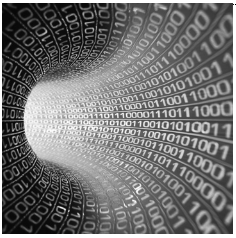
New Media & Society

http://nms.sagepub.com

As real as real? Macroeconomic behavior in a large-scale virtual world

Edward Castronova, Dmitri Williams, Cuihua Shen, Rabindra Ratan, Li Xiong, Yun Huang and Brian Keegan

New Media Society 2009; 11; 685
DOI: 10.1177/1461444809105346



new media & society

Los Angeles, London, New Delhi, Singapore and Washington DC
Vol 11(5): 685–707 [DOI: 10.1177/1461444809105346]

ARTICLE

As real as real? Macroeconomic behavior in a large-scale virtual world

EDWARD CASTRONOVA
Indiana University, USA

DMITRI WILLIAMS
CUIHUA SHEN
RABINDRA RATAN
LI XIONG
University of Southern California, USA

YUN HUANG
BRIAN KEEGAN
Northwestern University, USA

Abstract

This article proposes an empirical test of whether aggregate economic behavior maps from the real to the virtual. Transaction data from a large commercial virtual world – the first such data set provided to outside researchers – is used to calculate metrics for production, consumption and money supply based on real-world definitions. Movements in these metrics over time were examined for consistency with common theories of macroeconomic change. The results indicated that virtual economic behavior follows real-world patterns. Moreover, a natural experiment occurred, in that a new version of the virtual world with the same rules came online during the study. The new world's macroeconomic aggregates quickly grew to be nearly exact replicas of those of the existing worlds, suggesting that 'Code is Law': macroeconomic outcomes in a virtual world may be explained largely by design structure.

Key words

behavioral data • computer games • economics • Everquest 2 • mapping • MMORPG • supercomputing analysis • video games • virtual world

In *The Treachery of Images* (1928–9), René Magritte painted a picture of a pipe above the words: 'Ceci n'est pas une pipe' ('This is not a pipe'). In virtual worlds, there are no pipes either, yet users manage to use virtual objects and characters in a vast array of social transactions. When observers see a virtual sword on the computer screen, the paraphrase of Magritte's admonition – 'Ceci n'est pas une épée' ('This is not a sword') – surely comes to mind. However, is the sword only an image? Or does it become invested with some kind of socially constructed realness as a result of playing a role in human communication and exchange?

Such questions often have been posed at the level of philosophy, but the advances in technology which have given rise to virtual worlds have given them a practical and even empirical dimension (Jankowski, 2007). This article is a test of those practical aspects. By drawing on mainstream economic theories, it uses data from within a virtual world to test the 'realness' of those many swords and items. Here, this phenomenon is termed 'mapping', which is defined as the extent to which offline phenomena apply in a virtual world. By examining transactions from a large commercial virtual world with hundreds of thousands of players, the current mapping test concerns whether the items and economic behaviors within a virtual world function in the same way that they would in the real world – where, it is noted, currency is also largely representational (a dollar is only a piece of paper, but is assumed to have a value in gold backed by the US treasury). Do the same behaviors that make a US dollar bill economically 'real' also make a virtual sword economically 'real'? Does economic behavior map from the real world into the virtual?

Macro-level economic behavior in these spaces has existed since virtual worlds first rose to prominence in the late 1990s, but access to players and behavioral data access has always been problematic. Researchers do not own or control a large-scale virtual world and thus they can only study existing commercial worlds, unless they are able to make one (Castronova et al., 2008). However, commercial virtual worlds typically place restrictions on aggregate data collection by players in order to reduce the likelihood of fraud and exploitation. At the same time, to protect their business practices, companies are often unwilling to release the data that they do collect.

For example, efforts to collect survey data are hampered by researchers' inability to gain access to lists of players from which to draw random samples. Lists of player contact information are a precious resource to companies.

This article takes advantage of the first ever large-scale database access of a virtual world for academic use. Sony Online Entertainment allowed access to data from the virtual world game *EverQuest 2* (*EQ2*), a massively multiplayer online role-playing game (MMORPG). Researchers were granted access to a large trove of datafiles that the game's managers had created. The dataset is a comprehensive capture of the actions and communications of hundreds of thousands of players over time. This dataset contains more than 300 million individual transactions: lists of thousands of items sold, with purchase amounts and prices. In structure, these data points are identical to those used by real-world governments to calculate measures of aggregate economic activity, such as gross domestic product (GDP) and the inflation rate.

Yet Magritte's admonition hovers in the background: is this an economy or not? Being comprehensive and generated by the game system directly (and therefore relatively assured of accuracy), the data allow a direct test of whether an economic 'mapping' can be said to have occurred. If indeed it is possible to calculate aggregate economic statistics such as virtual GDP and virtual inflation, and if these follow the same rules of behavior as the corresponding real-world aggregates, this would constitute the first piece of evidence that some sort of real-to-virtual mapping does occur at the level of aggregate human behavior – at least in one virtual world. The following sections provide a basic context for the virtual world under study, and then outline the theories and predictions of these behaviors.

