, the dynamics are remarkably similar. These are both estimating the same thing using entirely different methodology but both tell generally the same story.

Figure 1.8 plots how item characteristics have evolved over the sample using the

---

[6]Since item-level characteristics are fairly well defined in this context – item quality, character class equipablility, and broad item type – it may be informative to run a simplified hedonic regression which eliminates time-variation in the β coefficients. Results from such a model could be interpreted as the average value placed on each observable characteristic for items in our sample and are presented in Appendix 4.1.5.
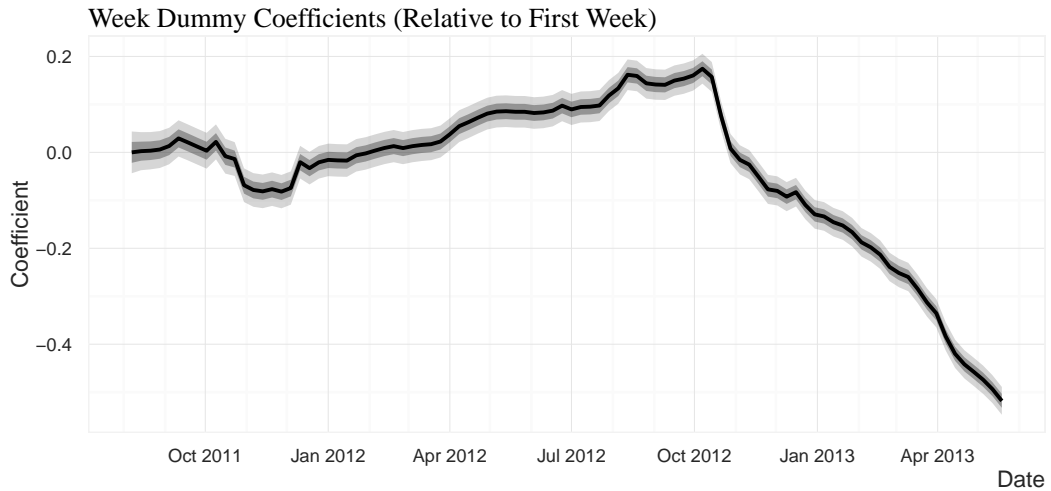
Figure 1.7: Time dummy estimate from simple hedonic price model.

Note: Dark and light gray ribbons reflect first and second standard error bands respectively. Coefficient estimates on time dummies from model Equation 4.1.

hedonic model from equation 1.3. In Figure 1.8 we see the evolution of value premiums based on item quality. For example, haunted items tend to have their highest premiums around Halloween (technically, we observe haunted items' least discounts around Halloween – haunted items are essentially identical to unique items, except for their text color and quality designation), but haunted quality items otherwise tend to trade at a discount relative to unique items. Unusuals clearly trade at a consistent and increasing premium relative to uniques and other qualities. Interestingly, in the weeks preceding Halloweens, unusuals exhibit an increase in their value premium. This is possibly due to the introduction of a number of highly coveted Halloween themed visual effects (e.g. circling ghosts, cauldron bubbles, and ''Demonflame'') at this time. Vintage items exhibit a consistently increasing premium relative to uniques. Vintages are defined by having been in existence prior to the introduction of item trading. These likely show steady

29
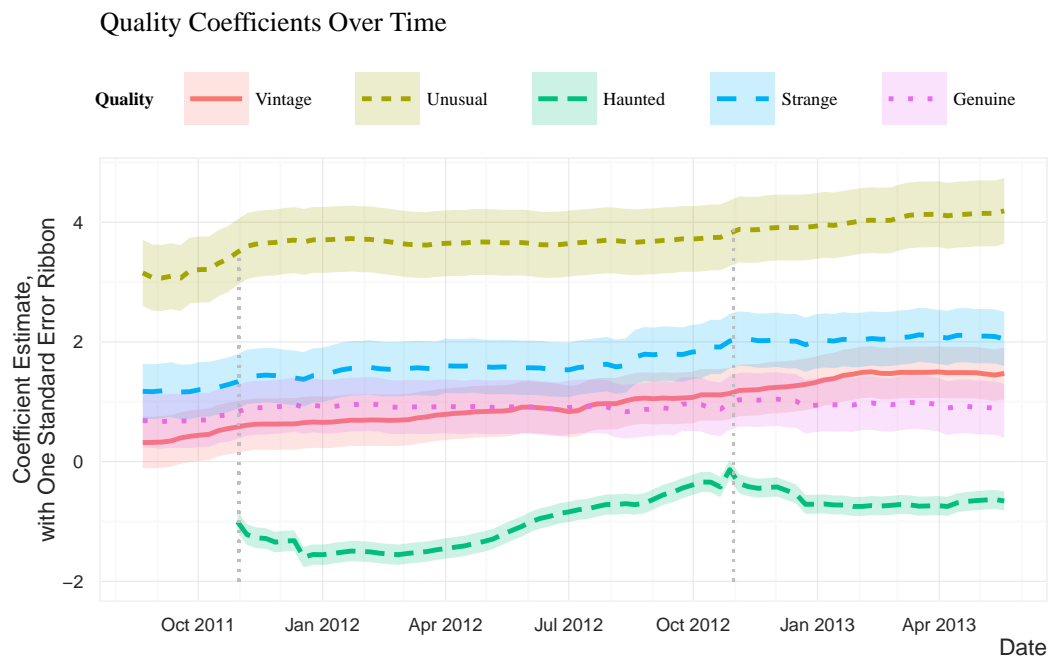
Figure 1.8: Coefficient estimates on time dummies interacted with item quality.

Note: Showing how premiums relative to unique have evolved over the trading sample. Standard error bands shows in transparent ribbons. Halloween 2011 and 2012 are indicated by vertical grey dashed lines.

increase due to the fixed nature of their supply.

1.6.3   Aggregate Value and Growth

Figure 1.9 shows the total nominal value of all items in active players' inventories (what we call "market capitalization") on a daily basis. This is calculated by taking the daily price of each item multiplied by the outstanding quantity in active players' inventories, and summed over all items. We estimate that on the last day of our sample the total value of the economy was approximately 10 million keys – or using a very conservative US Dollar value exchange rate of $2 per key (keys are available on the store at a price of $2.49, which acts as a price ceiling) – $20 million. Expanding stocks to include
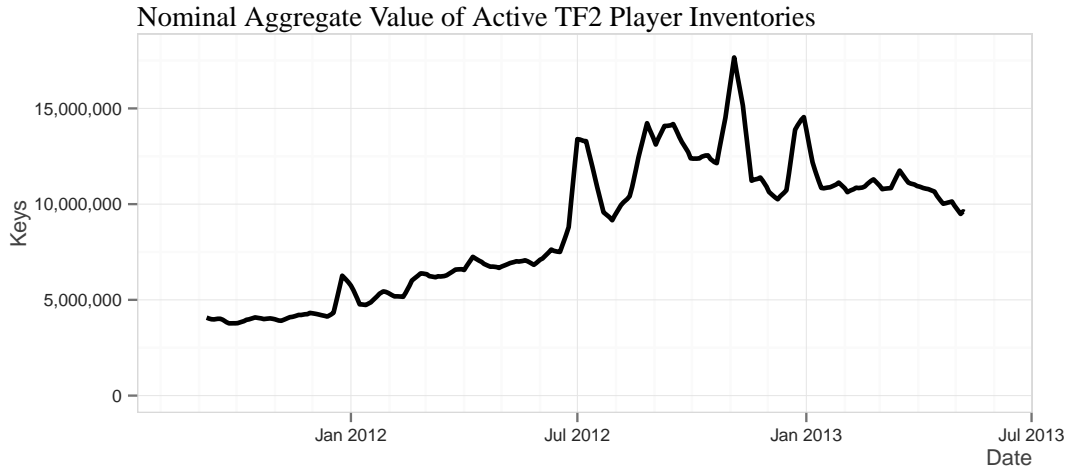
Figure 1.9: Nominal aggregate value of active TF2 player inventories.

Note: Keys are sold on the store for $2.49 each

all TF2 items from all users' inventories, not just active players, market capitalization on the last day is over 50 million keys, or over $100 million. Note that towards the end of our two year sample there appears to be a decline in aggregate value. This is explained by the decline in price level causing the bulk of commonly-held items (usually traded for metals) to drop in value with respect to our numeraire.

Table 1.2: Summary Statistics of Relevant Macro Variables

Date	Traders	Mean IQR	Nom. Trade Value	Chained Price Index
$2011 - 10 - 01$	16179	3.031	45181	109.68
$2012 - 06 - 01$	56493	.380	158427	126.9
$2013 - 04 - 01$	163122	.290	197266	106.7

Table 1.2 displays total traders, mean IQR of prices across all items, total nominal value traded, and the value of the price index for three different dates. There appears to be a trend of increasing population, decreasing price dispersion, and increasing total trade value. We also see that initially the economy experiences significant inflation

31
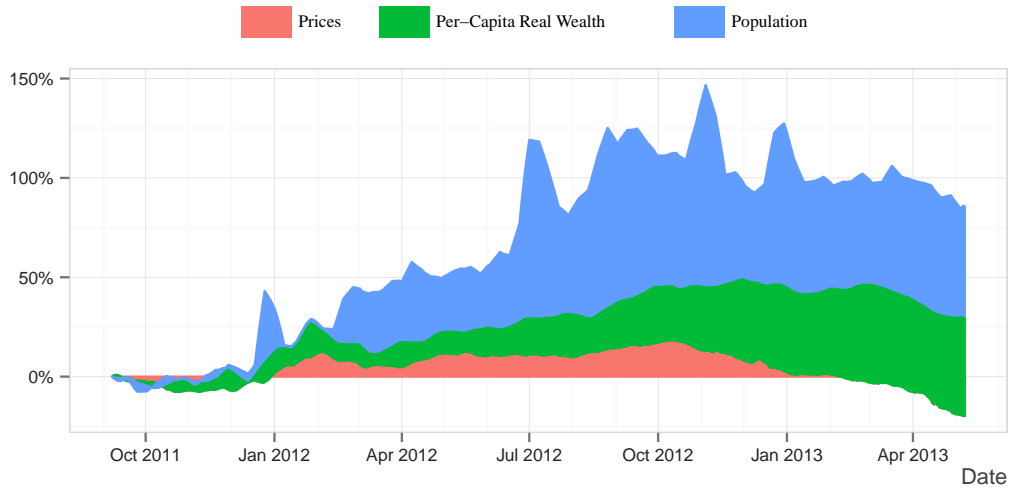
Figure 1.10: Growth of nominal active player wealth

Note: Nominal growth since August 2011. Aggregate nominal value of active player wealth is the product of prices, population, and per-capita real inventory values. The natural log of nominal wealth is thus plotted as the stack of these logged components.

but that this is replaced by deflation by the end of the time period.

In previous sections, we elucidated the trends of the price level and per-capita real wealth. Applying those along with data regarding changes in active population to the decomposition presented in Equation 1.2 results in Figure 1.10. The levels displayed are all in percentage terms with respect to the levels in period 0. E.g. at the beginning of July 2012, the nominal economy is approximately 120% larger than it was at the beginning, of which approximately 10% can be attributed to growth in the price level, 35% of which can be attributed to growth in real per-capita wealth, and the remainder attributed to growth in the number of active players.

We see that real per-capita inventories generally displayed a slowly increasing contribution to the total growth for the duration of our sample. It also shows that

32

practically all of the volatility displayed in Figure 1.9 can be explained by volatility in the population of active players and that there is actually a steady and increasing contribution to economic growth from the real per-capita component. This signifies a healthy and growing economy, even during periods which players are rapidly switching between being active and inactive.

Prices consistently increase after January 2012 until a peak in October 2012, thereafter steadily pulling down net growth until the end of the sample. It can be seen that the contribution from prices disappears (and in fact becomes negative) on precisely the date just after January 2013 at which the price index in Figure 1.6 shows that the price level dips below its starting point of 100. The reason that Per-Capita Real Wealth appears to be negative there is that it compensates for the negative total contribution of the price level on nominal growth starting at that point; this can be interpreted to mean that total contribution of the price level to nominal growth was approximately -15% at the end of the time period.

1.6.4   Nominal Rigidities and the Decline of the Price Level

Here, we present a plausible case in which this decreasing value of metals can translate to a decreasing aggregate price level. We observe that items tend to be primarily traded for a single currency. Low value items tend to trade for metals, mid value items tend to trade for keys, high value items tend to trade for Bill's Hats, and very high value items tend to trade for Earbuds as a result of the indivisibility of these currencies. It is therefore difficult to profit from currency arbitrage across "value-tiers" of
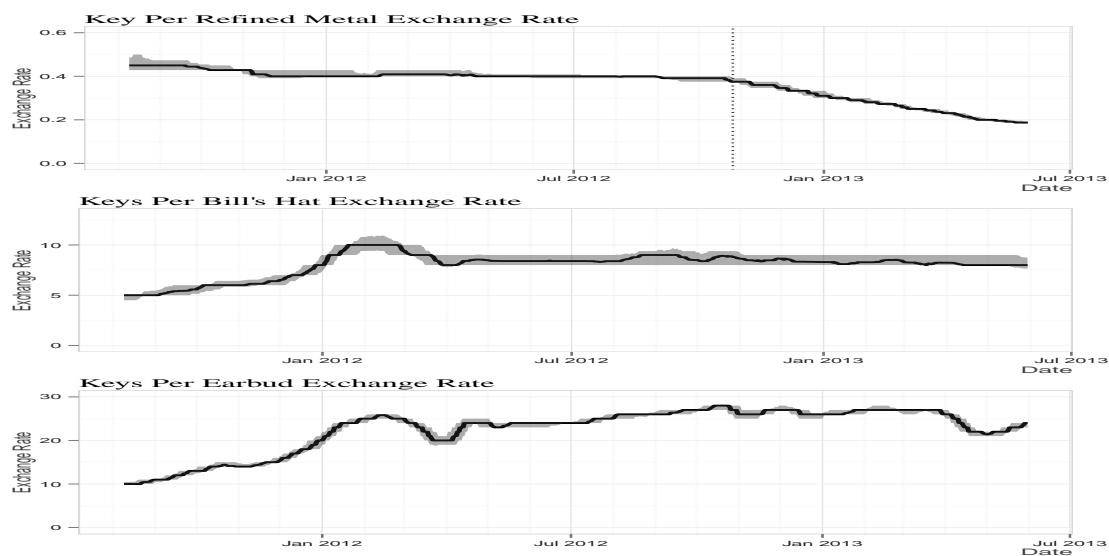
Figure 1.11: Inter-money exchange rates

Note: Daily median exchange rate with three-week smoothing. Grey ribbons reflect first and third quartiles of observed daily exchanges, meaning 50% of trades occurred within gray ribbon. 31 Oct 2012 indicated by a black dotted line in the top figure.

items. It is this combination of price rigidities across currency denominations along with depreciation of metals that may have led to the sustained deflation we observe.

Our best explanation for the deflation towards the end of the sample is monetary and due to the quirks of a barter system with multiple de facto commodity currency goods. See Figure 1.11 for the daily spot exchange rates between keys and each alternative currency. Notice that decline in the price level starts at the end of 2013 – as seen in the price index in Figure 1.6 – syncing up with a sustained appreciation of keys against metals in Figure 1.11. This appreciation is quite significant: at the beginning of our sample it took a little more than two refined metals to purchase a key, but towards the end it took nearly six metals. Thus, the metal-price of keys more than doubled over this period. Also interesting to note is that the path of Bill's Hats/Key and Earbuds/Key

34

exchange rates track each other closely (with a few exceptions near the end of the sample). This may imply that the higher-value currency goods are better substitutes for each other than the low-value metals, and is also likely due to the fixed nature of supply of these goods compared to the increasing supply of metals and keys. A more complete analysis of this potential source of depreciation is presented in Appendix 4.1.6.

To illustrate this point, consider how profitable arbitrage would occur if one currency (metal) is becoming devalued relative to the other currencies but metal prices remained fixed. One would trade metals for goods, then trade those goods for non-metal currencies, then trade the non-metal currencies back for more metal than they started with. This is only worth it if costs associated with trading the goods for non-metal currency is lower than the surplus from completing the cycle.

If these search and transactions costs are large enough, it is not worth it to engage in the arbitrage that would keep prices constant across all currencies. We see that as metal-key exchange rates decline and the value of metal to decreases, this does not appear to fully translate to the metal-price of metal-denominated items. Indeed, we see that for most metal-denominated items, their key-prices fall as metal depreciates. Thus, as the key-price of metals drops, the key-prices of metal-denominated items tend to drop with it. This leads to the component of our quantity bundle which consists of items that are primarily traded for metals to drop in lock-step with the metal depreciation. If this component of the aggregate quantity index is "large", it alone can drive large movements in our aggregate prices.

We argue that this is due to frictions imposed by a barter market. If buyers

were equally willing to pay with keys as metals for the purchase any good, it is likely that the prices of goods as denominated in the more consistently valued currency would be constant and there would be an increase in the price in terms of the currency which sees a declining value. But, if most traders will only offer metals for some subset of goods because it is impractical to trade for goods which are worth a tenth of a key or less using keys or higher value currencies, such a scenario is plausible.

We now present evidence for the presence of nominal rigidities discussed above, which would imply that items which happen to be priced in terms of metals – likely due to their low value and therefore difficulty in trading with indivisible higher value commodity currencies – have their value linked to the value of metals.

We investigate this by linking the frequency that metal is used to pay for items to the price change from Oct 2012 to the May 2013. We estimate the following weighted OLS model:

$$\rho_i = \beta_0 + \beta_1 \cdot m_i + \varepsilon_i$$

In this regression, m_i represents the value proportion of SM trades for item i in which the item trades for metal and thus $1 - m_i$ is the value proportion of trades which the item was exchanged for non-metal currencies. For example, an item that always traded for metals would have an m_i of 1 and an item for which half of the value of trades was from metals and half was from keys, m_i would be .5. The regression relies on value share percentages derived from October 2012 observations and these value share percentages hold a 0.95 correlation with observations in May 2013, implying that these

value shares seem relatively stable over our time horizon. The dependent variable ρ_i represents the percent change of the price of item i with respect to this item's price in October 12, 2012, just before the start of the deflationary period.

The model is weighted by the total value of each item i in the month of October 2012, thus more economically significant items were given heavier weights. We only looked at items for which prices were observed in both Oct 2012 and May 2013, there were 1,288 such items. We remove observations for which percent price changes were above the 99th or below the 1st percentile, leaving 1256 items with prices in both periods.

The interpretation of this regression is straight forward: the sign of the coefficient on m_i tells us if items which were primarily traded for metals tended to undergo price increases (positive β_1) or price decreases (negative β_1) over the period of deflation which started in October 2012.

Table 1.3: Regression Estimates from WOLS of Price Change on the Trading Value Share of Metal.

	Dependent variable:
	Percent Change in Price
Metal Value Share	-0.1867^{***} (0.0406)
Constant	-26.0422^{***} (1.4763)
Observations	1,256
R^2	0.0166

Note: *p<0.1; **p<0.05; ***p<0.01

Our regression coefficients are reported in Table 1.3. It shows that on average, items that traded 100% with metals tended to experience an 18% decrease in price

37
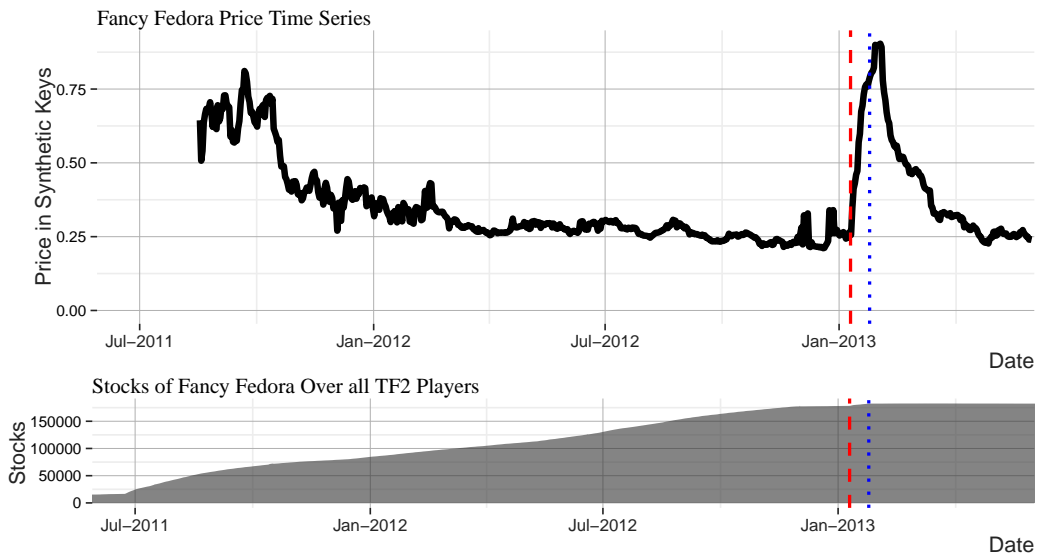
Figure 1.12: Price time series of the unique (normal) quality Fancy Fedora.

Notes: The date of the store retirement announcement is indicated by a dashed red line, and the actual retirement date is indicated by the second dotted line, in blue.

compared to items which never traded for metals. This is evidence that items which trade primarily for metals tend to have a corresponding decline in price. But the decline in price is also less than the decline in the exchange rate of metals (approximately 50% from October 2012 to May 2013, as can be seen in Figure 1.11) which means that this is likely only a part of the whole story.

1.6.5 Response of Individual Items to Micro-level Shocks

We conclude our results with a discussion of the impact on individual items of micro-level shocks. Notice in Figure 1.12, the price of the Fancy Fedora starts high and over a few months drops down and stabilizes, as is typical for newly introduced items.[7]

---

[7]The price time series was generated using trailing price estimates rather than the centered prices discussed above. This was because centered prices cause price estimates to increase before the announcement which is not representative of what was happening in the market on this day.

But at the beginning of 2013, there is a sudden spike in interest. This is driven by a 10 Jan 2013 announcement, as indicated by a red dashed line, that this hat and 8 others would be "retired" on 25 Jan 2013 as indicated by a blue dotted line. Retirement of these items means that they are no longer acquirable except by trading with other players and thus the total supply would be capped at the current level on 25 Jan.

This announcement led to rampant speculation on these items which drove up the price by approximately 120% over the two week time period between announcement and retirement. But this price boost ultimately proved to be temporary as the price falls almost as rapidly as it surged in the first place. This represents the first evidence of a possible speculative bubble in a barter market that we are aware of.

Figure 1.12 also shows the stocks of Fancy Fedoras. On January 10, 2013 there were 178,400, which increased by 2.26% to 182,440 by January 25th. Our best explanation is that there was a sufficient quantity of these hats in existence to satisfy the demand for them for the purpose of durable consumption at the price of approximately 0.2 keys, but the retirement announcement caused a positive demand shock as market participants anticipated a negative future supply shock, driving up current prices (red dashed line). Soon after this negative supply shock took place (blue dotted line), it became clear that the act of fixing supply did not actually do much to shrink quantity available and – as well as the fact that there are likely a large number of close substitutes and the influx of supply by speculators after the January 10 announcement – meant that people interested in durable consumption of the item could simply buy a different hat that didn't see the price more than double. Thus, the announcement and subsequent

39

retirement did not effectively change long run demand and had a small but positive effect on long run supply, so the price returned to its initial level and the speculators that went long on them figuratively lost their shirts.

Another item example demonstrating clear market responses to micro-shocks is the strange Scattergun, a strange-quality version of the default class weapon of the Scout. Strange quality items are notable because they record some sort of player statistic while the player uses the item (e.g. a counter that tracks total number of other players killed with the gun).

On 9 October 2012, as shown in Figure 1.13 with a red dotted line, these stranges were suddenly made available from a newly introduced and particularly ubiquitous series of locked crates and found inside these crates with a probability of approximately 20%. The effect of this policy shock on supply can be seen in Figure 1.13. The total stock on 9 Oct 2012 was approximate 71,000 and had been increasing at the rate of approximately 50 per day for months. After this new crate was introduced the rate of increase of the inventory stock suddenly exploded: after one month there were more than 101,000 strange Scatterguns in existence. And after three months, the stock had doubled.

The impact of this sudden large positive supply shock can clearly be seen in the price of the gun, depicted in Figure 1.13. In contrast to the story of the Fancy Fedora, in which the retirement of an item did not appear to have a long lasting impact on the market supply or demand leading to long run prices being the same as before the retirement, this event obviously actually impacts the long-run supply which causes an
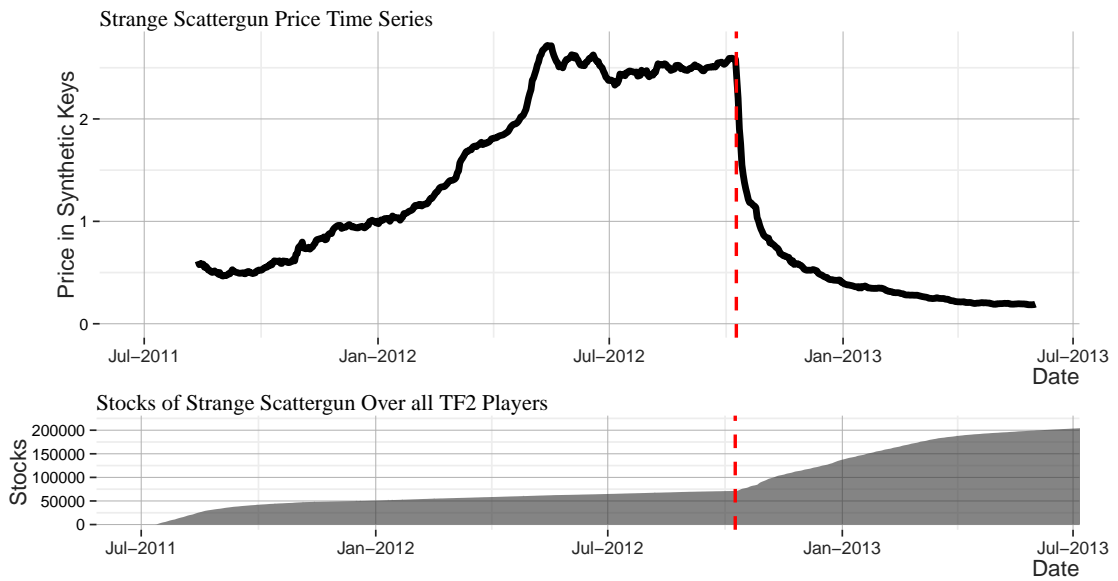
Figure 1.13: Price time series of the strange quality Scattergun.

Note: The red dotted line is at October 9th, the date which the item became more widely available.

unambiguous decrease in long-run market price. Thus, individual prices in the economy do indeed appear to respond to specific micro-level shocks in the ways consistent with basic microeconomic intuition given the direction of the shocks to supply and demand.

1.7   Conclusion

With this work, we present an examination of an economy which is interesting for at least two reasons. First, it is a remarkably rich dataset which documents a true barter market, the likes of which have been pondered by economists for centuries. Second, it's a virtual economy consisting entirely of non-tangible goods which people nonetheless assign value to.

Our primary goal was to calculate macroeconomic growth in this novel envi-

ronment and concluded that an increasing component of nominal growth was due to increases in real per-capita holdings. Per-capita real wealth displays a slow and steady growth for the duration of our sample and most of the volatility in aggregate economic value can be explained by volatility in the active player population.

We presented a hedonic pricing model which we used to impute prices for a Törnqvist price index. We show that not all classes are created equal when it comes to item values. The index indicates that the price level tended to rise until October of 2012, at which point the price level starts declining due at least in part to the declining value of metals. We then traced the source of this depreciation of metals to a shock to both the stocks of metals and keys as well as the rate of increase of their respective supplies.[8] We then demonstrated that items which trade for metals tended to have prices that decreased as the value of metals declined, indicating possible nominal rigidities. But the price decline was less than the decline of the value of metals, so this is likely not the only thing affecting these items. Thus, we did find evidence that macro-indicators responded to macro-level shocks.

Finally, we find in these virtual economies evidence of the same sorts of forces which evidently influence "real world" markets in our micro-level case studies. If a credible central authority makes some decree that could increase expectations of future prices, prices move in that direction. If there is a sudden exogenous positive supply shock in the market for a specific good, the price of that good falls. These goods are but two of many items that have been impacted by idiosyncratic shocks, and their behavior

---

[8]See Appendix 4.1.6

is mirrored in similar goods which were subjected to similar shocks. None of this news should be surprising, but it supports our position that other such virtual economies (which are certainly only going to become more common in the coming years) are fertile ground for further research and the fact that these virtual economies will typically have impeccable record keeping should be enough to get researchers excited.

Future work will investigate the emergence and evolution of number of fundamental market institutions in the tradition of Radford (1945), Burdett et al. (2001), and Lankenau (2001) and we will search for the origin of media of exchange and the development of trade intermediaries by mapping trade networks and behaviors of these intermediaries. In doing so, we hope to answer questions related to how much surplus such intermediary activity brings to the economy as a whole, and how is that surplus is distributed amongst various types of users, deep questions that go to the heart of classic economic inquiry[9] and are issues which many modern economists have struggled to answer empirically.

---

[9]See Smith (1776), Jevons (1885), and Menger (1892)

Chapter 2

Emergenge of Networks and Market Institutions in a Large Virtual Economy

Written with coauthors Curtis Kephart and Daniel Friedman.

2.1 Introduction

How and when do new institutions emerge to facilitate trade, and how can we measure their impact? Such questions are classic but have new urgency in the early 21st century, as markets more tightly bind together economic activity across the planet, and mobile communications enable new ways to transact.

This paper makes a small empirical contribution pertaining to those large questions. In September 2010, Valve Corporation launched a high-performance pure barter trading platform for the user community of one of their more popular games, Team Fortress 2. Our data cover every transaction on that platform over a 661.4 day interval,

involving thousands of different types of goods and 1.9 million traders.

We analyze those data with several classic questions in mind. Given its best conceivable shot, how stable is barter? Does the institution of bilateral barter unmediated by currency survive or does a monetary institution push barter exchange to obscurity? Were Adam Smith (1776, Book 1 Chapter 4), William S Jevons (1885) and Karl Menger (1892), among others, correct in predicting that commodity money will emerge to solve logistical problems inherent in barter? Do we see a unique medium of exchange? Do trade specialists emerge as the market grows, as would seem to follow from the opening argument in Smith (1776)? If so, what kind — dealers (who carry inventory)? brokers (who don't)? speculators? In general, do we see institutions emerge that lower transactions costs?

Despite recent theoretical elaborations such as Kiyotaki and Wright (1989) or Ostroy and Starr (1990), and agent based simulations such as Howitt and Clower (2000), these classic questions have provoked remarkably little empirical work. Perhaps the best known is Radford (1945), who showed that cigarettes emerged as medium of exchange in a WW2 POW camp.

Our empirical investigation is also motivated by network-theoretic questions. What network architectures characterize barter versus monetary exchange? Or direct trade versus intermediated trade? Which network metrics can best demonstrate how goods networks evolve over time? Or how trader networks evolve?

A large and heterogeneous literature on economic networks has recently begun to emerge; see Jackson (2010) and Easley and Kleinberg (2012) for recent overviews.

This literature begins to address some of the classical questions. For example, Choi et al. (2014) develop a network model in which nodes offer intermediation services at posted prices. For given network architectures, they characterize equilibrium prices and argue that node centrality is key to understanding efficiency and division of surplus.

Unfortunately, network theory has not yet advanced sufficiently to provide sharp predictions on network evolution that can be tested on our rich data set. Our empirical study is therefore largely descriptive. By the same token, an advantage of our study is that it may spur theoretical advances to deal with distinctive aspects of the data. These include (a) large, non-uniform, changing networks rather than static networks that are uniform (e.g., cellular automata) or small, (b) two different but interrelated networks — for goods and for traders — defined endogenously by actual transactions data. We find modifications of recent network metics that seem to capture the classical notions of money as a medium of exchange and of various sorts of intermediation, and document how they evolve as the virtual economy matures.

We are aware of three related empirical lines of research, all of them mainly descriptive. A series of articles including Kirman (1997) and Kirman and Vriend (2001) study the market for fresh fish in Marseilles. For this differentiated perishable commodity, the authors focus on the stability of trading relationships between a few dozen sellers and several dozen buyers, using a sparse sample of periodic data. We need different techniques to study essentially continuous trade (on average, nearly a trade every second) in our much larger networks for exchange of homogeneous durable goods.

Bech and Atalay (2010) use federal funds (overnight) loan data to construct a

46

trader network among US banks. They look mainly at a directed unweighted network averaged over time (although sometimes they consider weighted edges or time trends), and confirm stylized facts about node degree distributions. Our network metrics overlap with theirs, but our focus is on undirected weighted networks (although sometimes we consider directed or unweighted networks) and how they change over time. Some of our findings, including those on node degree distributions, stand in contrast to stylized facts established for other sorts of economic or social networks.

Castronova (2001) is among the first to examine the economics of on-line "virtual worlds", applying standard economic metrics to players of the video game Everquest. Chapter 1 similarly studies Team Fortress 2, generating prices from pure barter transactions and generating macroeconomic aggregates. Lehdonvirta (2005) and Lehdonvirta (2010) critique existing studies of markets for virtual items tied to on-line games. Our study emphasizes the rather different insights that can be obtained from network analysis.

Section 2.2 sketches relevant aspects of modern network theory, including metrics such as node degree and strength, network assortativity, and betweenness and closeness centrality. It also shows how a set of barter transactions can be used to define a goods network as well as a trader network. The following section describes the data briefly; see Chapter 1 for more details. Section 2.4 presents results, beginning with an overview of trading volume, whose US dollar value averages well over 2 million per week over the second half of the sample. The analysis of trader networks discloses economically interesting violations of scale-free distributions for node strength, and the emergence of

47

several different sorts of trade specialists. The analysis of goods networks discloses the emergence of commodity monies — not one, but several. Price dispersion decreases over time but remains substantial. A concluding discussion is followed by an Appendix with self-contained formal definitions of network metrics and with supplementary data analysis.

2.2 Network Concepts

We first informally present some useful network metrics. Then we sketch how to construct empirical networks from a set of transactions, and note the differences we would see between classic monetary trade and classic barter, and between direct exchange and intermediated exchange. Formal details are collected in the Appendix.

2.2.1 Network Metrics

A network consists of a finite collection of nodes $i = 1, ..., I$ and edges (or links) $y_{ij} \geq 0$ between ordered pairs of nodes. The most general sort of network we will consider is called weighted and directed, meaning that we keep track of the numerical value (or "weight") of each edge ij, and that the weight from node i to node j may differ from the weight in the opposite direction, from node j to node i. We work mainly with undirected networks, for which these two weights are always equal. Sometimes we consider the familiar subcase of unweighted (or binary) networks, where $y_{ij} = y_{ji}$ is either 0 (no edge connects the two nodes) or 1 (that edge does exist).

Node strength and degree. The strength of node i in an undirected weighted

48

network is the sum of its edge weights. Its degree k_i is the number of edges of positive weight that include that node, a nonnegative integer.

Assortativity. Do strong nodes tend to connect directly to other strong nodes, rather than to weaker nodes? A positive answer suggests that the network may be like our galaxy, with a weighty core and gossamer periphery. A sufficiently negative answer, on the other hand, may hint at some sort of specialization, e.g., internet service providers and customers.

An assortativity metric is, in essence, the correlation of the strengths of each edge's two nodes. In familiar binary networks, node strength is simply node degree, and it is customary to define assortativity there as the correlation across all existing edges of excess node degree, that is, netting out the edge in question. This removes a positive bias that otherwise would give random networks a positive assortativity. As explained in the Appendix, the same logic requires using excess strength for computing the assortativity A of a weighted network, and we do so below without further comment.

Centrality. A node is "central" if it is in some broad sense relatively important, e.g., if it is on lots of shortest paths. Shortest paths in a binary network are found simply by counting edges, but in weighted networks one should take into account the edge weights. We follow standard practice in adding the reciprocals of edge weights to obtain path lengths. The betweenness centrality $B(n)$ of node n is the fraction of all node pairs ij that have a shortest path that goes through n. For example, consider a star-shaped network, for which all other nodes have edges with, and only with, a special node n^*. Then $B(n^*) = 1$ and $B(n) = 0$ for all other nodes n.

49

An alternative intuition is that a node is central if on average it has a short distance to other nodes. We define the closeness centrality $C(n)$ of node n as the sum of inverse distances to all other nodes in the network, normalized by the size of the network, which is not constant in our data. Thus $C(n)$ will increase as distances shorten between typical pairs of nodes.

2.2.2 Transactions, Trader Networks and Goods Networks

Suppose that trader i initiates net trade $x = (x_1, ..., x_N) \in R^N$ with counterparty j. That is, there are N different goods, and i transfers a bundle x^- to j and receives bundle x^+ in exchange, where (by convention) those two bundles have non-negative components and the net trade vector is $x = x^+ - x^-$. Given an N-vector p of positive prices, the value of the bundle i acquires is $v^+ = p \cdot x^+$ and the value of the bundle j acquires is $v^- = p \cdot x^-$. The transaction is budget-balanced at p if $v^+ = v^-$ or, equivalently, if $0 = p \cdot x \equiv \sum_{i=1}^{N} p_i x_i$.

Suppose that transactions $k = 1, ..., K$ are observed over some given time interval. Given price vector p, the observed trader network is a weighted directed network constructed as follows. The nodes consist of all individuals who participated at least once in those transactions, either as initiator or counterparty. The directed edge weight from node ℓ to node m is the total value of the bundles that individual ℓ acquires from m. If all trades are budget-balanced, then the trader network is automatically undirected.

If some transactions are not budget balanced, then we can recover an undirected

network by replacing each v^+ and v^- in the total value calculation by

$$v = \max\{v^+, v^-\}. \tag{2.1}$$

This convention makes good sense if the main reason for an imbalance is that the record keeping system missed an element of the transaction such as an explicit or implicit promise to repay. If instead the main reason were random noise in goods valuations (perhaps due imprecise pricing or indivisible units) then a better convention would be $v = (v^+ + v^-)/2$. Of course, we can always recover an unweighted undirected ("binary") network by assigning weight 1 to any edge with positive weight (or, alternatively, any edge with weight exceeding some specified positive threshold) and keeping the other edge weights at 0.

The same set of transactions also defines a goods network. The nodes of this network are the subset of $n = 1, ..., N$ that have a non-zero component in at least one of the K transactions. The edge weights reflect the total value of transactions in which one good is part of the exchange for another, and are automatically undirected when we apply convention (2.1). The edge weight is clear when just one good appears on each side of the trade. The Appendix shows how to allocate values across pairs of goods when there are several different goods on one or both sides of a transaction.

The Appendix also shows that two matrices, one for goods and the other for traders, characterize the network structure we employ here. Data permitting, a more complete characterization would be as a tensor of a higher order that specifies, for each

transaction, the characteristics of both traders involved and of the two bundles that they exchange.

2.2.3 Classical Networks

Pure barter can be idealized as transactions in which each good is equally likely to be traded for any other. Then (apart from sampling error) the goods network would be completely connected with edge weights proportional to the strength of each node. Assortativity for the network would be near zero, and the distributions of betweenness $B(n)$ and closeness $C(n)$ would be be broad and unimodal if the goods' overall value shares are uniformly distributed over a wide support.

At the other extreme, money mediation can be idealized as a single good, $n^* = 0$ say, such that transactions take either the form $(-m, x^+)$, i.e., the initiator buys the bundle x^+ for $m > 0$ units of good 0, or the form $(m, -x^-)$ in which the initiator sells the bundle x^-. This defines a star-shaped goods network around node 0. Assortativity would be quite negative, betweenness would be 1.0 for the money good and zero for everything else, and closeness would be high for all goods, especially the money good.

Trader networks can reveal institutions that facilitate transactions. At one extreme, there could be uniform bilateral trading relationships, whereby any trader is equally likely to trade with any other. This again would give us a fully connected graph with metrics similar to those just described for barter. At the other extreme, there could be a universal store, designated as trader 0. If the other traders sell their bundles only to the store and buy bundles only from the store (or even conduct barter transactions

Figure 2.1: Item Barter Trading Platform

Here the initiating trader (You) offers a game license and one unit of refined metal and the counterparty (Test4321) offers ten different hats. The transaction occurs when both traders check the central blue box.

but only with the store), then again we have a star-shaped network. Less extreme forms of intermediation will leave traces in the network metrics, but supplementary analysis will be necessary to distinguish among various sorts of intermediaries such as market makers, speculators and brokers.

2.3 The Data

As explained in more detail in Baumer and Kephart (2015), our data pertain to the video game Team Fortress 2 (TF2), sponsored by Valve Corporation. Launched in October 2007 using a standard computer game business model (revenues mainly from players purchasing permanent rights to access the game), Valve made TF2 "free-to-play" — i.e., zero price for game access rights — in July 2011, gaining revenue from a company store (dubbed the "Mann Co. Store") where some popular items could be purchased for US dollars or other national currencies. It should be noted that the traded items in TF2 with appreciable value, unlike those in many other games, offer little or no direct advantage in playing the game; they are mainly to establish identity or to make a style statement.

TF2 took a new turn in September 2011 with the advent of the barter ("Steam Trade") platform shown in Figure 2.1; beta versions had been seen a few months earlier. Our data consist of all barter transactions beginning in August 2011, when the current accounting system was first introduced, through May 2013 — over 40 million bilateral barter transactions involving nearly 2 million unique trader identities and over 1000 distinct types of items (or goods), gathered in a relational database of more than 0.5 terabytes. Table 2.1 shows a tiny truncated extract.

In most respects the data align well with the theoretical structure introduced in Section 2.2, but there are some notable quirks. Writing $x \in R^N$ there suggests that goods are perfectly divisible, while in TF2 each good has indivisible units (each with

Table 2.1: Example data snippet.

TradeID	PartyA	PartyB	Time	AppID	AssetID	NewID	Origin	EconAssetClass
1	1203	1876	234	440	3881	4120	1	100
2	4256	172	245	440	3942	4136	0	1949
2	4256	172	245	440	4135	4137	1	1585
3	993	8384	250	440	4133	4138	0	207

Party A is the initiating trader i, Party B is the counterparty j for the trade at Time (only last three digits shown) t_k, where the TradeID k appears in the first column. AppID 440 refers to TF2 transactions, AssetID and NewID are tracking numbers for particular units of the item (or good) specified in EconAssetClass. Origin is the indicator variable that the good is a positive element of x^-. Note that a transaction as defined in Section 2.2.2 may correspond to several lines in the database, e.g., trade $k = 2$ here consists of the second and third rows.

its own AssetID). Section 2.2.2 discussed pure exchange of perfectly durable units in fixed total supply, but in TF2 new units of goods appear randomly or after achieving certain game milestones, and certain goods can be purchased from a company store. Some goods can be produced by consuming others. For example, a player can convert two weapons into one unit of scrap metal, convert three scrap metals into to one unit of reclaimed metal or the reverse, and convert three units of reclaimed metal into one unit of refined metal or the reverse. These metals can be combined with other goods to produce new units of designated goods via known production recipes. Also, a treasure chest (or crate) can be opened via a purchasable key to produce its (heretofore hidden) contents, with the key and crate irreversibly consumed. In a sense these crates are like scratch-off lottery tickets: for a modest price — one key and one crate — you can immediately see what prize that you won; the dollar value of the prize ranges from a few pennies to (very rarely!) over $10,000.

The barter trading platform allows trade in items other than those used in TF2, such as the Left4Dead2 game license seen in Figure 2.1. Of the 70 million transactions

we observe, 44 million involve at least one TF2 item, and nearly 41 million involve only TF2 items. Transactions involving non-TF2 items were more common near the end of the period covered by our data, and are excluded from our analysis.

2.4 Results

Before analyzing how network metrics evolve over time, we take a look at the overall growth of the exchange economy. Figure 2.2 shows that the number of trades K rose from below 100,000 in the first week to more than 500,000 per week a year later, and remained above that level for the rest of the sample. Similarly, the number of unique trader identifiers active each week was approximately 25,000 in the first few weeks and increased to 200,000 within a year, leveling off thereafter. (For the Steam Trading platform as a whole, growth trends continued unabated but, as noted earlier, mainly for games other than TF2.)

Figure 2.3 shows roughly similar trends for the weekly value of trade. Values are determined using the daily price vector, the construction of which is described in Chapter 1. The unit of account is a key, which over the entire sample period could be purchased at Valve's store for US$2.49 or the equivalent in other national currencies. A substantial fraction of transactions are not budget balanced; indeed 41% are one way transactions, with either $v^+ = 0$ or $v^- = 0$.

Weekly trade value bounces around in the 1 - 1.5M key-equivalent range during the last year, or about US$2.5 - 3.75 million. Trade value is highly correlated with trade
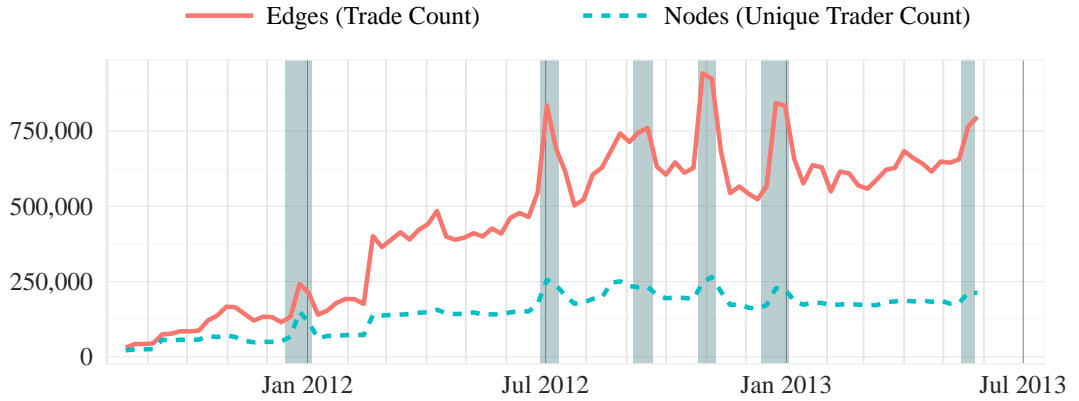
Figure 2.2: Weekly transaction counts and unique trader count

Vertical shaded bars indicate promotional events

count ($\rho = 0.966$) and with unique trader count ($\rho = 0.938$), but the trade values are less sensitive than trade counts to special promotions. It seems that most promotional items have low prices, and those with high prices have low trade volume. Promotions also seem to bring an inflow of new market participants who trade relatively low value items.

2.4.1 Trader Network

Node Degree Distribution. Figure 2.4 shows the distribution of node degree (k_ℓ = number of counterparties of trader ℓ) in trader networks obtained from all transactions in a single week. Each panel shows the degree histogram in log-log scale. For all 6 weeks shown (and all other weeks as well), most of the histogram declines roughly linearly (in logs), suggesting that the Pareto distribution (known by physicists as power law and by network theorists as scale-free) dominates here as it does in so

57

Figure 2.3: Weekly total trade values

Values are measured in key-equivalents; they can be multiplied by 2.5 to get approximate US Dollar equivalents.

many technology and social networks (e.g., Barabási and Albert, 1999, Pastor-Satorras and Vespignani, 2001, Liljeros et al., 2001). The slope seems perhaps less steep in later weeks, suggesting that the Pareto exponent γ may decrease over time.

Beginning in Panel C, we see something different and more economically interesting. Starting in early October 2011 a group of traders emerge who trade with an order of magnitude larger set of counterparties. This subpopulation tends to become larger and more disconnected from the main mass over time, and represents a qualitative departure from the usual Pareto distribution. We are witnessing the spontaneous emergence of large traders, who have thousands of counterparties every week.

Assortativity. How can we assess the economic impact of large traders? Our first step is to check trends in assortativity. Figure 2.5 traces the standard (excess degree) measure of assortativity $A(W^b)$ in the unweighted undirected ("binary") trader network W^b computed weekly from transaction data. In the first few weeks, assortativity is surprisingly positive, indicating that at first traders with many counterparties tended

58
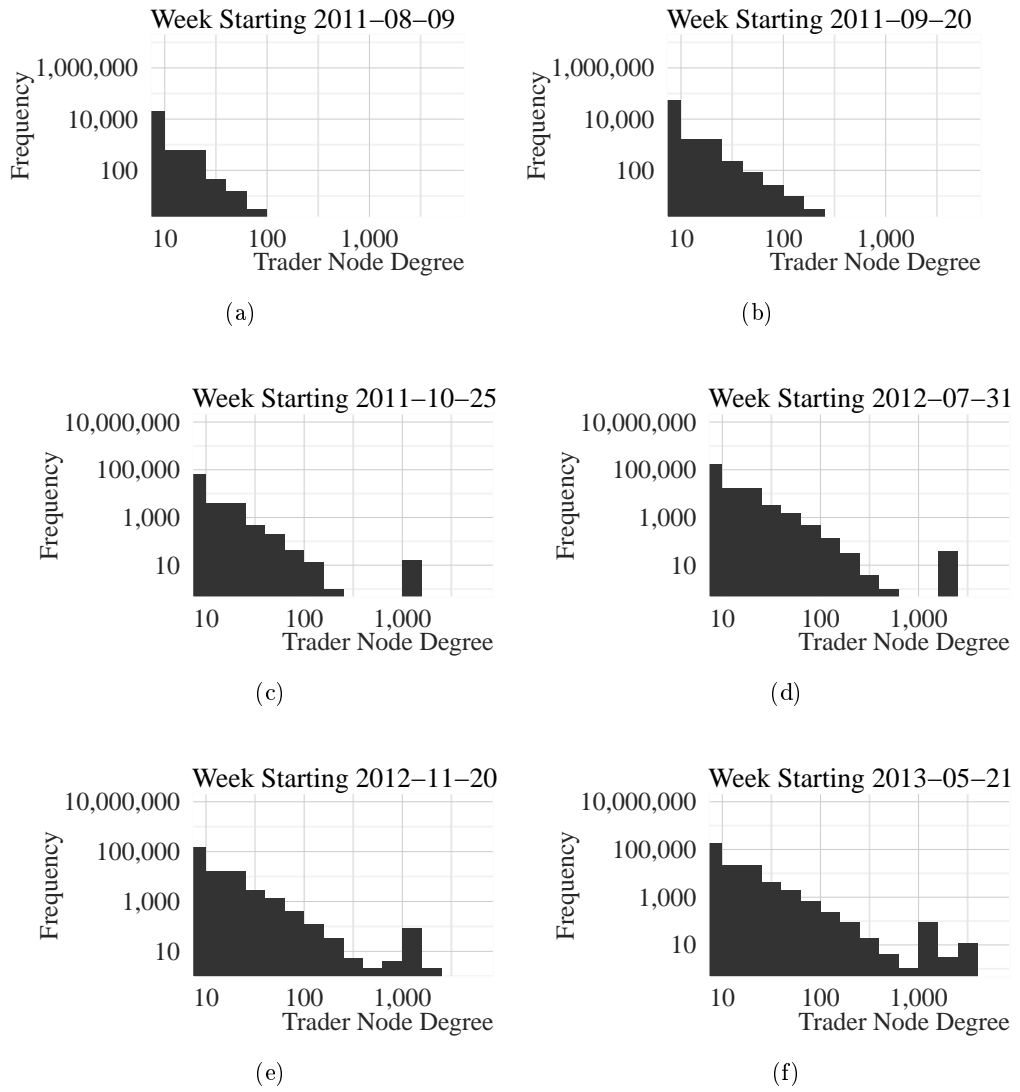
Figure 2.4: Degree distribution in trader network (log scales).

Figure 2.5: Weekly assortativity

$A(W^b)$ in the unweighted trader network. Horizontal dashed lines show benchmarks from Newman (2002), and vertical shading indicates promotional events that attract new traders.

to trade with others of high degree, and traders with few counterparties tended to trade with each other. The level is comparable to the most positive benchmarks in Newman (2002), including movie actors. (In the network whose edges indicate whether the actors have appeared together in the same film, the big name actors tend to work with other big names, hence the positive assortativity.) The zero assortativity benchmark is a random graph.

Over the next six months, trader network assortativity plummets towards Newman's lowest benchmarks, world wide web links (-0.065) and internet wiring (-0.189). (The latter is negatively assortative since individual homes and businesses mostly connect to internet backbones.) The downward trend ends with a 2011 Winter holiday event that brought an influx of new traders. After February 2012, assortativity mostly

60

bounces around in a negative range bounded by the internet wiring and hyperlink benchmarks. Most upward jumps are short lived and coincide with special promotions. (A likely mechanism for those jumps is that a raft of new promotional goods and the accompanying influx of new small players increases the average willingness of small traders to search and match with other small traders.)

The weighted network shows similar trends. Weighted assortativity in the first month of our sample was 0.1295, very significantly positive relative to a random graph with identical edge count and edge weight distribution ($p < 0.0001$). Weighted assortativity then turns quite negative and is usually in the range between world wide web links and internet connections; in the final month $A(W^s)$ is -0.0959, again quite different from zero ($p < 0.0001$). See Appendix formula (4.6) and following discussion for details.

What drives the broad trends? Figure 2.6 takes a finer grained look, computing weighted assortativity in undirected weighted trader networks built from two different subsets of the transaction data. The solid red line shows assortativity $A(W^{hi})$ based on the transactions more valuable than or equal to the median for that week, and the dashed blue line does the same for the network W^{lo} constructed from remaining (lower than median v) transactions.

In the first few weeks assortativity is between 0 and 0.15 for both subsets. Thereafter, the red line bounces around a modestly upward trend, suggesting that big trades tend to occur mainly between big traders. At the same time, the blue line quickly trends down and eventually settles in modestly negative territory. Thus it appears that,

61
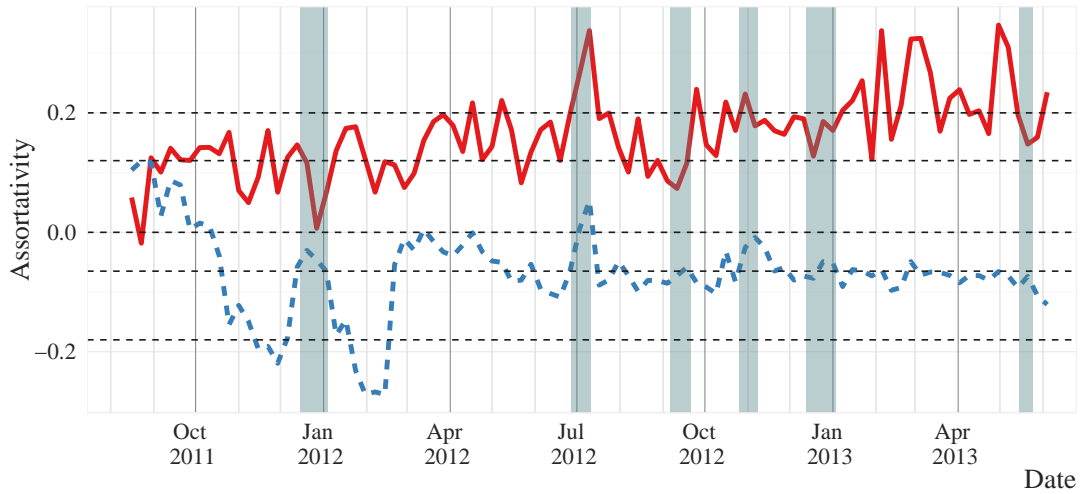
Figure 2.6: Weekly assortativity

$A(W^s)$ for weighted trader networks built from high value ($s =$ hi, solid red line) and from low value ($s =$ lo, dashed blue line) transactions.

after the market matures, traders who want to exchange goods of relatively low value turn to large traders, perhaps specialists, who are willing to accommodate them. Both of these subgraphs are statistically different to a random graph with identical edge count and edge weight distribution ($p < 0.0001$).

Market Makers. To better understand the large traders, we sorted them by characteristics such as weekly transaction count and value, frequency of one-way trades, and profitability. The group that emerged in October 2011 consisted of 88 unique trader id's, each of whom trade "inhuman" quantities, working 24 hours a day 7 days a week. We call them the Clump because they all move closely together in exploratory animations of trader activity. All evidence (see the Appendix for more details) indicates that the Clump consists of 88 automatons controlled by a single economic entity. The clump

initiated over 17.5% of all TF2 transactions in our sample and had an overall gross profit margin (value received minus value delivered divided by the sum of value received and delivered) of roughly 2.1%, with a slight declining trend. Players' first trades with the Clump almost always (in 96.9% of cases) were one-way inward, delivering value to the Clump. Subsequent trades were commonly one-way and about half were outward, and only about 18% of these were for a good previously delivered to the Clump. We infer that the Clump provides some warehousing services but is primarily an inventory-carrying market maker for a broad range of goods, and that it grants trade credit secured by customers' deposits.

A second sort of large trader emerged in October 2012 that specialized in 2-way trades. On closer examination, these traders predominantly accepted piles of junked weapons in exchange for metal at or near the conversion rate (2 weapons $\Leftrightarrow$ 1 scrap metal) available to ordinary traders. Apparently these specialists are not really exchange intermediaries, but rather offer a convenience on the production side, analogous to CoinStar machines at grocery stores that give dollar bills in exchange for piles of coins. (In Chapter 1, it was noted that a sustained depreciation of metals relative to keys began in October 2012. Was that a coincidence? Available evidence is inconclusive.)

Closely associated with this CoinStar entity is a single account that specializes in trading metals for keys and the reverse, beginning in December 2012. Essentially all of its 1,500 counterparties are among the 10,000 accounts that utilized the CoinStar service. We refer to this account as the MoneyChanger for reasons that will be apparent in the next section. We note here that the weightiest edge in the goods network is

keys-metals, and that ever since it first appeared MoneyChanger has accounted for 10% of this edge weight. MoneyChanger's weekly average spread between buying and selling prices is about 2%, and it initiated all but one of its 37,000+ trades in our sample. We surmise that MoneyChanger is the primary market maker in the TF2 economy's thickest market.

Another sort of large trader emerged in late December 2012, eventually controlling 6 trader IDs. Of the nearly 400,000 trades involving these IDs, only a few hundred were one way and the vast majority were budget balanced or very nearly so; gross profit margin was less than half of one percent on average gross trade value of 1.42 keys. These traders had over 30,000 unique counterparties, who averaged more than a dozen transactions, though the distribution was quite skewed and the median was only three transactions. Inspection of individual trades suggests a familiar business model: buy and sell goods at a narrow price spread for a range of standard goods, avoiding large inventories. That is, the six IDs were employed by a market maker who used spot transactions, and did not offer trade credit or take deposits. This spot market maker's share of transactions trended up relative to the Clump, but remained less than half as large, and covered a somewhat narrower range of goods, almost none of them of high value.

Speculators. A (long) speculator builds up inventory of a particular item whose price he expects to rise, and later sells it off, making a profit if his expectation is correct. (Short speculation seems infeasible in TF2.) To detect speculative behavior, we applied the Wald-Wolfowitz runs test (Bradley, 1968) to a given trader's sequence

64

of transactions in a given good. A run ends and a new run begins each time the trader breaks a sequence of consecutive buys with a sell, or breaks a sequence of consecutive sells with a buy. We classify the trader as a suspected speculator if the Wald-Wolfowitz runs test Z-score falls in the $p = 0.001$ lower tail under the null hypothesis of exchangeability (essentially, serial independence). That is, suspected speculators tend to have relatively few (hence relatively long) runs, suggestive of inventory accumulation and liquidation.

Table 2.2: Proportion of All Trades of Selected Items Involving Speculators By Item Quality

Item	All Qualities	Normal	Vintage	Strange	Unusual
Alien Swarm Parasite	2.0%	3.8%	0.8%		
Batter's Helmet	0.3%	0.3%	**16.7%**		0.0%
Demoman's Fro	0.2%	0.4%	0.0%		0.0%
Fancy Fedora	0.1%	0.1%	0.2%		0.4%
Fast Learner	0.0%	0.0%			
Football Helmet	0.4%	0.6%	**12.8%**		1.0%
Ghastlierest Gibus	1.9%	3.4%	0.0%		
Killer's Kabuto	0.2%	0.1%	1.1%		0.0%
Merc's Pride Scarf	0.1%	0.1%			
Mining Light	0.1%	0.2%	0.0%		0.0%
Professional's Panama	0.2%	0.2%	**9.3%**		0.1%
Prussian Pickelhaube	0.2%	0.1%	**11.3%**		0.0%
Pyro's Beanie	0.2%	0.0%	**8.0%**		0.0%
Soldier's Stash	0.2%	0.4%	0.1%		0.0%
Solemn Vow	2.2%	2.5%			
Triboniophorus Tyrannus	0.7%	1.3%	1.1%		0.2%
Trophy Belt	0.1%	0.1%	0.0%		0.1%
Your Eternal Reward	1.1%	1.3%		0.1%	

Suspected speculator trade volume as a percentage of total trade volume for selected items. The four columns furthest right report, respectively, normal (unique), vintage, strange and unusual quality items; first column shows percentage over all four qualities. Bold values are greater than 5%.

We examined the 72 items (18 goods each with 4 qualities) shown in Table 2.2, chosen because price trends seemed conducive to speculation. The share of trades by suspected speculators is mostly well under 5% of overall trading volume. The only

65

exceptions are five of the vintage quality goods where that share reached 5 to 16.7%; all five exceptions are attributable to a single account that traded over 1000 of each of those hats but never had a maximum position of greater than 100. So this suspect was not really a speculator, it seems, but perhaps instead was a hobbyist or an erratic middleman.

So far we haven't found any traders making large profits via speculation; all our evidence suggests that speculation is not a major TF2 economic activity. This may help explain the puzzle of how it was possible to sustain for so many months a steady depreciation of metals against keys.

Brokers. Another sort of trade specialist facilitates trade of valuable items between two parties who trust the specialist but not each other. For example, trader A may agree to send US$100 to trader C in exchange for a very special hat. Trader B (a broker or escrow agent) might agree, for a modest fee, to hold the money transfer until C sends him the hat, and then to send C the money and send A the hat. We have no data on outside money transfers, but can observe B engaging in two one-way trades in rapid succession for the same good; the signature is a short holding time and specialization in particularly high value goods.

We screen for brokers by analyzing the complete transaction history of goods available in unusual quality with a particular effect that tends to command prices of at least 20 keys. We then search for individual accounts which displayed at least twenty episodes of receiving a high-value good in a one-way trade which is then delivered in a second one-way trade within 48 hours. In the first three such accounts we detected, we
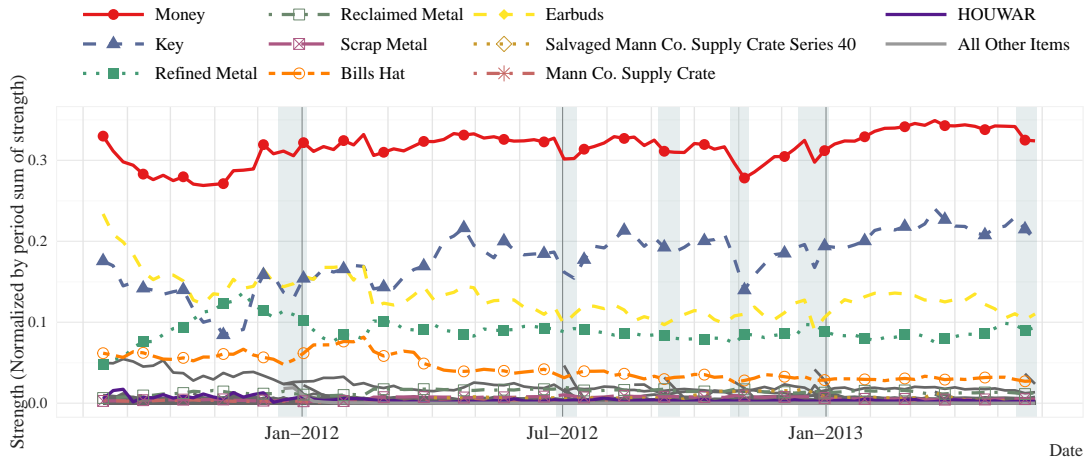
Figure 2.7: Weekly normalized item strength s_i in weighted goods networks

very conservatively estimate the value of brokered exchange at 5000 keys or US$12,500; the median holding time was 7 minutes.

Our sampling of unusual grade goods so far has detected tightly-defined brokerage in 9 to 15% of all of the trades involving these high value goods trades. We conclude that brokerage plays a substantial but not dominant role in TF2's markets for high-end goods.

2.4.2 Goods Network

For goods networks, the main questions are whether traders eventually abandon barter in favor of monetary exchange, and whether transactions costs decline substantially as the market matures.

Strength. As a first clue on barter versus money mediated exchange, we consider item strengths. Figure 2.7 graphs the strengths of every item in the TF2 goods

67

network, normalized so that the strengths of all items always sum to 1.0. Putting aside for the moment the top line, we see that Keys account over 20% of the trade value by the end of the sample. Both Earbuds and Refined Metal are about as important as Keys in early 2012, but by the end of the sample each is around 10%. (Earbuds traded for about 40 units of refined metal at the start of the sample and over 140 units by the end, so refined metal's high strength also indicates very large trade quantities.) These three items are strongest, but another three also show consistent strength: Bill's Hat and Reclaimed and Scrap Metal. The other items that show up on the graph are Max's Severed Head (actually a hat), the Hat of Undeniable Wealth and Respect (HOUWAR), as well as blips for special event keys and crates.

Since metals are convertible in both directions at a fixed rate (three units of lesser-quality metal to one of next higher value type), it makes sense to combine the three grades of metal into a single composite. We tentatively define money as the combination of metals, keys, earbuds and Bill's hats. Although relative prices can vary among these four components, we will see later that they all serve as media of exchange.

Simply summing the four components' strengths, we approach the 50% benchmark for idealized monetary exchange. Does this mean that barter has disappeared? Not necessarily; some of the sum comes from "currency market" trade, i.e., from edges within the set of four (originally 6) nodes. A better indication of the extent of monetization comes from collapsing these four nodes into a single node and calculating its strength in the resulting goods network, as shown in the top red line of Figure 2.7. We see that the four money items jointly account for about 30 to 34% of value flows in the

68
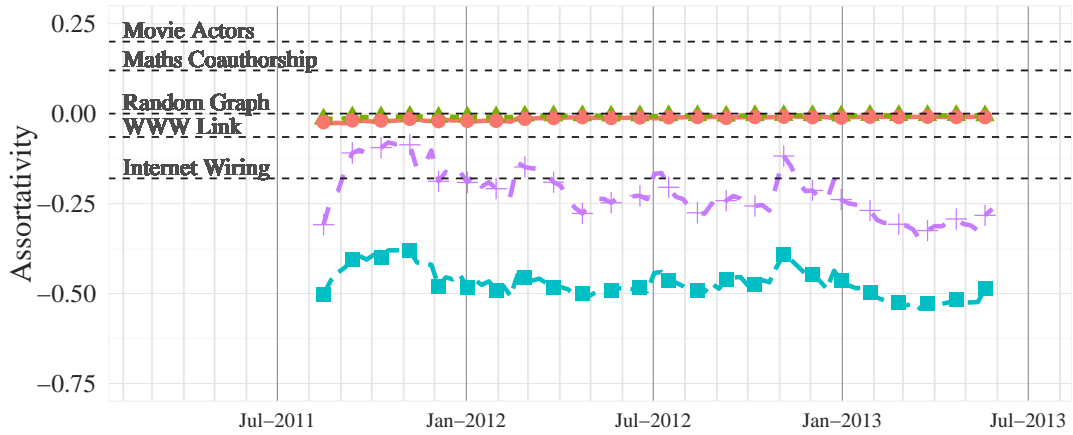
Figure 2.8: Weekly assortativity $A(Z)$ in goods networks

Connected orange dots and green triangles indicate respectively the unweighted (binary) full network and reduced (single money node) network values, while dashed blue squares and dashed purple indicate the corresponding weighted network values.

reduced network — quite a lot but substantially below the benchmark.

For an alternative perspective, see Figure 4.4 in Appendix A1, which shows a combination of node degree and node strength. As one might expect, the lower denomination currencies look more important with this metric, and there is an overall upward trend due to increasing number of distinct traded items over time.

Assortativity. Figure 2.8 shows that unweighted goods networks have assortativity $A(Z^b)$ that remains very close to zero, but that assortativity $A(Z)$ in the weighted networks, computed using equation (4.6), is surprisingly negative, even compared to the internet wiring benchmark. We just saw that the tentative money items have weighty edges with each other, so we again collapse them into a single composite good and obtain extremely negative assortativity, around -0.5. This is a very strong hint that other goods tend to trade with this composite good, so it may indeed be the main

69
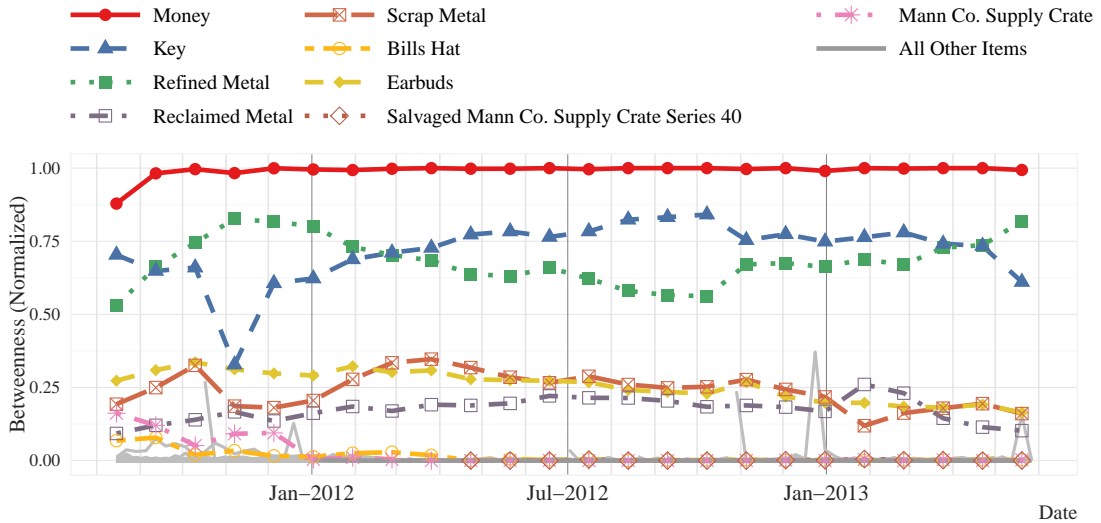
Figure 2.9: Weekly betweenness centrality $B(n)$ in weighted goods networks

The top line pertains to the network with six nodes collapsed to one ("Money") node; other lines pertain to the full goods network.

medium of exchange.

Betweenness. The most definitive evidence on money comes from the betweenness metric $B(n)$. Figure 2.9 shows that $B(n)$ is essentially zero for the vast majority of goods. All long-lived exceptions are among the tentative money goods, especially keys and refined metal, each with betweenness usually in the 60-80% range. (Short-lived exceptions are mostly event keys and crates, which seem to be substitutes for ordinary keys as media of exchange — their up spikes in Figure 2.9 coincide with down spikes for keys, especially around the 2012 holidays and the end of the sample.)

Once again, we reconstructed the goods network with the same composite money good as before. The top red line of the Figure shows that even in the first week over 85% of trade by value went through composite money, and within a few weeks,
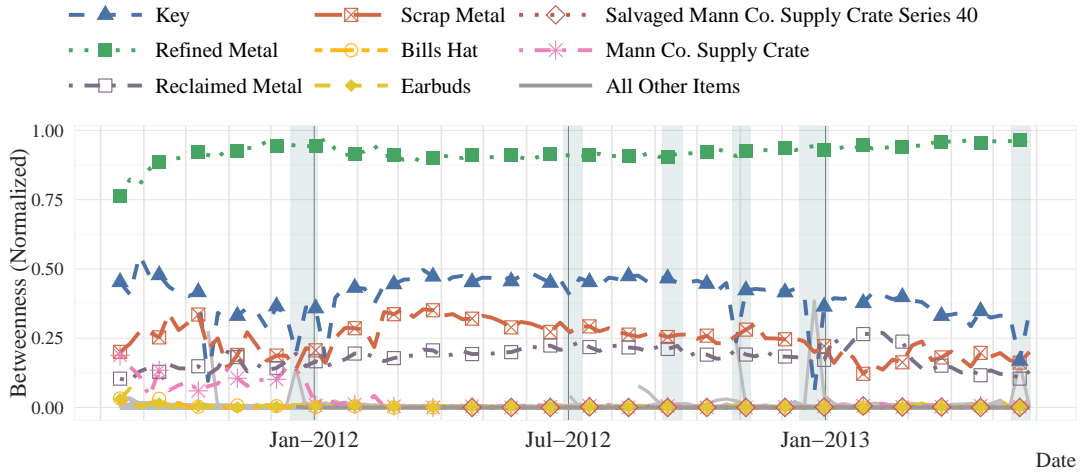
Figure 2.10: Weekly betweenness centrality $B(n)$ for low value transactions

Involving only items valued below 0.95 keys

virtually all trade did so for the rest of the data sample. We interpret this as conclusive evidence that the TF trading platform completed its transition to monetary exchange by October 2011.

The question now remains, why are there four money goods (or six, including separate grades of metal) rather than just one? Some relevant evidence appears in Figures 2.10 - 2.12, which break down the goods networks by the maximum price of the (non-money) goods involved in the transaction.

Metal, especially refined metal, is the main sort of money used in the low value transactions. Figure 2.10 shows that keys also play a role here, and that special event keys are close substitutes when they appear. For mid-tier items, Figure 2.11 shows that keys have dominated since late 2011; refined metal plays a supporting role that diminishes over time. For top-tier items, Figure 2.12 shows that earbuds and Bill's hats

71
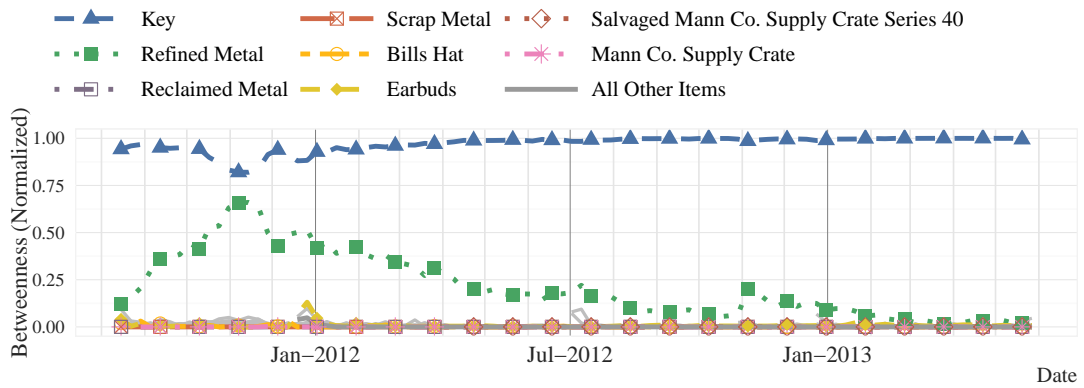
Figure 2.11: Weekly betweenness centrality $B(n)$ for medium value transactions

Involving only items valued valued between 0.95 and 5 keys.
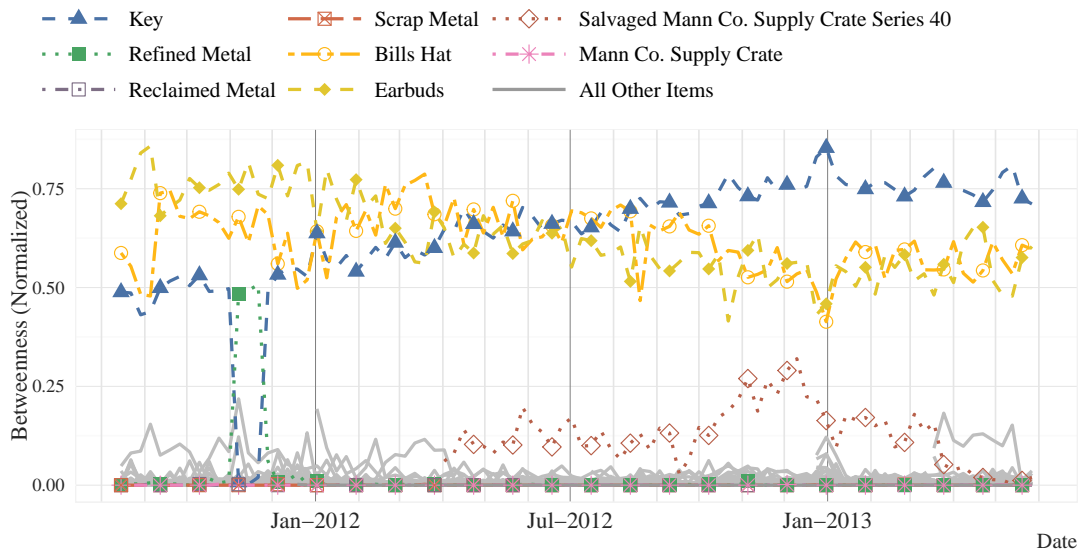
Figure 2.12: Weekly betweenness centrality $B(n)$ for high value transactions

Involving only items valued valued above 5 keys.
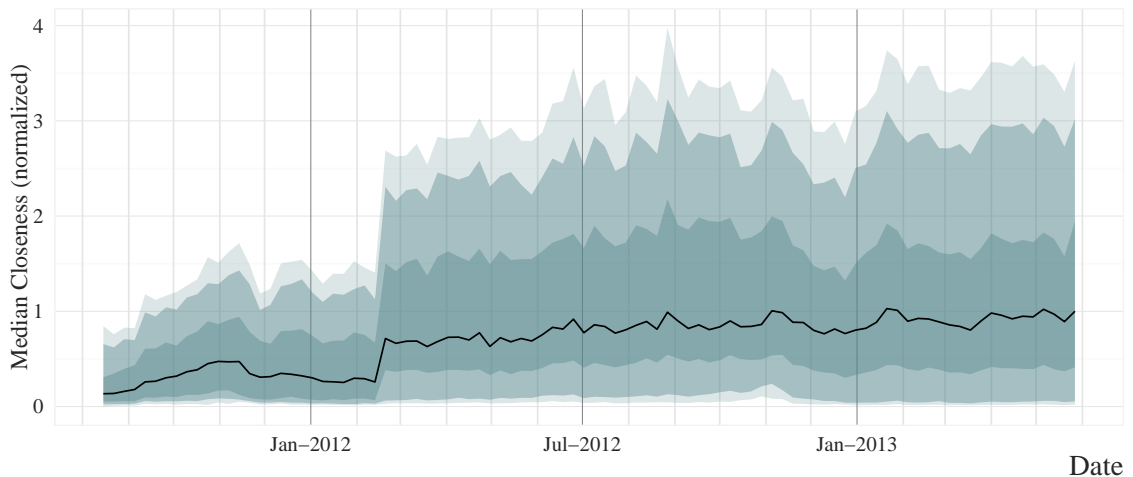
Figure 2.13: Weekly normalized closeness distribution for the weighted goods network

Computed via equation (4.8) divided by the weekly number of goods, and scaled so that median is 1.0 in the final week. Line is median and shaded ribbons emanating from median span 25th - 75th, 10th - 90th, and 5th - 95th percentiles.

come into their own, though keys are almost as important early on and eventually attain the highest betweenness centrality even in this segment.

Closeness. Figure 2.13 shows the distribution of normalized closeness $C(n)$ on a weekly basis. The distributions are usually unimodal and don't have especially long or fat tails. (By contrast, the betweenness centrality distributions are very skewed as $B(n)$ is quite large for a handful of goods and very close to zero for everything else.) Between the first few weeks and the last, median closeness increases seven fold while the shape seems to change little. The big jump in mid February 2012 seems to be due to mainly to allowing trade in previously untradeable items and improvements in the user interface; the number of trades and traders nearly doubled at this time. Overall, the Figure shows that shortest paths between goods became weightier and shorter over

73

time, i.e., it became easier and easier to trade one arbitrary good for another.

2.4.3   Price Dispersion

Did the emergence of money and specialist traders improve efficiency? More specifically, did transactions costs decline and prices become more unified? In a monetary economy, the most direct measure of transactions cost is the spread between bid and ask prices, the current cost of a round-trip transaction. That is, the direct measure is the net loss (as a percentage of the mean price) when you sell an item at the highest bid price and immediately repurchase it at the lowest ask price. Since bid and ask prices are not part of our data set (nor did they exist in TF2 during this time), we need an observable proxy for transactions costs.

As explained in the Appendix, we believe that SIQR, the **i**nter**q**uartile price **r**ange **s**caled by the median price, is a good proxy for transaction cost, as well as a robust direct measure of price dispersion. It also aggregates well across goods (see equation (4.17) and surrounding discussion), so we take its value-weighted average as in Figure 2.14.

The Figure shows that overall SIQR is quite high in the Summer of 2011, but by Winter 2011 it declines to under 50 percent, and is mostly in the 25-35 range in 2012. We infer that a typical (in terms of value) round trip trade would return only about 100-75 = 25 percent of its original value in early months, but would return 65-75% in 2012. The overall SIQR spikes briefly during promotion events, probably due to the influx of new traders and new goods of unclear value. In 2013 SIQR declines modestly
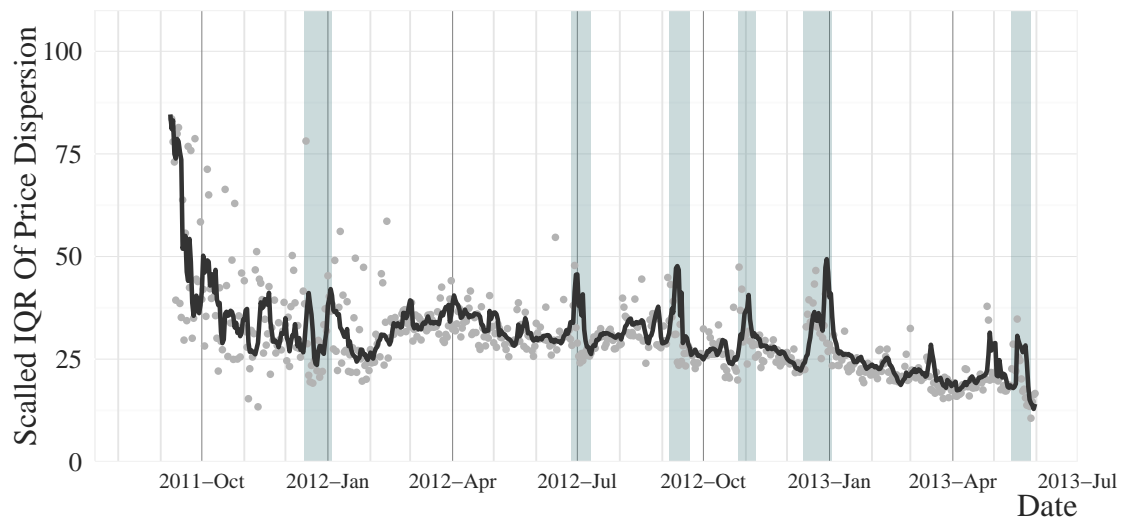
Figure 2.14: Overall price dispersion as a percentage of mean price

Gray dots show daily value-weighted averages of normalized interquartile range, and black lines show trailing value-weighted 7-day averages, over all items with at least 100 price observations in previous 30 days. Vertical shading indicates promotion events.

and mostly remains below 25, and in the last few days of our sample it falls to about 16, suggesting that round-trip costs are about one sixth of original value.

The emergence of the Clump in October 2011, and its sudden disappearance for ten days at the end of July 2012, had no discernible effect on SIQR. As noted earlier, the thickest bilateral market is keys-for-metal, and there SIQR drops below 10 percent in October 2011. From January 2012 until the end of our sample, the SIQR for that money conversion rate mostly bounces around in the 3-8 percent range, while that for earbuds is mostly around 5% and that for Bill's hat is mostly around 12%.

75

2.5   Discussion

What conclusions can we draw from our results? Although Valve engineers created a trading platform that was entirely egalitarian, we found that several sorts of specialists soon emerged to facilitate trade. The node strength distribution in the trader network grew a longer and fatter upper tail, and then calved off sets of very large traders with different specialities and different business plans.

One set of nodes evidently was controlled by a single entity (we call it the Clump) that traded actively in a wide range of commodities and maintained modest inventories. We see the Clump as a classic intermediary. It earned a modestly profitable spread between buy and sell prices and offered trade credit secured by deposits. An apparent competitor emerged later, a market maker in low- to mid-value goods via spot transactions. Another specialist (MoneyChanger, as we call it) made the market for key-metal exchanges, the weightiest edge in the entire TF2 economy. Evidence so far suggests a substantial role for brokers (or escrow agents) in the market for high-end goods, but little role for speculators. The dramatic drop in trader network assortativity in the first year suggests that, taken together, the specialist traders indeed facilitate trade, reducing small traders' search costs and frictions.

The results on goods networks are equally enlightening. Although Valve designers created perhaps the most efficient barter platform that the world has ever seen, the evidence indicates that nevertheless indirect, monetary exchange soon evolved. Betweenness metrics show that the composite money good was already quite prevalent by

the time our data sample begins, and became essentially universal in exchange within a few more months, consistent with classical economists' writings on money.

We found that the money composite in TF2 consisted of 6 distinct goods (or 4, if one lumps together the three different grades of metal). In our interpretation, the multiplicity of commodity monies is due to the indivisibility of TF2 goods. Indivisibility limits the competition between low- and high-denomination commodity moneys; it is awkward to trade dozens or hundreds of units of a low denomination money for a valuable good, or to make change when paying for a cheap good using a unit of a high denomination money. Thus a high denomination money good may be a complement rather than a substitute for low denomination money.

Indeed, although only one good, the dollar, is money in the US, it is also true that coins and bills are indivisible. Four popular denominations (quarters, 1-dollar bills, 5's, 20's) span two orders of magnitude in value. Likewise, in TF2 trading, proto-money goods may compete within each denomination range — Max's Severed Head, HOUWAR, earbuds and Bill's hats seem to have competed with each other in the high range with only the latter two surviving as media of exchange, but none of these items seemed to compete with metals in the low range.

The sustained increase in the closeness metric suggests easier trading as TF2's market institutions matured. More direct evidence comes from our measure of price dispersion, SIQR, which also serves as a proxy for transactions costs. The overall value of SIQR dropped sharply during the first year from around 75% to well under 50%, and by the end of our sample was below 25%. Even in the thickest markets, SIQR

remained substantially above the 0% level implied by a strict Law of One Price. Our interpretation is that indivisibilities remained important for low to medium value goods, and that markets remained thin for high-end goods.

Each economy, including the modern global economy that we all inhabit, has its own peculiarities, and one must be cautious in generalizing. Our paper contributes some new evidence on how economies can self-organize, a new data point to combine with those already available. This new data point may be especially useful because it comes with unprecedented detail on transactions, and is relatively independent of those already known by historians. We cannot conclude too much from the TF2 economy alone, but it does add new support (and new caveats) to classical perspectives on money, and to the view that institutions emerge spontaneously to reduce transactions costs and facilitate trade.

Much work remains. Towards the end of our sample, the Steam Trading platform supported considerable trade for virtual goods for games other than TF2, and some trades crossed the boundary between different games. We conjecture that these data may provide a new perspective on international finance questions, especially those concerning what happens when previously separate economies begin to interact with each other. Valve and its user community both continue to innovate, so the story continues.

Our main technical contributions are to propose new ways to construct two different networks from barter transactions data, and new ways to adapt existing network metrics in order to describe how these networks evolve. We hope that our readers are inspired to further refine and extend these metrics, and to build testable models of how

architectures change endogenously in large networks.

Part II

Part 2: Coordination in Lab Experiments

Chapter 3

Minimum effort coordination in continuous time - An Experimental Analysis with changing payoff structures

Written with coauthors Thomas Campbell and Maren Tonn

3.1   Introduction

The minimum effort game has been the subject of experimental inquiry following the archetypal example of Anarchia from Hirshleifer (1983). Anarchia is a circular island which is split into a set of equally sized wedge pieces with corresponding identical coastlines. Each wedge is owned by a different denizen. It is periodically threatened by floods, which require residents to build walls on their coastlines. If your wall is shorter than your neighbors, then the entire island is put at risk of flooding. If you build it

taller than your neighbors, then you have wasted effort. But not all examples of the minimum effort game are so contrived: Target stores famously had their payment network breached in 2013, leading to the loss of hundreds of millions of customers' personal information including credit card numbers. The hacker that was responsible broke in through a third party HVAC system which Target had given network access. As it turns out, network security is only as strong as the most vulnerable access point. Obviously, the HVAC vendor whom Target had granted network credentials did not put forth the same amount of effort in their network security that Target had.

Literature on minimum effort coordination games is extensive, starting with Van Huyck et al. (1990). The fact that it only takes one player to cause a collapse of Pareto optimal coordination makes this game especially interesting. This means that the risk dominant equilibrium is the opposite of the payoff dominant equilibrium, so there is no possibility for players to increase their own payoff at the cost of their neighbors. Since everybody's payoffs necessarily either rise together or fall together, one might think that it should be an environment conducive to coordination. But experimental work has demonstrated that coordination in a repeated version of the game is in fact generally quite poor. It is this puzzle which has motivated a wide body of examination, including allowing communication between participants and increasing the number of discrete periods. We hope to illuminate another situation which may solve the coordination problem which has motivated previous research. Our approach is to have participants interact with each other in a novel way.

One of our primary contributions is the application of a continuous time struc-

ture rather than repeated discrete periods used in most prior exploration of the minimum effort game. Continuous time interaction may be sufficient to solve the coordination difficulties experienced in much of the existing literature because it effectively turns every period into an infinite number of repetitions, and increasing repetitions has been shown to increase coordination rates. It also, when combined with high information about other players' decisions, effectively offers participants a non-explicit way to communicate through signaling. Continuous time is also a natural approach for experimentation in general as it is closer to the speed of adjustment in real world dynamic systems. To our knowledge, there are only two studies focusing on coordination games in continuous time (Deck and Nikiforakis, 2012; Leng et al., 2016). Both studies implement a minimum effort game, but vary substantially from our study. They vary available information and compare results to discrete time.

Continuous time in experimental research is still relatively unexplored. Previous experimental studies that implement decision making in continuous time included the Prisoner's Dilemma (Bigoni et al., 2015; Friedman and Oprea, 2012). Both studies show high cooperation rates.

The final inspiration for our work comes from the growing experimental research of Gradualism as in Ye et al. (2014) and Kamijo et al. (2016) as a tool to improve coordination. They examine the effect of slowly adjusting the environment in which experimental participants act and find that it tends to produce better group coordination. Due to this evidence, we believe that a combination of continuous time along with Gradualism approaches can provide a solution to the fundamental problem of coordination

83

failure in this environment.

We test whether the extension of the game into continuous time without Gradualism is sufficient to induce socially optimal coordination. We alter payoff function structures of the minimum effort game in continuous time in two ways. First, we compare two levels of a penalty parameter that determines the extent to which deviation from minimum play is punished. Second, we focus on gradual changes of the payoff function structure by changing the penalty parameter within period. Our motivation for using this sequence of payout structure transformations is to manifest a norm for coordination first in an easier version of the game before we change it to a different game.

Our results show significant differences with respect to different payoff structures in the first part. Continuous time alone is not sufficient to reach coordination when the punishment for deviating from minimum play is high. When implementing two different forms of payoff changes, we do not find high rates of group coordination on socially optimal equilibria as expected, but we do observe more successful coordination when compared to constant severe punishments lasting the full period.

The rest of the paper is organized as follows. First, we introduce relevant literature; next we present our design and hypothesis; finally we present our results for both parts, and discuss our findings

3.2   Literature

Van Huyck et al. (1990) are credited with one of the first studies of minimum effort games, where the minimum play in a group determines individual payoffs. They used large groups of 14 and 16 subjects. The only information made available to players over the 10 repetitions of the game was the previous group minimum. They experimented with adding a 'penalty parameter' to punish deviation from the minimum. They found convergence to the lowest Pareto-ranked equilibrium when deviation from the group minimum is punished. They also experimented with group size with a penalty for deviation. With pairs of subjects they found that nearly all of these pairs converged to the Pareto-dominant equilibrium, whereas larger groups exhibited more trouble coordinating on higher Pareto-ranked equilibria.

A variety of studies introduce mechanisms to improve efficiency in coordination games. Cachon and Camerer (1996) improve coordination in minimum and median effort games by charging a participation fee to enter into the game. In Weber et al. (2001), an appointed leader giving a speech before the decision does not lead to higher efficiency . Van Huyck et al. (1997) reports evidence that a good start is a good predictor of better group coordination as in for median effort games, indicating that there may be path dependency in outcomes.

Further minimum effort games show that repetition improves coordination and information about the decisions of other players increase minimum play (Berninghaus and Ehrhart, 1998, 2001). Both studies underline the potential of continuous time games

with high information availability as continuous time can be conceived as an infinite number of differential time periods.

The only studies on the minimum effort game in continuous time are Deck and Nikiforakis (2012) and Leng et al. (2016). Deck and Nikiforakis vary information about the other players in a real-time game that is repeated 10 times. Participants can change their play at any time during the period and everyone the lowest effort level. But only the decisions at the last moment count towards the payoff, so signalling was effectively costless. This experiment uses the payoff structure of Van Huyck et al. (1990) and generally find a high level of coordination. Leng et al. (2016) compares continuous and discrete time formulations of the minimum effort game and find coordination on higher equilibria in continuous time, but only when subjects receive information on all other subjects' decisions.

The final body of literature from which we draw inspiration is the work on "Gradualism", including Kamijo et al. (2016) and Ye et al. (2014). Kamijo et al. (2016) introduce a minimum effort coordination game that gradually changes and gets more difficult from period to period. They compare an exogenous form of Gradualism to forms with different mechanisms that are endogenously determined. They limit the game to a 2x2 matrix game in the first period and adjust the number of strategies available to participants. In the exogenous version one additional choice is added every period while in the endogenous treatments additional choices are added or removed depending on the level of coordination in the preceding round. They found that gradual treatments were associated with a higher degree of coordination when compared to controls.

In Ye et al. (2014), the difficulty and profitability in a coordination game with binary choices and groups of four players are varied using sharp and gradual changes. They use a discrete environment and a constant game structure. The found higher efficiency in the Gradualism treatments than when changes happened suddenly.

3.3   Experimental Design and Procedure

This experimental study analyzes decision making in a minimum effort coordination game that is implemented in continuous time. First, we introduce the basic game followed by the exact payoff structures that depends on our treatments.

The Game   Subjects are placed within a group of 4 players. Their payoffs are determined by their individual decision along with minimum effort of their group. The flow payoff π_i at a specific point of time of player i is determined as follows:

$$\pi_i(e_i, e_{-i}; \beta) = \hat{e} - \beta(e_i - \hat{e}) \tag{3.1}$$

where $e_i \in \{1, 2, ..., 10\}$ is the "effort" chosen by player i, $\hat{e} = \min_j e_j$ for $j \in \{1, 2, 3, 4\}$ is the minimum effort out of all four players in player i's group, and $\beta \geq 0$.

Because this is a continuous time environment, the payoff earned in each period is calculated by taking the time integral of the flow payoffs with the period duration

normalized to one. For example if a player has a flow payoff of 2 for half of a period and a flow payoff of 4 for the other half, the total payoff from flows for that period will be 3. Players start each period with an initial endowment of three points which is added on to their total payout resulting from the flows in each period.

We refer to β as a "penalty parameter". It represents how punishing it is for a player to select a position higher than the group minimum. A large β means that it is relatively more costly to deviate from the group minimum.

At a penalty parameter of zero (as examined in Van Huyck et al. (1990)), all players have a weakly dominant strategy to chose the maximum effort level. But for any positive penalty parameter, there is no longer a weakly dominant strategy and any set of strategies such that $e_i = e_j$ for all $i \neq j$ is a Nash Equilibrium. For all of our treatments we implement β, so that $0.1 \leq \beta \leq 2.5$. We select the lower bound of β to conform with previous experiments and choose a larger upper bound than is used in any literature to implement a particularly strenuous test of coordination. Next, we will introduce our treatments that vary β between and within treatments. Table 3.3 shows an overview of the treatments.

Constant Treatments. In a first step, we vary the penalty level β between periods in a within-subject design. In treatment MILD the penalty parameter is always equal to 0.1. In treatment SEVERE it is always equal to 2.5. These penalty parameters lead to possible payoffs displayed in Table 3.1 and 3.2. The MILD penalty parameter is in a similar range as standard coordination games have implemented Van Huyck et al.

Table 3.2: Payout with penalty parameter of .1 (top) and 2.5 (bottom)

		Group Minimum									
		10	9	8	7	6	5	4	3	2	1
	10	10	8.9	7.8	6.7	5.6	4.5	3.4	2.3	1.2	0.1
	9		9	7.9	6.8	5.7	4.6	3.5	2.4	1.3	0.2
	8			8	6.9	5.8	4.7	3.6	2.5	1.4	0.3
	7				7	5.9	4.8	3.7	2.6	1.5	0.4
	6					6	4.9	3.8	2.7	1.6	0.5
Player's decision	5						5	3.9	2.8	1.7	0.6
	4							4	2.9	1.8	0.7
	3								3	1.9	0.8
	2									2	0.9
	1										1

		Group Minimum									
		10	9	8	7	6	5	4	3	2	1
	10	10	6.5	3	-0.5	-4	-7.5	-11	-14.5	-18	-21.5
	9		9	5.5	2	-1.5	-5	-8.5	-12	-15.5	-19
	8			8	4.5	1	-2.5	-6	-9.5	-13	-16.5
	7				7	3.5	0	-3.5	-7	-10.5	-14
	6					6	2.5	-1	-4.5	-8	-11.5
Player's decision	5						5	1.5	-2	-5.5	-9
	4							4	0.5	-3	-6.5
	3								3	-0.5	-4
	2									2	-1.5
	1										1

(1990). The SEVERE penalty parameter shows a different payoff structure with losses being possible.

Changing Treatments. In a second step, we implement two additional treatments. We vary how the penalty parameter is changing from MILD to SEVERE within each period. We introduce the treatments JUMP and GRADUAL. In JUMP, the penalty parameter discretely changes from 0.1 to 2.5 at the 15 second mark. In GRADUAL, the penalty parameter adjusts from 0.1 to 2.5 continuously over a 30 second interval, starting at the 15 second mark. Table 3.3 presents a summary of all treatments.

Each session consists of 12 periods, 6 periods of each of the two treatments. In

Table 3.3: Treatment Overview

	Treatment Name	Description
1	MILD	Penalty Parameter is equal to 0.1 for the duration of the period
2	SEVERE	Penalty Parameter is equal to 2.5 for the duration of the period
3	JUMP	For the first 15 seconds of the period, the penalty parameter is equal to 0.1, then is set to 2.5 for the rest of the period
4	GRADUAL	Penalty parameter is 0.1 in first 15 seconds, then over the next 30 seconds increases linearly to 2.5 where it remains for last 15 seconds.

one set of sessions MILD and SEVERE are implemented. In the other set of sessions GRADUAL and JUMP are varied (compare Table 3.4). Prior to the first period, participants are randomly assigned to one of two matching groups of size 8. In each period participants are rematched into subgroups of 4. Each period subjects are playing with a different group of players and we receive two independent observations per session. This results in eight independent observations per treatments.

Table 3.4: Sessions

Session	Treatment	No. of Subjects
1-4	MILD & SEVERE	16 per session (8 independent observations)
5-8	JUMP & GRADUAL	16 per session (8 independent observations)

At the beginning of each period, each player is randomly assigned a starting position from (2,4,6,8) and no position is repeated in a single group. This differs from many other studies such as Leng et al. (2016) which allowed participants to select their starting position and Deck and Nikiforakis (2012) which started participants at the minimum. From Van Huyck et al. (1997), we know that there is evidence of path-dependency (the higher the initialization position of subjects, the more likely the Pareto-dominant outcome is chosen) in the outcome of the group, so the starting value is an important design consideration.

90

Movement from one position to another is not instantaneous and occurs at a fixed rate of two units per seconds. This process has been termed a "speed limit" in Kephart and Rose (2016). Subjects select their position target asynchronously and the software updates every 0.17 seconds. This speed limit is analogous to a time cost of adjustment; using the network security example, it takes time and resources to adjust the amount of security protecting one's network, whether they want to change it to a higher level or a lower level. Both of these design decisions are also intended to emphasize the continuous time action space and distinguish it from a discrete choice game.

Procedure. We conduct 8 sessions with 16 participants each for a total of 128 participants. The experiment is conducted at the LEEPS Lab at UC Santa Cruz in February 2016. Subjects are students of UC Santa Cruz and recruited via ORSEE (Greiner, 2015). The experiment is implemented with Redwood 2 (a successor of ConG, Pettit et al. (2014)). Subjects receive a show-up fee of seven USD. A second experiment was conducted immediately following this coordination game and the show-up fee is paid for participating in both. On top of that, participants receive $0.70 times their mean points earned per period over the session. Therefore, all periods contribute to the final earnings of each participant. Average payments excluding the show-up fee were $4.32 in constant treatments and $4.47 in sessions with changing penalty parameter. Paper instructions were distributed and all subjects had the opportunity to read the instructions at their own pace. Afterwards, subjects were given a series of questions on the screen designed to ensure that everybody understood how payoffs were determined. There were no practice

periods. On average the experiment took 30 minutes (including instructions and control questions).

3.4 Theoretical Predictions and Hypotheses

This experiment examines a minimum effort game that is implemented in continuous time. Within a group of four participants, payoffs are determined by the individual decision and minimum play of the group. During a period players can change their decision at any time and have full information about the current effort of the other three players that they are matched with for that period. Chosen effort e can be any integer between 1 and 10.

Therefore, there are ten static Nash equilibria in this game which occur at any combination of $e_{i \in N}$ for which $e_i = e_j$ for all i and all j. The Nash equilibria in this game are Pareto ranked and higher equilibrium group minima generate higher social welfare.

To evaluate the results we use the minimum group play M and mean absolute deviation (2). Mean absolute deviation is the average deviation from the minimum across a particular group and thus equals 0 when a group is playing at any equilibrium.

$$\text{Mean Absolute Deviation} = \frac{\sum_{n=1}^{4} (\text{deviation from minimum})_n}{4} \qquad (3.2)$$

We study averages over the whole period, the last 15 seconds and the endpoints of a period. In addition, we examine the development over the time within periods.

In the first set of sessions, we only vary the penalty parameter between periods so that it is constant within each period. In the second set of sessions, the penalty parameter changes within the period. The first question this study seeks to answer is whether continuous time and high information availability is sufficient to induce coordination to near Pareto optimal levels. Previous studies (Van Huyck et al., 1990) have shown that coordination on Pareto-superior equilibria is more difficult to achieve when a penalty is imposed for deviations away from the group minimum. Our treatments include information about the other players position, this can facilitate coordination (Berninghaus and Ehrhart, 2001; Knez and Camerer, 1994).

Berninghaus and Ehrhart (1998) show that the degree of coordination to superior outcomes is sensitive to the number of iterations of a discrete time coordination game, even if overall session duration is fixed. When subjects play 90 rather than 10 periods, they tend to coordinate on better outcomes. Continuous time action effectively has our subjects act in an infinite number of differential time periods, since at each moment they have the opportunity to choose an action. In this light, our subjects effectively have much more than 90 periods in which to act.

Hypothesis 1: Our first hypothesis is that players coordinate on equilibria with higher outcomes when the penalty parameter is MILD than in when it is SEVERE.

When the penalty parameter is relatively smaller, the penalty for off-equilibrium play is relatively lower as well. Attempts to persuade other participants in one's group to increase their effort by playing above the group minimum are less costly. We hypothesize that participants will be more willing undertake these sorts of forward looking strategies when it is relatively less costly to do so, in which case we anticipate coordination to higher average group minima in MILD compared to SEVERE treatments.

Knez and Camerer (1994); Van Huyck et al. (1997) report evidence of path-dependency in equilibrium selection. If people initialize with a high group minimum, that high degree of coordination tends to carry through the full session. We hypothesize that starting a period with a mild penalty level will induce coordination to "good" equilibria and that this coordination will carry through to the end of the period even if the penalty parameter is increased later in the period. In this case, we would expect the average group minimum to be higher for the final 15 seconds of periods in the GRADUAL and JUMP treatments than it would be in the case that the penalty was high from the beginning.

Hypothesis 2: Our second hypothesis states that when shifting the penalty parameter from MILD to SEVERE (in treatments JUMP and GRADUAL), during the last 15 seconds, where the penalty parameter is severe in all three treatments, coordination will be higher in JUMP and GRADUAL than in the constant SEVERE treatment.

Kamijo et al. (2016); Ye et al. (2014) both found that slow and gradual adjustment led to higher levels of coordination in their discrete time games when compared to no adjustment or fast adjustment. This supports our Hypothesis 2. In a next step we focus on the difference between GRADUAL and JUMP. Our GRADUAL treatment involves the cost of deviation gradually increasing over time and our JUMP treatment involves it instantaneously increasing to its maximum value. In line with the results of Ye et al. (2014), we hypothesize that the gradual increase will allow participants to more comfortably coordinate to higher minimum levels of effort and thus we expect that the minimum group play at the end of the period in GRADUAL periods will be higher on average than in JUMP periods.

Hypothesis 3: Our third hypothesis states that observed group minima are higher in GRADUAL compared to JUMP during the final 15 seconds of the period, when the penalty parameter is the same in both treatments.

3.5 Results

3.5.1 Constant Penalty Parameter

In the first part of the analysis, we focus on the minimum effort in treatments MILD and SEVERE. Subjects need time for orientation and there were no practice periods. This leads to learning over time. Figure 4.5 (in the Appendix) underlines the differences between earlier and later periods. Therefore, the analysis of our results

focuses on periods 5 to 12. This way we do not take into account the results from two periods per treatment in the analysis. For completeness, results for periods 1 to 4 are displayed in the Appendix.

Table 3.5: Average minimum effort in constant treatments

Treatment	Full period minimum	Last 15 seconds minimum	Endpoint minimum
SEVERE	2.90	2.89	2.79
MILD	6.23	7.40	7.51

Table 3.5 shows the results of the constant treatments. Over the whole period the average minimum play is 2.90 in SEVERE and 6.23 in MILD. When only looking at the period end points this difference is even bigger with 2.79 in SEVERE and 7.51 in MILD. During the last 15 seconds of the period, average minimum play is equal to 2.89 and 7.40, respectively. This difference is statistically significant (Wilcoxon signed-rank, p: 0.008, we use matching groups, as independent observations, 8 per treatment). While we do find coordination on high group minima in MILD, continuous time and full information was not sufficient to lead to high levels of coordination when introducing a high penalty parameter in SEVERE.

We can look at the development of minimum play during a period. Figure 3.1 shows the development of the average group minimum during the 60 seconds. At the beginning of the period the minimum is always equal to 2 because one of the players in each group randomly starts at the point. In MILD the minimum sharply increases during the first 20 seconds. After this, we observe only a small increase. In SEVERE the minimum only increases for a few seconds and does not increase any further.

96
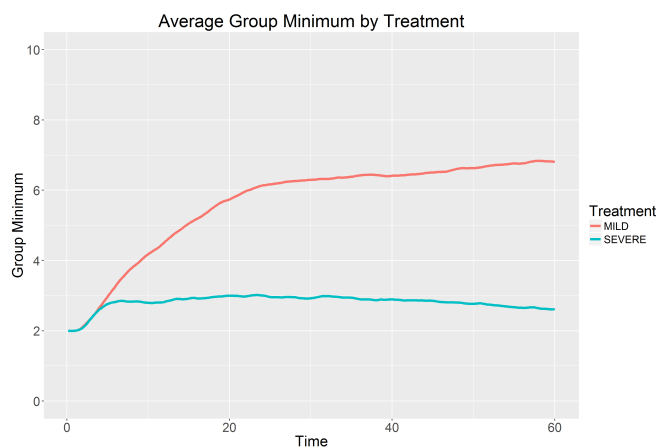
Figure 3.1: Average Group Minimum Play by Treatment

These results are also underlined by Figure 3.2, which shows the distribution of group minima at the final moment of each period. While a majority of groups coordinate on a minimum of 10 (the Pareto optimal) in MILD, an even larger share of groups coordinate on 1 (the Pareto pessimal) in SEVERE.
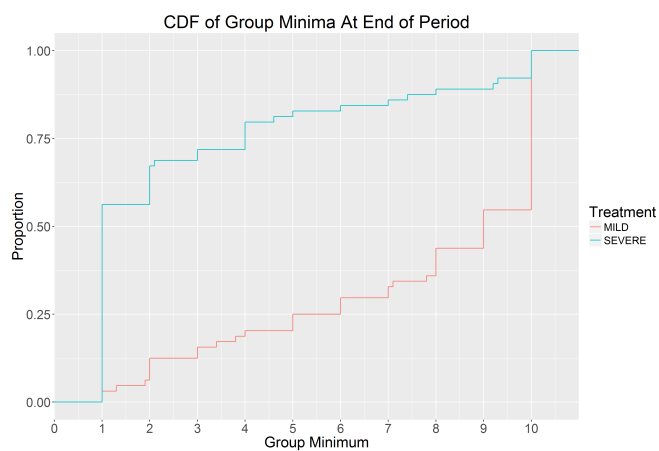
Figure 3.2: Distribution of group minimum play at period end

Result 1: Subjects coordinate on higher minimum play in MILD (in comparison to SEVERE). This holds for the average plays across whole period, the average plays

during the final 15 seconds of each period and the average end points of each period. The difference between the treatments increases towards the end of the period.
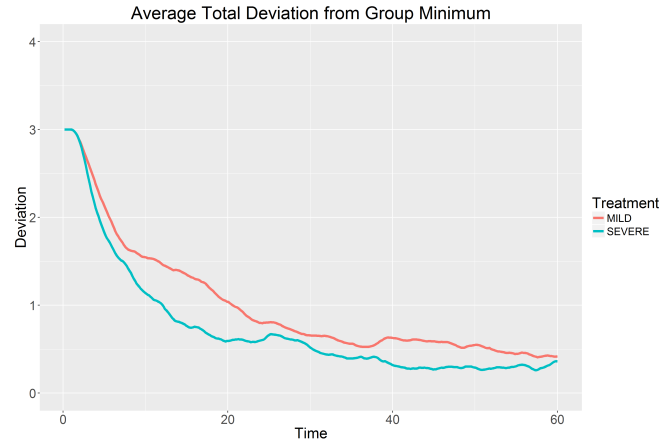
Figure 3.3: Total deviation, constant treatments

In a next step, we focus on equilibrium play. As long as all players in a group decide on the same number they are playing a Nash Equlibrium. Figure 3.3 shows mean absolute deviation. This deviation decreases over time within each period in both treatments. At the beginning the average deviation is always 3 by construction, since the four starting positions are fixed across all periods. At the end of a period groups coordinate well with an average deviation below 1 and mainly agree on one number. The mean absolute deviation curve is higher in the MILD penalty treatment than in SEVERE. This is consistent with people being induced to play over the minimum at a higher rate when the costs of doing so are relatively low. But average deviation in both cases is around 0.5 by the time 30 seconds have elapsed, indicating that equilibriation is mostly complete by this time. We did not find any significant correlation between the starting position and the future plays on an individual level (comparing starting position

98

with ending position), so we believe that our randomization was effective.

Overall, we can show that coordination can arise in a minimum effort game in continuous time, but a SEVERE penalty parameter leads to coordination on an inefficient outcome. Therefore, we observe similar results as literature focusing on discrete time suggest. In the following part, we will analyze whether this coordination failure can be overcome with changing the penalty parameter during the period, starting with MILD.

3.5.2 Changing Penalty Parameter

In our second set of sessions, we vary how the penalty parameter β changes during the period to see how coordination is affected and whether we can reach coordination on a higher equilibrium than when participants were subjected to a SEVERE penalty parameter for the full period. We vary whether the parameter jumps from MILD to SEVERE discretely (JUMP) or slowly over time (GRADUAL). This change occurs after 15 seconds of play every period and treatment order is varied randomly within-subjects. The summary statistics are presented in Table 3.6. During the last 15 seconds the average minimum play is equal to 3.75 in JUMP and 4.32 in GRADUAL. At this time the change of the penalty parameter is over in both treatments. At the endpoint minimum play values are slightly lower with 3.56 (JUMP) and 3.96 (GRADUAL).

Table 3.6: Summary Statistics: Gradually Changing Treatments

Treatment	Full minimum	Last 15sec minimum	Endpoint minimum
JUMP	4.27	3.75	3.56
GRADUAL	4.78	4.32	3.96

99
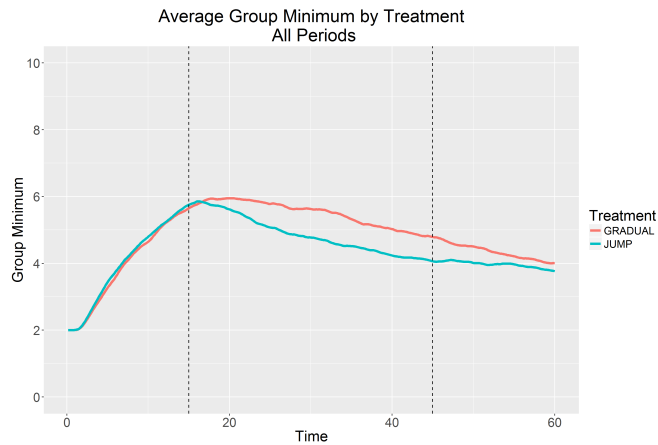
Figure 3.4: Average Group Minimum Play by Treatment

Figure 3.4 shows the development of minimum play during the 60 seconds period. Prior to second 15 (until first vertical line), minimum play is increasing similar to the MILD constant treatment (this similarity is displayed explicitly in the appendix). At second 15, the penalty parameter changes to 2.5 in JUMP treatment periods and slowly adjusts over a period of 30 seconds in GRADUAL. The path of observed group minima correspondingly changes direction when the penalty parameter shifts and starts decreasing, albeit more slowly in GRADUAL compared to JUMP.

In general, we did not find high coordination as expected. In the previous MILD constant treatment an end of period average of 7.40 was observed. This is significantly higher than both treatments with changing penalty parameter (Mann-Whitney U test, p: 0.02 for both comparisons, MILD vs. JUMP and MILD vs. GRADUAL, we aggregate over matching groups, 8 per treatment).

Figure 3.5 shows the distribution of group minima at the end point of each period for all treatments. We find that the cumulative distribution function (CDF)

100
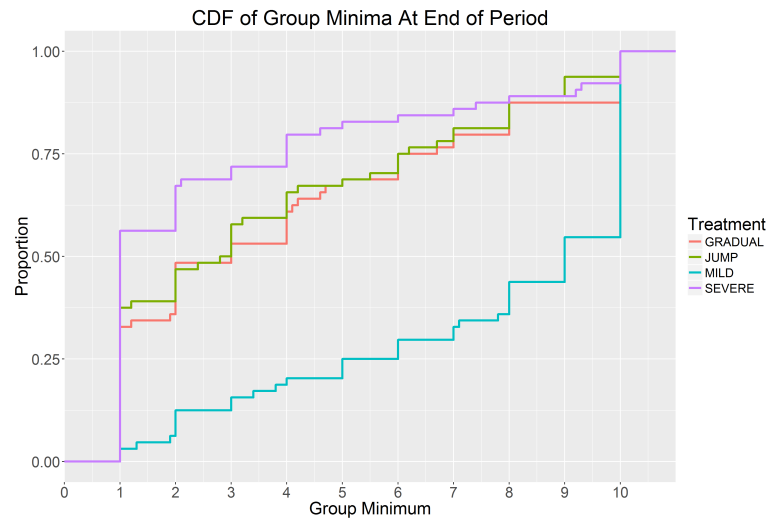
Figure 3.5: Cumulative distribution function of Group Minima at the end of each period

representing JUMP and GRADUAL treatments both are significantly below that of SEVERE (Kolmogorov-Smirnov (K-S), p: 0.047 (GRADUAL) and p: 0.030 (JUMP)) and above that of MILD (K-S, p<.001 for both). In line with Figure 3.5, we use the less conservative version and use individual group minima. The results indicate that the observed distribution of group minima tends to have higher density at higher numbers in both of the changing payoff treatments. Therefore, there is an overall beneficial effect of both, GRADUAL and JUMP, when compared to the constant SEVERE treatment. Still subjects were not able to generate positive outcomes as found in the constant MILD treatment. A Mann-Whitney test with matching groups as independent observations only supports the significant improvement of GRADUAL when comparing to SEVERE (p: 0.065), but fails to find significant improvements when comparing JUMP and SEVERE (p: 0.235).

Result 2: Subjects coordinate on significantly lower minimum play in JUMP

and GRADUAL than in MILD. But subjects do coordinate on higher minimum play in GRADUAL and JUMP than in SEVERE.

Mean absolute deviation is decreasing in both treatments similar to the constant penalty parameter treatments SEVERE and MILD. Figure 3.6 shows this development over time. About halfway through the average period, participants have an average deviation of less than 0.5 in both treatments and it stays at approximately that level until the end of the period, compared to 0.7 absolute deviation in the SEVERE treatment, and 0.9 in the MILD treatment with constant penalty parameters.
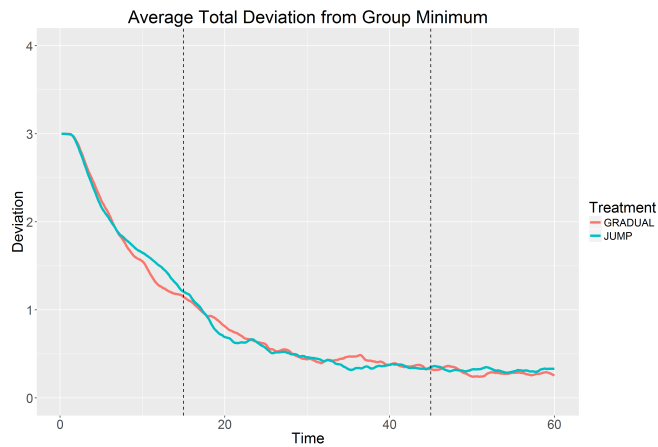
Figure 3.6: Total deviation

When focusing on the differences between GRADUAL and JUMP, we can see that with respect to deviation the graphs are nearly identical. Overall, we cannot find meaningful differences between the two treatments GRADUAL and JUMP. Only when testing using play through the full 60 seconds, do we find a significant difference between GRADUAL and JUMP (Wilcoxon signed-rank p: 0.023). But, the penalty parameter is lower in JUMP than in GRADUAL for the middle 30 seconds of the period, so this result

is not much stronger than what we found from comparing SEVERE and MILD. When we test the group minima from the last 15 seconds and the penalty levels are identical in both treatments, the difference is not statistically significant (Wilcoxon signed-rank, p: 0.195).

Result 3: Over the whole period minimum play is significantly higher in GRADUAL than in JUMP, but this difference diminishes at the end of the period.

Contrary to Hypothesis 3 and the previous experimental findings on Gradualism, we find that there was no discernable difference in outcomes when there was a slow gradual adjustment upward of the penalty level compared to a sudden shift.

Figure 3.7: Distribution of group minimum play at period end - GRADUAL and JUMP

Figure 3.7 shows the distribution of group minima at the final moment of the period in the changing parameter sessions. This points in the direction that GRADUAL leades to higher group minima than JUMP but the difference is not significant at the $p = .05$ level in the conservative K-S test or paired Wilcoxon Rank Sum test.
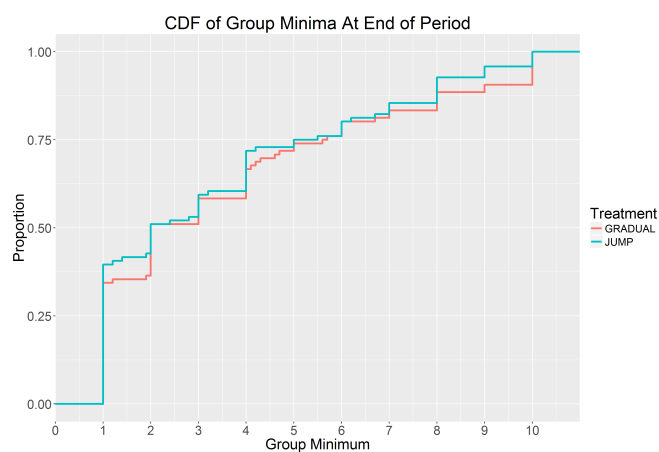
103

3.6 Discussion and Conclusion

In this study, we have implemented continuous time in a minimum effort game. Our results show that when there is a relatively mild penalty for deviation from the group minimum, groups tend to coordinate to a highly efficient outcome. In fact, they coordinate to a group minimum of at least 8 in over half of the periods. But when deviation is punished severely, continuous time along with full information is not enough to induce high levels of coordination and a large majority of the time group minima are 3 or lower.

We introduced two mechanisms to overcome these problems. In GRADUAL and JUMP, we change the payoff structure from MILD to SEVERE within the period. This form of Gradualism has not been implemented in continuous time games before. We find that the GRADUAL treatment results in significantly higher group minima than JUMP when looking at full periods, but that there is no significant difference with respect to outcomes during the last 15 seconds, when GRADUAL and JUMP have the same value of the penalty parameter. We conclude that slow adjustment of the penalty parameter did not induce a higher level of cooperation in our continuous time environment relative to a sudden jump, in contrast to previous work which have documented beneficial effects of slow adjustment in discrete time coordination experiments.

We did, however, find a beneficial effect of both of the Gradualism treatments GRADUAL and JUMP relative to a constant SEVERE penalty when comparing the distributions of end-of-period group minima for each treatment. We also find further

support that path dependency is an important component of minimum effort experiments when played in continuous time as players being given a period of relatively low penalty that induces high levels of coordination are more likely to carry their higher coordination through for the full period even when the penalty level becomes relatively high.

Future research might include the impact of different starting values and speed of movement. In addition from our results it can be assumed that it is necessary to establish more stable coordination before changing the payoff structure. Therefore longer periods or a longer time before introducing a change might be interesting. Nevertheless, this paper has shown new possible solutions to coordination failure that arises when implementing continuous time. Further possibilities would be the implementation of explicit communication, as it is usually only one player that inhibits coordination. The other players might be able to convince this player of changing their decisions if explicit communication was allowed.

Part III

Appendices and References

Chapter 4

Appendices

4.1 Appendix A

Chapter 1: Pricing in a Barter Economy

4.1.1 Price Weighting Methodology

We use a number of different approaches in generating weights to assign to individual observations in estimating daily prices. Broadly, these approaches fit into two categories: "Centered" and "Trailing" (or Leading).

4.1.2 Centered Prices Weighted Mean

To calculate an item's mean price for a specific day, we start with an interval of seven days. We collect all SM price observations from three days previous to three days into the future and remove any price observations above the 9th decile and below the 1st decile. We drop these extremes because almost all items have many price observations which are clear outliers and means are sensitive to such outliers. We then apply a triangular (or, more precisely a trapezoidal) weighting function as illustrated in Figure 4.1

There are initially three days on either side of the day which we are estimating prices for. Many items are very high volume and thus we have lots of price observations but for some items, there is relatively low enough volume such that even including a full week does not give us a large enough number of observations that we are confident in their prices.
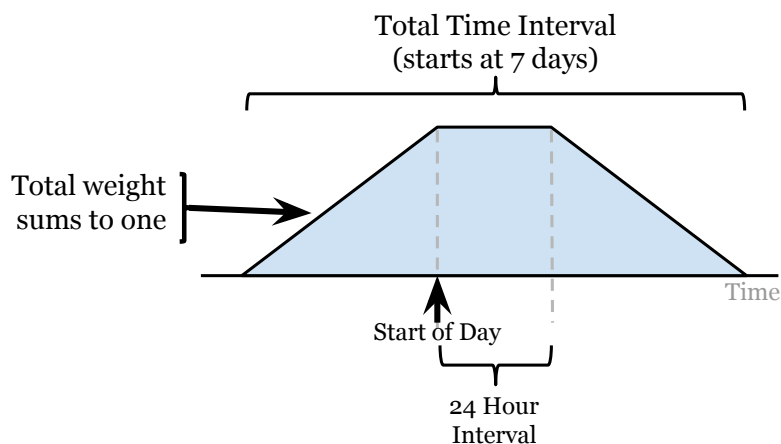

Figure 4.1: Weighting function

To account for this issue, we define a control system which utilizes the coefficient of variation: $c_v = \frac{\sigma}{\mu}$, where μ and σ are the mean and standard deviation of our sample. Our control system sets a cutoff value for coefficient of variation c_v^* and we calculate the coefficient for each item in a given time period c_v^{it} and if i is true that $c_v^{it} > c_v^*$, we increase the window for that item on that day by one day and recalculate. This process is repeated until the window includes sufficient observations such that $c_v^{it} \leq c_v^*$. The cutoff we use for this process is $c_v^* = .5$ a this number appears to consistently select an appropriate window width.

4.1.3 Representative Basket Derivation Methodology

In consumer inflation indexes like CPI these quantities strive to reflect typical consumption baskets. In contrast, quantities reflect producer purchases in input producer price indexes and in the Gross Domestic Product deflater they reflect production quantities. Our quantity index reflects the bundle of goods held by a "representative player."

4.1.4 Methodology

These representative player inventories were generated by drawing random samples of users from the active player population, where an active player is defined as one who logged into Team Fortress 2 within ninety days of the sample date. We identify the average quantity of each TF2 item held in the sampled inventories. But there are some unique issues with our sampling in this environment due to the presence of an upper

tail of inventory value distributions composed of people with very large inventory values. These HNWIs are rare enough that we almost certainly will not have a good balance of them represented in each time period's active player sample. Increasing our sample size sufficiently beyond 1% of the population is also technically infeasible given the number of active players (typically more than 250,000 each week) many of whom possess scores of items. Without adjustment, the price index could exhibit big movements from one period to the next due more to sudden shifts in the quantity index than shifts in price.

Our approach to dealing with these HNWIs is first to tag the top proportion of wealth-holding individual users as HNWIs, where we define the inventory value cutoff as a nominal inventory value above 800 keys, or approximately $1600. If an active player is classified as a HNWI in one of these censuses, their inventories are logged each week for the entire year and they are excluded from the non-HNWI sample for that year. These HNWI players account for approximately 0.3 to 0.4 percent of the active player population.

We then track inventories of all HNWIs each period along with the random 1% sample of non-HNWIs, and derive average item inventories for each group. The composition of the basket derived from these 1% samples does not fluctuate greatly from time period to time period. Finally, the HNWI and non-HNWI representative inventories are combined weighting item quantities based on each groups' relative proportion of the overall active player population at each period.

All inventory data excludes individuals who have marked their "Steam Profile" as private. Of the approximately 1,500 unique active players classified as HNWIs, 255

110

have been excluded due to this privacy restriction on their their backpacks. Our method-
ology thus assumes the omission of these privacy preferring players does not significantly
bias the representativeness of our HNWI and non-HNWI sample.

Once representative baskets are found for each tier, they are average together
weighted by the relative proportion of each group to the overall population.

4.1.5   Hedonic Estimates of Values of Item Characteristics

Equation 4.1 presents the hedonic model we estimate. We use this simplified
version because the model with time dummies has thousands of regression coefficients,
far too many to report in a single table. The full model from Equation 1.3, however,
was used to produce Figures 1.7 and 1.8:

$$ln\left(p_{it}\right) = \alpha_t + \sum_{k=1}^{K}(\beta_k \cdot x_{it}) + \varepsilon_{it} \qquad \text{for } t = 0, ..., T \qquad (4.1)$$

For item i in period t, price p_{it} is a function of weekly time dummies, K
time-invariant item characteristics, and an error process. Table 4.1 shows the coefficient
estimates of the hedonic regression.

All TF2 items are associated with a single "quality". We used the unique
quality for our regression as it is by far the most common as the baseline, and estimates
for each item are premiums or discounts relative to that item's unique version. These
results suggest that vintage items have tended to trade a full 180% above more normal
unique quality ones. All unique quality items that existed on or before September 20th

Table 4.1: Hedonic Price Model with Time Unvarying Characteristic Dummies

	Dependent variable:	
	log price	
Quality (mutually exclusive), Relative to Unique Quality		
Genuine	0.9035***	(0.0054)
Haunted	−0.9255***	(0.0077)
Other	5.5827***	(0.0621)
Strange	1.9038***	(0.0051)
Unusual	3.7819***	(0.0042)
Vintage	1.0300***	(0.0045)
Item Type (mutually exclusive) Relative to Action Items		
Cosmetic	0.4629***	(0.0098)
Tool	−0.0655***	(0.0111)
Weapon	−0.5792***	(0.0105)
Character Class Equippablility (non-exclusive)		
Spy Equippable	0.2507***	(0.0041)
Engineer Equippable	−0.0392***	(0.0042)
Soldier Equippable	0.1775***	(0.0037)
Sniper Equippable	0.0674***	(0.0043)
Demoman Equippable	−0.0869***	(0.0039)
Medic Equippable	0.0983***	(0.0043)
Pyro Equippable	0.0663***	(0.0038)
Heavy Equippable	−0.0318***	(0.0039)
Scout Equippable	0.2527***	(0.0039)
Widely Held Item ($>3\%$ of Active Players)	−1.5229***	(0.0033)

With Week Time Dummies	X
Observations	734,066
R^2	0.7449
Adjusted R^2	0.7449
Residual Std. Error	1.093(df = 733853)
F Statistic	1.88e+04*** (df = 114; 733853)

Note: Standard errors in parentheses, and   *p<0.1; **p<0.05; ***p<0.01

2010, when TF2 trading was introduced, were redesignated as vintage. Unusual quality items tend to attract the highest premium, a full 43000% premium above uniques. All unusual quality items possess some kind of visual effect, like flames, orbiting planets, or stinky-smelly lines. Unusual items are particularly rare, as they only appear with a very small probability from opening a crate and cannot come from any other source, and it is this rarity which is likely the reason which they command prices much higher than those of non-unusual items. Strange items, which will track in-game statistics, tend to exhibit a 571% premium above uniques. Quality "other" appears to attract the highest premium, however, items of this quality only appeared due to extremely unusual circumstances, akin to very rare coins minted with imperfections which make them very valuable to dedicated coin collectors but unavailable and inconsequential to everyone else, and accounting for only a negligible fraction of all coins. And, like coins, it is likely that some owners of the oddities are not even aware of the item's value. Thus we tend to see an extremely small number of transactions involving items of quality "other", but those transactions indicate that they are worth a small fortune. These, however, are not very representative of the broader economy.

All tools, weapons, and cosmetic items may be used by only one, some, or all character classes. Unlike quality or item type, this is non-exclusive designation. Character classes vary in speed, strength, and other abilities. For example medics are able heal teammates, and heavies are slow but may inflict and withstand a lot of damage. The highest equipability premiums come from items that can be carried by scouts and spies. Should an item be equipable by a scout, for example, it will tend to garner a 29%
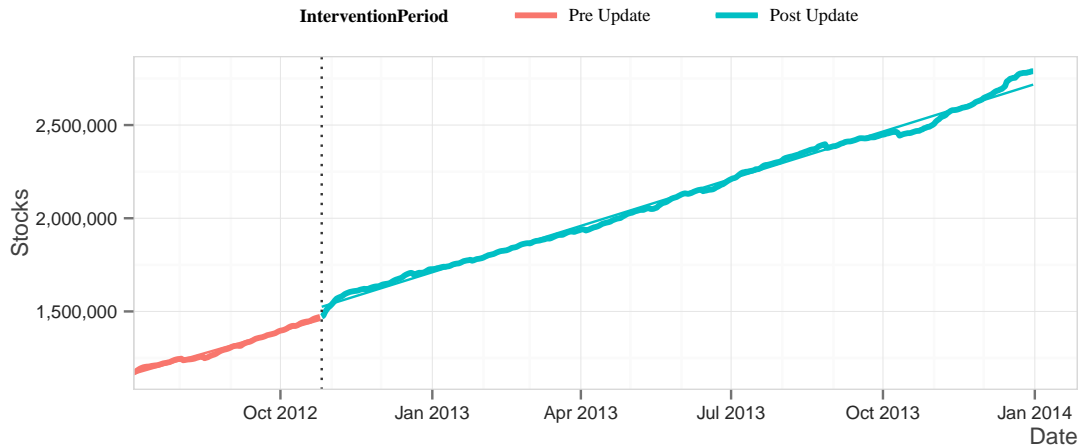
113

Figure 4.2: Regression discontinuity on Refined Metal stocks

premium relative to a similar item that is not equipable by the scout.

Items that are held by a relatively large percentage of active players, defined as items that appear in more than 3% of active player inventories, tend to exhibit a large discount, trading for 78% lower than items that are not so widely held. To ensure this characteristic is time unvarying, the percentage holding statistic is taken on average over the whole sample. This is included in the regression to account for the potential price implication of an item's relative scarcity.

4.1.6 Decline of the Aggregate Price Level

Figure 4.2 shows the outstanding item stocks of the primary currency metal, refined metals. Notice that right at the end of October, 2012, there is a sudden increase in the rate at which refined metals are entering into the economy. We argue that this is due to a major content update released on 26 October 2012. This date marks the
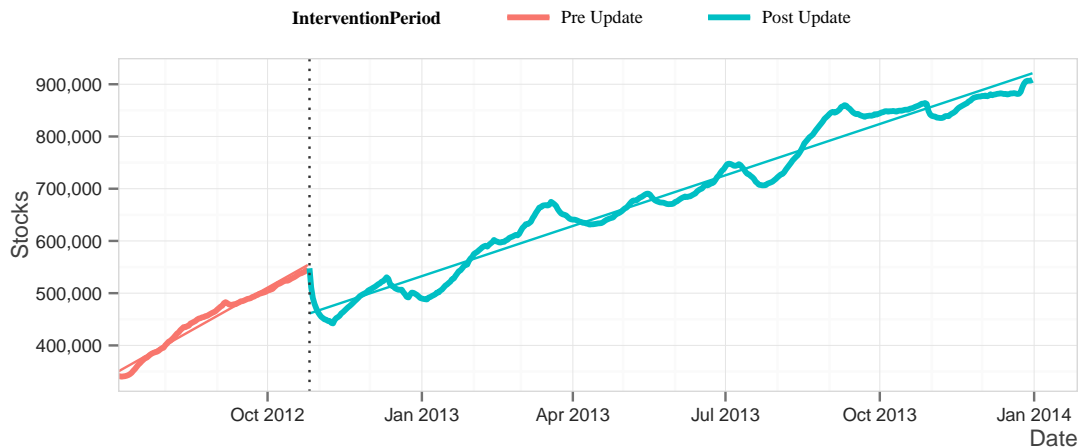
Figure 4.3: Regression discontinuity on Key stocks

point in time at which metals started their precipitous drop in value. This apparent positive supply shock appears to have been temporary, as the rate that metal stocks increase swiftly returns to a rate of increase only marginally higher than its rate prior to the event. But the fact that the rate of change of the stock of metals slowed did not correspond to a slowing rate of depreciation.

Figure 4.3 sheds some light on this question. Simultaneous with the sudden burst of metals that entered the economy, there was a sudden negative supply shock to our numeraire, keys. This corresponds to the major content update including new types of crates, leading to many more keys being consumed than purchased for a short period. The supply of metals suddenly increased and the supply of keys suddenly decreased; naturally, we would expect this to lead to an increase in the number of metals required to receive a key in exchange. The rate of increase of keys also appears to slow somewhat. This trend is confirmed by a regression discontinuity analysis using the following

115

regression equation:

$$Stock_i = \beta_0 + \beta_1 DaysToPatch_i + \beta_2 PostPatch_i + \beta_3 Interaction_i + \varepsilon_i$$

Where $DaysToPatch$ is the number of days until the update went live, $PostPatch$ is a dummy variable that takes a value of zero on days which were prior to October 26 and a value of one after, and $Interaction$ is the product of these two variables. Our regression estimates are presented in Table 4.2. β_1 can be interpreted as the pre-update rate of expansion of the money supply for each currency and $\beta_1 + \beta_3$ is the post-update rate of expansion of that currency, thus β_3 is the difference in the trends before and after the cutoff.

Table 4.2: Before and After Halloween Time Trends for Keys (1) and Metals (2)

	Dependent variable:	
	Item Stocks	
	(1)	(2)
DaysToPatch	1,810***	2,560***
	(59)	(47)
PostPatch	−90,924***	78,012***
	(4,552)	(3,649)
Interaction	−745***	129***
	(60)	(48)
Constant	572,952***	1,457,520***
	(3,976)	(3,187)
Observations	500	500
R^2	0.977	0.998
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 4.2 shows that the daily rate of increase in the stock of keys dropped

116

from 1,810 to approximately 1,065, a 41% reduction. Simultaneously, the rate of change of metals increases slightly from 2,560 to 2,689, an 5% increase. This could explain the sudden and continuous depreciation; before October 26, keys were entering the economy at a rate that was not too far off of the rate that refined metals were entering, but after this event the rates at which the goods entered the economy differed drastically, leading to relative supplies drifting further and further away from each other.

The depreciation of metals can clearly be seen when comparing the nominal and real values of our representative basket. These are presented in Figures **??** and **??**, respectively. The real value increases at an approximately constant rate as the average holdings of a representative agent increases over time, but the nominal value levels off and starts to decline right as the metal depreciation starts. This is due to the fact that the most commonly held items are generally metal-denominated and these items see shrinking key-values starting in October of 2012.

4.2 Appendix B

Chapter 2: Network Details and Supplementary Analyses

A directed weighted network is described by an $I \times I$ adjacency matrix $Y = ((y_{ij}))$. The entry $y_{ij} \geq 0$ is called the weight of the directed edge (or link) from node i to node j. By convention, all diagonal elements $y_{ii} = 0$, i.e., nodes do not connect to themselves.

Simplifying networks. Given network $Y = ((y_{ij}))$, replace y_{ij} by $y_{ij}^s = \max\{y_{ij}, y_{ji}\}$ for all $i, j = 1, ..., I$. The resulting adjacency matrix Y^s is symmetric, and so the network it defines is undirected. Likewise, replace every $y_{ij} > 0$ (or alternatively, every entry exceeding some positive threshold value) by $y_{ij}^b = 1$ to obtain the directed unweighted network $Y^b = ((y_{ij}^b))$.

Happily, the operation $Y \mapsto Y^s$ commutes with the operation $Y \mapsto Y^b$. That is, we arrive at the same undirected unweighted network Y^{sb} whether we first symmetrize using $y_{ij}^s = \max\{y_{ij}, y_{ji}\}$ and then binarize, or first binarize then symmetrize. This is not true if we symmetrize using $y_{ij}^m = [y_{ij} + y_{ji}]/2$. Hence, besides the empirical advantages mentioned in the text, the max convention also has a theoretical advantage, which can be helpful when using metrics (such as node degree) based on simple undirected unweighted networks.

Node strength and degree. The strength of node i in an undirected net-

work[1] is the sum of its edge weights,

$$s_i = \sum_{j=1}^{I} y_{ij}. \tag{4.2}$$

The degree k_i of any node i is the number of edges of positive weight that include that node, a nonnegative integer given by

$$k_i = \sum_{j=1}^{I} \mathbf{1}_{[y_{ij}>0]}, \tag{4.3}$$

where the indicator function $\mathbf{1}_e = 1$ if event e occurs and is 0 otherwise. In an unweighted network, of course, node strength coincides with node degree.

Occasionally it is helpful to distinguish nodes with lots of moderately weighty connections from nodes with just a few very weighty connections. For this purpose, following Barrat et al. (2004) and Opsahl et al. (2010), consider Cobb-Douglas combinations

$$s_{\alpha,i} = s_i^{1-\alpha} k_i^{\alpha}. \tag{4.4}$$

Figure 4.4 below reports the middling case $\alpha = 0.5$, which can be compared to the polar cases $\alpha = 0$ (so $s_{\alpha,i} = s_i$) and $\alpha = 1$ (so $s_{\alpha,i} = k_i$) reported in the text.

Assortativity. An assortativity metric is, in essence, the correlation (across edges) of the strengths of each edge's two nodes. Conceptually, the expression is straight-

---

[1] In a directed network, the expression in equation (4.2) is called the out-strength of node i , and the expression in the same equation with y_{ij} replaced by y_{ji} is called the in-strength.

forward:

$$A(Y) = \rho_{[ij]} = \frac{Cov[ij]}{Var[i]} \equiv \frac{E(s_i - Es)(s_j - Es)}{E(s_i - Es)^2} = \frac{E(s_i s_j) - (Es)^2}{E(s_i^2) - (Es)^2}, \qquad (4.5)$$

where the expectation operator E is understood to average over all edges ij.

The concept is easiest to implement in unweighted undirected networks, once it is understood that non-existent edges (ij such that $y_{ij} = 0$) are ignored. Newman (2002) noted that the mere fact that two nodes share an edge means that their edge counts will be positively correlated, biasing upward the Assortativity calculation. He therefore proposed replacing node degree by excess degree in such networks, netting out the edge in question. That upward bias also seems important for weighted networks, so we define $s_{i\backslash j} = s_i - y_{ij}$ as the excess strength of node i for edge ij, with expected value $\bar{s} = \frac{1}{H} \sum_{i,j=1}^{I} y_{ij} s_{i\backslash j}$, where $H = \sum_{i,j=1}^{I} y_{ij}$. Then the assortativity of an undirected weighted network Y is

$$A(Y) = \frac{\frac{1}{H} \sum_{i,j=1}^{I} y_{ij} s_{i\backslash j} s_{j\backslash i} - \bar{s}^2}{\frac{1}{H} \sum_{i,j=1}^{I} y_{ij} s_{i\backslash j}^2 - \bar{s}^2} \qquad (4.6)$$

We have not seen equivalent expressions in the literature (see Noldus and Van Mieghem, 2015 for a recent review). Leung and Chau (2007) uses edge-weighted averages and covariances but not excess strength in defining assortativity for weighted networks. Many authors follow Newman in using excess degree, but only in unweighted networks. A caveat: directed weighted networks, not used in the present paper, would require a separate definition of $s_{j\backslash i}$.

To verify the unbiased nature of definition (4.6), we computed $A(Y)$ for one

hundred random graphs with the same edge count and edge weight distribution as for the first week of our data. The mean $A(Y)$ is very close to zero with very small standard deviation (-0.0002 ± 0.0024).

Centrality. A node can be considered central if it is on lots of shortest paths. To formalize this intuition, define a path $p = (n_1, n_2, ..., n_k)$ as a sequence of adjacent nodes, i.e., nodes satisfying $y_{n_i n_{i+1}} > 0$ for $i = 1, ..., k-1$. Let $\mathbb{P}_{ij}$ be the set of all paths from node i to node j, i.e., paths satisfying $n_1 = i$ and $n_k = j$. Define the distance from i to j along path $p \in \mathbb{P}_{ij}$ to be the sum of the reciprocals of the edge weights, $L(p) = \sum_1^{k-1} 1/y_{n_i n_{i+1}}$, and define a shortest path from i to j to be any $p^*(ij) \in \text{argmin}\{L(p) : p \in \mathbb{P}_{ij}\}$. A shortest path is generically unique, and the distance from i to j is always uniquely defined by $d(i, j) = L(p^*(ij))$. The distance is always positive for $i \neq j$, and is smaller when the shortest path has fewer and weightier links (edges). By convention, the distance is $+\infty$ if $\mathbb{P}_{ij} = \emptyset$, i.e., if i and j belong to different connected components of the network.

Following the Brandes (2001) generalization of Freeman (1979), define the betweenness centrality of node n as

$$B(n) = \sum_{i,j \neq n} \mathbf{1}_{[n \in p^*(ij)]} \Big/ \sum_{i,j \neq n} 1 \in [0, 1], \qquad (4.7)$$

i.e., the fraction of all relevant node pairs ij that have a shortest path that goes through n.

An alternative intuition is that a node is central if on average it has a short dis-

tance to other nodes. Freeman (1979) and Newman (2001) define the closeness centrality of node n as

$$\tilde{C}(n) = 1/\sum_{n'\neq n} d(n,n').$$

(4.8)

A problem for our purposes is that if even one node n' is very weakly connected to other nodes (or is disconnected) then $\tilde{C}(n)$ will be pushed towards (or will equal) zero for all n. This creates problems in our empirical work, so we prefer to use the less standard Opsahl et al. (2010) definition

$$C(n) = \sum_{n'\neq n} \frac{1}{d(n,n')}.$$

(4.9)

As the sum of reciprocal distances instead of the reciprocal of summed distances, (4.9) is much less sensitive to the weight of the lightest edge. One can see that $C(n)$ is the sum of harmonic mean weights (divided by number of edges) along shortest paths from n to all other nodes. Thus $C(n)$ will increase as distances shorten, as is desirable, but also as the number $I-1$ of other nodes increase. So we will normalize it by dividing by $I-1$.

Transactions. We take as given a finite set of active traders $A = \{1, 2, ..., M\}$ and a finite set of tradable goods indexed $n = 1, ..., N$, and consider bilateral barter transactions observed over some finite time interval $[0, 1]$. Such a transaction is specified by naming the initiating trader $i \in A$, the counterparty $j \in A$, and the net trade vector $x \in R^N$.

Suppose that trader i initiates net trade x with counterparty j at time t. The

convention is that i's post-transaction holdings $\omega(i,t) \in R_+^N$ are related to her pre-transaction holdings $\omega(i,t-) = \lim_{\epsilon \downarrow 0} \omega(t-\epsilon) \in R_+^N$ via

$$\omega(i,t) = \omega(i,t-) + x. \tag{4.10}$$

Of course, j's holdings satisfy

$$\omega(j,t) = \omega(j,t-) - x. \tag{4.11}$$

Using the notation $x_n^+ = \max\{0, x_n\} \geq 0$ and $x_n^- = -\min\{0, x_n\} \geq 0$, the convention can be restated by saying that i trades bundle x^- to j and acquires bundle x^+ in exchange, so $x = x^+ - x^-$ is the net trade vector.

Without loss of generality (just drop exceptions from the lists), we can assume that all M traders transact and that all N goods are traded at least once. Since self-trades and null trades are meaningless, we can assume without loss of generality that $i \neq j$ and $x \neq 0$. Netting out transactions in which a trader both acquires and relinquishes positive amounts of the same good n, we can say without loss of economic content that $x^+ \cdot x^- = 0$. For convenience and with only slight loss of generality, we assume that each price $p_n > 0$, so the price vector is a point in the strictly positive orthant, $p \in R_{++}^N$.

Given a price vector $p \in R_{++}^N$, the value of the bundle i acquires is $v^+ = p \cdot x^+$ and the value of the bundle j acquires is $v^- = p \cdot x^-$. The transaction is budget-balanced at p if $v^+ = v^-$ or, equivalently, if $0 = p \cdot x \equiv \sum_{k=1}^N p_k x_k$.

Trader Network and Goods Network. Suppose that transactions $\langle i(t), j(t), x(t) \rangle \in$ $A \times A \times R^N$ are observed at times $t = t_1, t_2, ..., t_K$, where $0 \leq t_1 \leq t_2 \leq ... \leq t_K \leq 1$, so the transactions are indexed by k. Given price vector p, the observed trader network is a weighted directed network with node set A. The directed edge weight from node ℓ to node m is the value of the bundles that ℓ acquires from m. Thus the trader network is defined by the $M \times M$ adjacency matrix $W = ((w_{\ell m}))$ with entries

$$w_{\ell m} = \sum_{k=1}^{K} v^+(t_k) \mathbf{1}_{[i(t_k)=\ell] \& [j(t_k)=m]} + v^-(t_k) \mathbf{1}_{[j(t_k)=\ell] \& [i(t_k)=m]}, \qquad (4.12)$$

where $\mathbf{1}_e = 1$ if event e occurs and is 0 otherwise. Thus equation (4.12) ignores all transactions except those in which ℓ, as initiator or counterparty, acquires goods from m. Of course, adjacency matrix entries are nonnegative and, by convention, diagonal entries are zero. If all trades are budget-balanced, then (4.12) tells us that the adjacency matrix is symmetric so the trader network is undirected.

The same set of transactions also defines a goods network. The nodes of this network are $n = 1, ..., N$, and the edge weights reflect the value of transactions in which one good is part of the exchange for another. We want the $N \times N$ adjacency matrix $Z = ((z_{nn'}))$ to be symmetric because edges represent mutual exchange values of goods — if the value flows from good n to n' for the initiator then it flows the opposite direction for her counterparty, and there is no reason here to privilege one party over the other.

Specifying the edge weights $z_{nn'}$ takes some thought when trades involving n and n' also include other goods. For example, suppose that the value $v_{n'}^- = p_{n'}^- x_{n'}^-$ of

good n' constitutes half the value $v^- = p \cdot x^-$ of the vector x^- of goods sent by the initiating trader. Then it seems reasonable to assign half the value v_n^+ of good n acquired by the initiating trader to the edge nn', and the other half of v_n^+ to other edges nn'' connecting n to the other goods n'' sent by the initiating trader. More generally, given the v's associated with a trade vector x, we could assign the weight $v_n^+(\frac{v_{n'}^-}{v^-})$ to the edge nn' when good n is a positive component of x^+ and good n' is a positive component of x^-. Treating the initiator and counterparty symmetrically, we would add the term $v_n^-(\frac{v_{n'}^+}{v^+})$ to account for the case where n' is a positive component of x^+ and good n is a positive component of x^-. Of course, both expressions are 0 when the two goods do not appear on opposite sides of the transaction.

If a transaction is not budget balanced, then the denominators differ in the two expressions and symmetry is lost. To recover symmetry, we adopt the convention that both denominators are

$$v = \max\{v^+, v^-\}, \tag{4.13}$$

and define the contribution $v_{nn'}$ to that edge weight of a trade x at time t by the equation

$$v_{nn'}(t) = v_n^+ \left(\frac{v_{n'}^-}{v} \right) + v_n^- \left(\frac{v_{n'}^+}{v} \right) \tag{4.14}$$

Using that expression the goods network matrix entries are

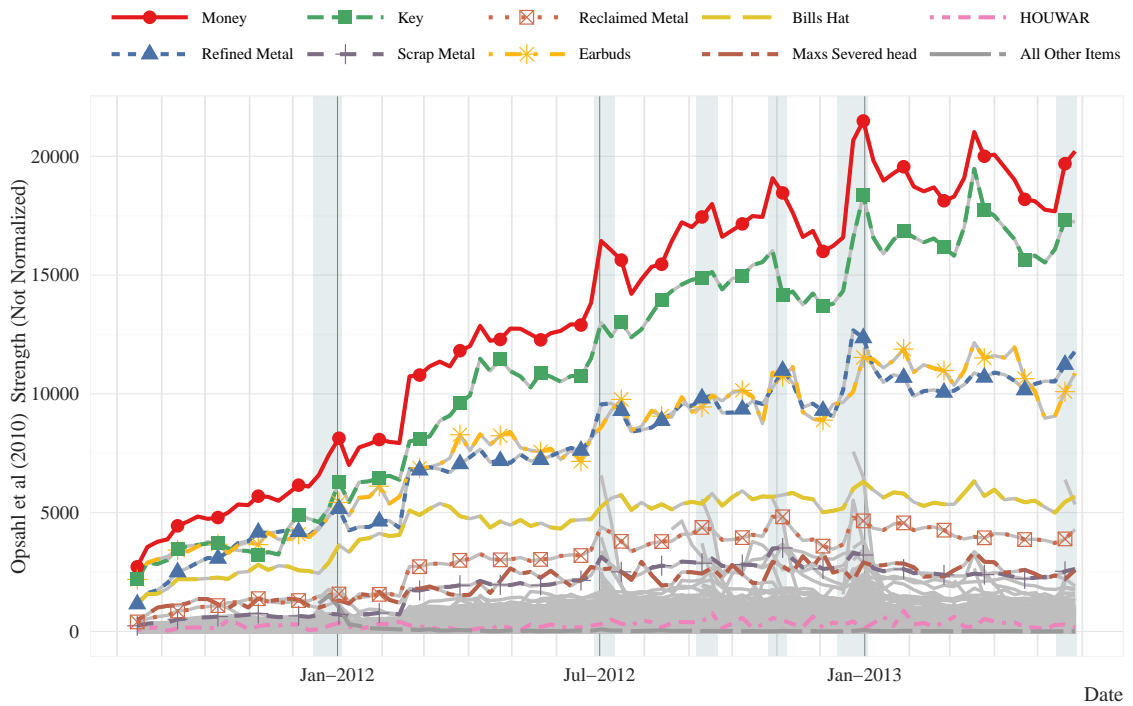$$z_{nn'} = \sum_{k=1}^{K} v_{nn'}(t_k). \tag{4.15}$$

Figure 4.4: Weekly strength $s_{\alpha,i} = s_i^{1-\alpha} k_i^\alpha$ for $\alpha = 0.5$

Money strength is computed for the reduced network (collapsing Money's six constituent nodes); all other strengths are for the full goods network.

Since a good can't appear with opposite sign from itself, the diagonal entries are zero.

Supplementary data analysis. The remainder of this Appendix reports supplementary empirical results. Figure 4.4 shows Opsahl et al. (2010) strength for all items in the economy. Composite Money (in the reduced goods network) is strongest, but all six of its constituents also exhibit considerable strength. (The weakest of them, scrap metal, tracks the strongest other item, Max's Severed Head, fairly closely.) Keys emerge as the strongest constituent, with refined metals and earbuds vying for second place.

The clump. Of the 88 members of the Clump, less than 0.00003 of their transactions are initiated by non-clump traders or are between two clump traders, and these exceptions are confined to a couple of days and have low value. Over our sample the clump transacted 7 million times with over 135 thousand unique counterparties, accounting for over 17.5% of all TF2 trading. Overall gross profit margin (value received minus value delivered divided by the sum of value received and delivered) was roughly 2.1%, with a slight declining trend. Most of its transactions, 94.7%, are one-way, and 91.7% of counterparties trade more than once. Of their first trades with the Clump, 96.9% were one-way inward, delivering value to the Clump. Subsequent one-way trades are increasingly likely to be outward; by the second trade 55% withdraw value. As the number of trades with the Clump increase the percent of trades that withdraw value approaches two-thirds, but only about 18% of these were for a good previously delivered to the Clump by that trader.

As an inventory-carrying market maker for a broad range of goods, what value does the Clump provide to customers? In a post on TF2 forum, an apparent customer explained that it "is fast and straightforward. -Prices set up upfront. -Don't have to delve into a forum/server looking for someone having what i want. -Don't have to chase a user i want to trade with. -No unnecesary [sic] haggling/price changing/offer changing/trade requests during the trade. If buying [at relatively] high [price] is what i have to pay for the convenience of automated trading, so be it. I'm not on it for the benefit, but for the hats. I consider it a price for the service offered." (Steam User Forum, October 31 2012)

Brokers. Our value estimate is conservative because unusual quality items are

not uniform — there are several different forms of unusual. Our single estimated price for unusual items understates the price of the most desirable sorts, which are more likely to be brokered. For example, burning flames is one sort of unusual effect, and recently a burning flames hottie's hoodie was valued at over 250 keys, but we priced it at 40 keys, the median across all unusual hottie's hoodies. Since lower value items tend to be traded more frequently, we believe that our median price estimates are in fact lower bounds for the valuations of the thinly traded unusuals. Because the services of a third party broker are more likely to be requested for relatively valuable items, we believe that the true total value flow that has been mediated by brokers may actually be much larger than the 5000 key estimate which we present.

Denomination ratios. Exchange rates differ from day to day, but over the second half of our sample, Bill's hats traded for about 8 keys, and earbuds usually for 21- 27 keys. This roughly 3:1 ratio is less than the 4:1 or 5:1 ratio for popular coins and bills, but there is a possible historical precedent for a compressed ratio. According to Wikipedia, "In the Great Recoinage of 1816, the guinea was replaced as the major unit of currency by the pound and in coinage with a sovereign. Even after the coin ceased to circulate, the name guinea was long used to indicate the amount of 21 shillings (£1.05 in decimalised currency). The guinea had an aristocratic overtone; professional fees and payment for land, horses, art, bespoke tailoring, furniture and other luxury items were often quoted in guineas until a couple of years after decimalisation in 1971."

Price dispersion as a proxy for transaction costs. Once most transactions go through a money good, it becomes much easier to detect and arbitrage price

discrepancies. Indeed, if anyone posts (or even hints) that they are willing to buy at a price higher than the price at which someone is willing to sell, then anyone aware of the two prices could accept both offers and pocket the difference.

This standard argument implies that every accepted ask price observed over a short period of time is above all accepted bid prices. Assuming equal numbers of the two sorts of transactions, we conclude that all prices above the median are accepted asks, whose median thus is at the 75th percentile, while the median of accepted bids is at the 25th percentile. Hence their difference, the interquartile range, is a proxy for round-trip transaction cost. To maintain comparability across goods, it makes sense to express the interquartile range as a percentage of median price, so for item n at time t, we define

$$SIQR_{nt} = \frac{100 * [p_{75,nt} - p_{25,nt}]}{p_{50,nt}}, \qquad (4.16)$$

where $p_{z,nt}$ is the zth percentile of the imputed prices associated with good n and time interval t. We concede that $SIQR_{nt}$ may somewhat overstate the round trip cost, but there is no reason to think that the degree of overstatement changes systematically over time or across goods. Of course, by definition, $SIQR_{nt}$ is a robust and direct measure of price dispersion.

To aggregate SIQR across goods, we take the value-weighted mean adjusted for sample size,

$$SIQR_t = \frac{\sum_n \eta_{nt} SIQR_{nt}}{\sum_n \eta_{nt}}, \qquad (4.17)$$

where the weight $\eta_{nt} = p_{50,nt} k_{nt}^{1.5}$ is the square root of the number of transactions k_{nt}

(to capture the sample precision) times a robust estimate of the relevant transaction value $p_{50,nt}k_{nt}$. The sum is taken over all goods n for which the time interval $[t-30, t]$ includes at least $k_{\min} = 100$ transactions involving good n. To aggregate $SIQR_t$ across time, we simply take a simple trailing average.

Our results are robust to a variety of other choices of $k_{\min}$; of course, lower choices of $k_{\min}$ generally result in choppier time series. Also, we find similar trends in SIQR when we replace $k_{nt}^{1.5}$ in the definition of the weight by $k_{nt}^{0.5}$, as would be appropriate if, instead asking how much should you expect to lose on a round trip for an item of typical value, you asked the question for a typical trade of whatever size.

Recent developments in the TF2 economy. A future avenue we hope to explore involves the collapse of earbuds which started around the beginning of 2015, dropping from a value of approximately 30 keys down to their current estimated value of 5 keys. We propose that Valve's introduction of a centralized dollar denominated posted-price marketplace replaced the use of earbuds as the preferred medium of exchange for high-value items and thus the value-in-exchange of earbuds became equal to their value-in-use as a cosmetic item which was much less than their exchange value when they were used as a primary currency.

4.3   Appendix C

Chapter 3: Minimum Effort Experiment

4.3.1   Order of Treatments

Table 4.3: Treatment Order for Constant β Sessions

Version	Period 1-4				Period 5-8				Period 9-12			
I	S	M	S	M	S	M	M	S	M	S	S	M
II	S	M	S	M	M	S	S	M	S	M	M	S
III	M	S	M	S	M	S	S	M	S	M	M	S
IV	M	S	M	S	S	M	M	S	M	S	S	M

Between session position of SEVERE (=S) and MILD (=M) is exchanged. The same randomization for JUMP and GRADUAL was implemented

The order in which participants were assigned each treatment level is listed in Table 4.3. These four patterns ensured that participants saw balanced treatment level orderings across sessions.
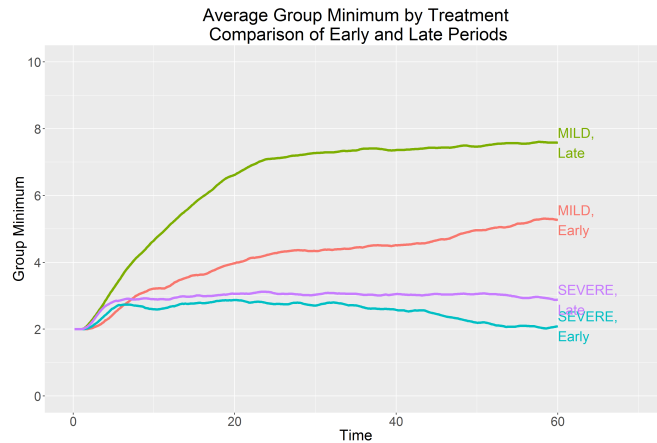
4.3.2   Additional Results

Figure 4.5: Average Group Minimum Play by Treatment, comparison between first four and last four periods
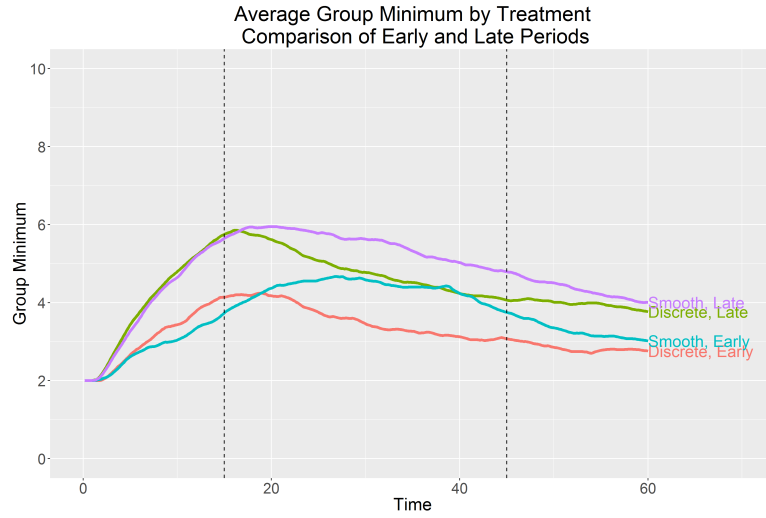
Figure 4.6: Comparison between Early and Late Periods for JUMP and GRADUAL Sessions

Figures 4.5 and 4.6 demonstrate that group minima were generally lower in early periods than in late periods. Because of this learning process, we restricted our analysis to the final 8 periods and excluded the first 4 to ensure that participants had the opportunity to understand the structure of the game.

Table 4.4: Proportion of Time Groups Spent at MILD and SEVERE Group Min, Constant Sessions

Treatment	Full $X^{min} \leq 3$	Full $X^{min} \geq 8$	End $X^{min} \leq 3$	End $X^{min} \geq 8$
SEVERE	0.74	0.09	0.75	0.13
MILD	0.24	0.41	0.17	0.62

Table 4.5: Proportion of Time Groups Spent at MILD and SEVERE Group Min, Changing Sessions

Treatment	Full $X^{min} \leq 3$	Full $X^{min} \geq 8$	End $X^{min} \leq 3$	End $X^{min} \geq 8$
JUMP	0.41	0.13	0.55	0.15
GRADUAL	0.36	0.20	0.51	0.21

Tables 4.4 and 4.5 show the proportion of the time that group minima were

"low" (3 or less) and "high" (8 or higher) in each treatment level for both full periods and the final 15 seconds of each period. This again shows that SEVERE tended to end up at poor equilibria more often than MILD and that there is little difference between JUMP and GRADUAL.

Table 4.6: Wilcoxon and Mann-Whitney, last 15 seconds

Comparison	p	statistic
Paired Constant SEVERE vs Constant MILD	0.008	0
Paired JUMP vs GRADUAL	0.195	8
Constant SEVERE vs JUMP	0.235	20
Constant SEVERE vs GRADUAL	0.065	14
Constant MILD vs JUMP	0.002	60
Constant MILD vs GRADUAL	0.001	61

Table 4.7: Kolmogorov-Smirnov Tests Comparing Treatments at the end of the period

Test	Alternative Hypothesis	p
GRADUAL versus SEVERE	CDF lies below	.047
JUMP versus SEVERE	CDF lies below	.030
GRADUAL versus MILD	CDF lies above	<.001
JUMP versus MILD	CDF lies above	<.001
GRADUAL versus JUMP	CDFs not equal	.996

Tables 4.8 and 4.7 show test statistics and associated p-values with the varios Wilcoxon Rank Sum, Mann-Whitney, and Kolmogorov-Smirnov tests we ran on our data.

4.3.3 Early Periods

This section displays all relevant figures when focusing on the first 4 periods, which were excluded from analysis in the main body. The results we find point in the same direction as the conclusions drawn above but are generally not as strong.
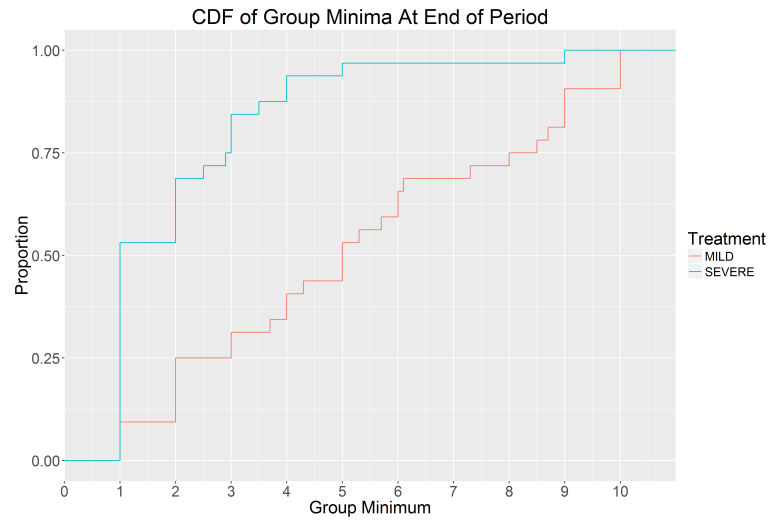

Figure 4.7: CDF of Ending Group Minima, Early Periods Constant Sessions

Table 4.8: Wilcoxon and Mann-Whitney, last 15 seconds, only early periods

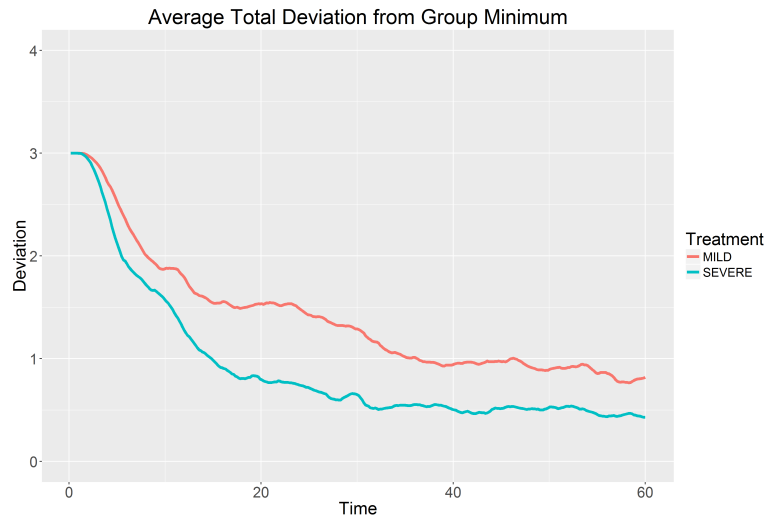| Comparison	p	statistic
Paired JUMP vs GRADUAL	0.383	11
Paired SEVERE vs MILD	0.008	0
SEVERE vs JUMP	0.798	29
SEVERE vs GRADUAL	0.083	15
MILD vs JUMP	0.065	50
MILD vs GRADUAL	0.028	53

Figure 4.8: Total Deviation by Treatment, Early Periods Constant Sessions
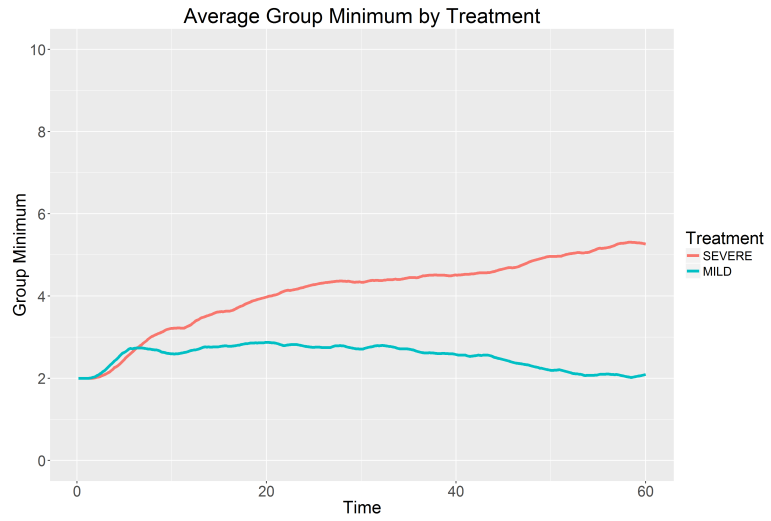

Figure 4.9: Mean Group Minimum by Treatment, Early Periods Constant Sessions
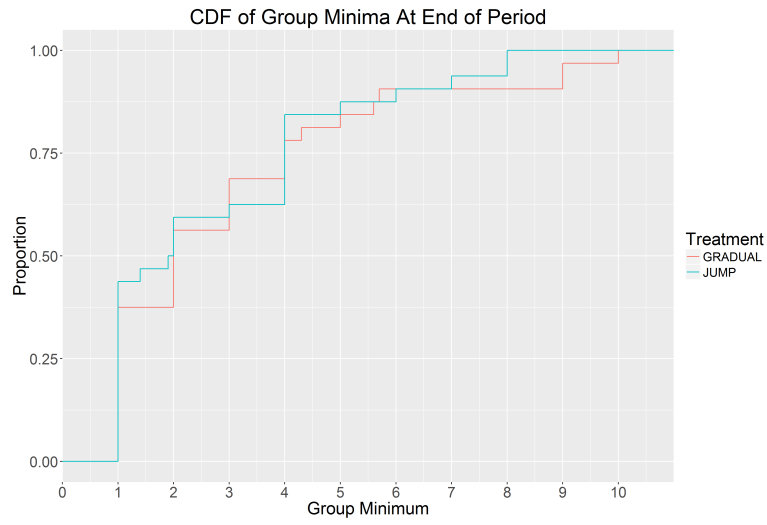
Figure 4.10: CDF of Ending Group Minima, Early Periods Gradual Sessions
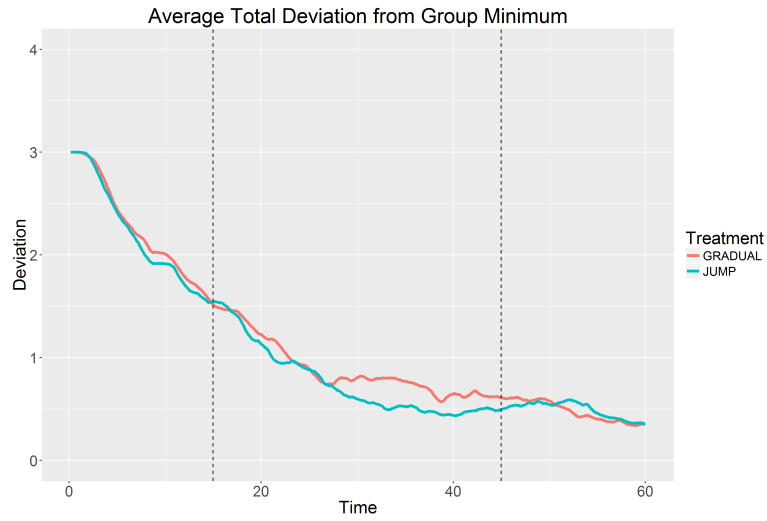

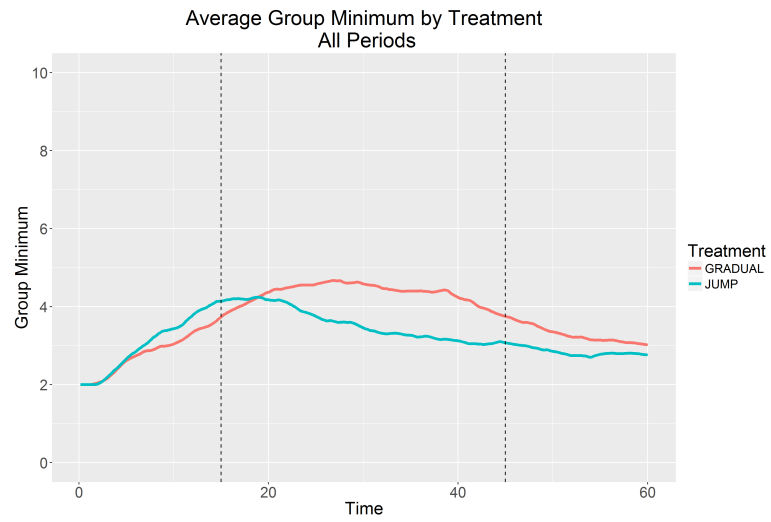Figure 4.11: Total Deviation by Treatment, Early Periods Gradual Sessions

136

Figure 4.12: Mean Group Minimum by Treatment, Early Periods Gradual Sessions

Comparing these figures to those from the first 8 periods, the same trends are all present.

Chapter 5

References

Bibliography

George A Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, pages 488–500, 1970.

Abhijit Banerjee and Eric Maskin. A walrasian theory of money and barter. *The Quarterly Journal of Economics*, 111(4):955–1005, 1996.

Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

Alain Barrat, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.

Matt Baumer and Curtis Kephart. Aggregate dynamics in a large virtual economy: Prices and real activity in team fortress. *University of California Santa Cruz Working Paper Series*, 2015.

Morten L Bech and Enghin Atalay. The topology of the federal funds market. *Physica A: Statistical Mechanics and its Applications*, 389(22):5223–5246, 2010.

S. K. Berninghaus and K.M. Ehrhart. Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior and Organization*, 37(2):231–248, 1998.

S. K. Berninghaus and K.M. Ehrhart. Coordination and information: Recent experimental evidence. *Economics Letters*, 73(3):345–351, 2001.

M. Bigoni, M. Casari, A. Skyzpacz, and G. Spagnolo. Time horizon and cooperation in continuous time. *Econometrica*, 83(2):587–616, 2015.

James V Bradley. Distribution-free statistical tests. 1968.

Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

Kenneth Burdett, Alberto Trejos, and Randall Wright. Cigarette money. *Journal of Economic Theory*, 99.1:117–142, 2001.

United States Bureau of Labor Statistics. *BLS handbook of methods*. US Department of Labor, Bureau of Labor Statistics, 2014. URL http://www.bls.gov/opub/hom/.

G.P. Cachon and C.F. Camerer. Loss avoidance and forward induction in experimental coordination games. *The Quarterly Journal of Economics*, 111(1):165–194, 1996.

Robert Cage, John Greenlees, and Patrick Jackman. Introducing the chained consumer price index. In *International Working Group on Price Indices (Ottawa Group): Proceedings of the Seventh Meeting*, pages 213–246. Paris: INSEE, 2003.

Edward Castronova. Virtual worlds: A first-hand account of market and society on the cyberian frontier. 2001.

Edward Castronova. A test of the law of demand in a virtual world: Exploring the petri dish approach to social science. Technical report, CESifo working paper, 2008.

Edward Castronova, Dmitri Williams, Cuihua Shen, Rabindra Ratan, Li Xiong, Yun Huang, and Brian Keegan. As real as real? macroeconomic behavior in a large-scale virtual world. *New Media & Society*, 11(5):685–707, 2009.

Syngjoo Choi, Andrea Galeotti, and Sanjeev Goyal. Trading in networks: theory and experiments. *Cambridge-INET Working Paper*, 8, 2014.

C. Deck and N. Nikiforakis. Perfect and imperfect real-time monitoring in a minimum-effort game. *Experimental Economics*, 15(1):71–88, 2012.

W Erwin Diewert. Hedonic regressions: a review of some unresolved issues. In *7th Meeting of the Ottawa Group, Paris, May*, pages 27–29, 2003.

Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

David Easley and Jon Kleinberg. Networks, crowds, and markets, 2012.

Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1979.

D. Friedman and R. Oprea. A continuous dilemma. *American Economic Review*, 102 (1):337–363, 2012.

International Monetary Fund. Export and import price index manual: theory and practice. INTERNATIONAL MONETARY FUND, 2009.

B. Greiner. Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125, 2015.

James Heckman. Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy*, 109(4):673–748, 2001.

J. Hirshleifer. From weakest-link to best-shot: The voluntary provision of public goods. *Public Choice*, 41(3):371–386, 1983.

Peter Howitt and Robert Clower. The emergence of economic organization. *Journal of Economic Behavior & Organization*, 41(1):55–84, 2000.

ILO-IWGPS. Consumer price index manual: Theory and practice. 2004.

Matthew O Jackson. An overview of social networks and economic applications. *The handbook of social economics*, 1:511–85, 2010.

Matthew O Jackson et al. *Social and economic networks*, volume 3. Princeton University Press Princeton, 2008.

William Stanley Jevons. *Money and the Mechanism of Exchange*, volume 17. 1885.

Y. Kamijo, H. Ozono, and K. Shimizu. Overcoming coordination failure using a mechanism based on gradualism and endogeneity. *Experimental Economics*, 19(1):202–217, 2016.

C. Kephart and L. Rose. Stability in competition? hotelling in continuous time. *Working Paper*, 2016.

Alan Kirman. The economy as an evolving network. *Journal of evolutionary economics*, 7(4):339–353, 1997.

Alan P Kirman and Nicolaas J Vriend. Evolving market structure: An ace model of price dispersion and loyalty. *Journal of Economic Dynamics and Control*, 25(3):459–502, 2001.

Nobuhiro Kiyotaki and Randall Wright. On money as a medium of exchange. *The Journal of Political Economy*, pages 927–954, 1989.

M. Knez and C. Camerer. Creating expertational assets in the laboratory: Coordination in "weakest link" games. *Strategic Management Journal*, 15:101–119, 1994.

Stephen E Lankenau. Smoke'em if you got'em: Cigarette black markets in us prisons and jails. *The Prison Journal*, 81(2):142–161, 2001.

Vili Lehdonvirta. Virtual economics: Applying economics to the study of game worlds. In *Proceedings of the 2005 Conference on Future Play (Future Play 2005), Lansing, MI*, 2005.

Vili Lehdonvirta. Virtual worlds don't exist: Questioning the dichotomous approach in mmo studies. *Game Studies*, 10(1), 2010.

A. Leng, L. Frisen, K. Kalayci, and P. Man. A minimum effort coordination game experiment in continuous time. *Working Paper*, 2016.

CC Leung and HF Chau. Weighted assortative and disassortative networks model. *Physica A: Statistical Mechanics and its Applications*, 378(2):591–602, 2007.

Fredrik Liljeros, Christofer R Edling, Luis A Nunes Amaral, H Eugene Stanley, and Yvonne Åberg. The web of human sexual contacts. *Nature*, 411(6840):907–908, 2001.

Karl Menger. On the origin of money. *The Economic Journal*, 2(6):239–255, 1892.

Mark EJ Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132, 2001.

Mark EJ Newman. Assortative mixing in networks. *Physical review letters*, 89(20): 208701, 2002.

Rogier Noldus and Piet Van Mieghem. Assortativity in complex networks. *Journal of Complex Networks*, page cnv005, 2015.

Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

Joseph Ostroy and Ross Starr. Money and the decentralization of exchange. *Econometrica*, 42(6):1093–1113, 1974.

Joseph M. Ostroy and Ross M.' Starr. The transactions role of money. In B. M. Friedman and F. H. Hahn, editors, *Handbook of Monetary Economics, Ed 1*. North-Holland Publishing Company and American Elsevier Publishing Company.

Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200, 2001.

J. Pettit, D. Friedman, C. Kephart, and R. Oprea. Software for continuous game experiments. *Experimental Economics*, 17(4):631–648, 2014.

Robert A Radford. The economic organisation of a pow camp. *Economica*, pages 189–201, 1945.

Sherwin Rosen. Hedonic prices and implicit markets: product differentiation in pure competition. *The Journal of Political Economy*, pages 34–55, 1974.

Steven Salop and Joseph Stiglitz. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *The Review of Economic Studies*, 44(3):493–510, 1977.

Steven Salop and Joseph Stiglitz. The theory of sales: A simple model of equilibrium price dispersion with identical agents. *The American Economic Review*, 72(5):1121–1130, 1982.

Adam Smith. An inquiry into the nature and causes of the wealth of nations: Volume one. 1776.

Leo Törnqvist. The bank of finlands consumption price index. *Bank of Finland Monthly Bulletin*, 10(1936):1–8, 1936.

J. B. Van Huyck, R. C. Battalio, and R. O. Beil. Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, 80(1):234–248, 1990.

J. B. Van Huyck, R. C. Battalio, and R. O. Beil. Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *The Quarterly Journal of Economics*, 106(3):885–910, 1991.

J. B. Van Huyck, J.P. Cook, and R. C. Battalio. Adaptive behavior and coordination failure. *Journal of Economic Behavior and Organization*, 32(4):483–503, 1997.

J. B. Van Huyck, R. C. Battalio, and F. W. Rankin. Evidence on learning in coordination games. *Experimental Economics*, 10(3):205–220, 2007.

Hal Varian. A model of sales. *American Economic Review*, 70(4):651–659, 1980.

R. Weber, C. Camerer, Y. Rottenstreich, and M. Knez. The illusion of leadership: Misattribution of cause in coordination games. *Organization Science*, 12(5):582–598, 2001.

R.A. Weber. Managing growth to acheive efficient coordination in large groups. *American Economic Review*, 96(1):114–126, 2006.

Steve Williamson and Randall Wright. Barter and monetary exchange under private information. *The American Economic Review*, 84(1):104–123, 1994.

M. Ye, S. Asher, and P. Nikolov. One step at a time: Does gradualism build coordination? *Working Paper*, 2014.