UNDERSTANDING ECONOMIC BEHAVIORS IN A VIRTUAL WORLD

EQ2 is a large-scale fantasy world in which hundreds of thousands of players interact to advance their characters, create communities with one another and sometimes engage in conflict with others (Williams et al., 2008). The underlying story of the game revolves around a battle between two factions: good and evil. Players are allowed to create multiple characters, shown on screen by visual representations known as 'avatars'. These may play many different roles and can have a range of appearances, but every player is subject to a basic need to acquire and use goods and services such as swords and blacksmiths, tailors, food vendors and most of the items one would expect in a very detailed fantasy world.

Because there is only so much virtual 'space' yet a large number of players, the game is divided into dozens of parallel and identical versions of the game world, commonly known as 'servers'. Each server can be considered as a stand-alone, parallel virtual world with a stable population of characters. Currently *EQ2* is operating on 25 servers, 19 of which are for players located in the USA.

On its face, *EQ2* bears the hallmarks of a functioning economy, with the buying and selling of goods, banking, the creation of goods and the equivalent of income, derived from taking currency and items from defeated monsters. For this reason, this study introduces a series of standard economic measures and a simple, well-known economic model of behavior, all of which will be systematically tested within *EQ2*. The most basic measures of an overall state of an economy are its GDP, price level ('prices'), inflation and the money supply (Landefeld et al., 2008).[1] GDP is the main measure of the size of an economy:[2] it is defined as the value of all goods produced in an economy in a period of time and can be split into nominal and real GDP. When calculating GDP, one must apply prices to the produced goods in order to determine their value. If one uses the prices that are in effect at the time of production, the result is nominal GDP. If instead one uses a set of prices from a fixed time period, the result is real GDP, set to a base level associated with a particular time period. The difference between nominal and real GDP is that nominal GDP changes as a result of changes in production amounts as well as the price of goods. Real GDP changes only as a result of changes in the amount of production. Prices – the values of the produced goods – are fixed. Economists agree that under certain conditions, these aggregate measures should move in certain ways (Landefeld et al., 2008). These measures and basic laws of motion are the key to identifying any macroeconomic mapping.

The price level is an abstract index set to equal 100 in a given time period. One takes the goods consumed in an economy in that time period and treats them as a 'market basket', or a fixed bundle of goods. In each time period thereafter, the prices that exist in a new time period are compared to the first basket. Thus if the basket costs 10 percent more than the next period, then the price level in the next period is 110 (10% more than 100). The two remaining measures are relatively simple. The inflation rate is defined as the percentage change in the price level. The money supply is defined as the number of units of currency in an economy at a point in time. Typically, these measures are applied at the level of countries and in a timescale of years, but they are applicable in principle on a variety of scales. In the real world, as in the virtual, one could apply these measures to neighborhoods, towns, cities, provinces, countries and continents. Further, they could be applied to any time period – day, week, month, year or decade.

The quantity theory of money (QTM) is a basic theory of motion using these measures along with population and stability. It states that under simple conditions, the price level is proportional to the ratio of money in the system and to the amount of goods being traded. For example, the QTM states that if the amount of money goes up, all else being equal, the price level should go up. If the amount of traded goods goes up, all else being equal, the price level should go down. If the population increases, usually there will be more trading, and this should lower the price level (if nothing else changes). Normally, the aggregates of an economy do not fluctuate wildly from period to period. This is an empirical observation, but not a theoretical one. In the QTM, if the quantity of money doubles or triples, the price level should double or triple immediately. Such events are rare in practice, because the inputs to the economy usually do not change so abruptly. Nonetheless, it has happened. In the German hyperinflation of 1922–3, the money supply was vastly increased in a very short timespan and the price level increased rapidly (Paarlberg, 1992). In general, economies usually have relatively stable prices.

Contemporary economists debate the extent to which the QTM holds (Cunningham, 1992; Hafer and Wheelock, 2001; Sargeant and Wallace, 1981), but a virtual economy is a unique test case because there is no banking or financial sector, international trade, government sector or taxation. All four of these factors are cited as reasons why a simple quantity theory is not sufficient to explain price dynamics in real economies (Brown et al., 1963; Samuelson and Solow, 1960). In other words, a virtual economy offers a controlled environment to test the quantity theory of money which is not possible in the real world.

Taken together, GDP, prices and the QTM constitute a simple and robust approach to measuring aggregate economic behavior. However, they remain imperfectly validated on large scales, even in the real world, because the total measurement of an entire economy has been rarely possible unless the economy in question was small enough to observe. For example, researchers have been able to study token-based economies within small, closed settings such as prisons and mental health facilities, where cigarettes were often used for exchange (Kazdin, 1977; Lankenau, 2001; Radford, 1945). These studies provided strong support for the economic theories described here (Cox, 1997), but also required researchers to be on hand to gather the data. This limits the approach to small-scale work and introduces measurement error from the presence of the researchers themselves.

For large economies, calculating exact values for macroeconomic metrics (e.g. GDP) is technically impossible because it requires obtaining all of the prices and quantities of goods and commodities transferred within an economic system. Even the Bureau of Economic Analysis, the only institution sanctioned by the US Department of Commerce to calculate US GDP, relies on estimates in order to perform this calculation (Fixler and Grimm, 2008). For example, during non-benchmark years, spending on cars is calculated as the number of cars sold multiplied by the ‘average list price with options adjusted for transportation charges, sales tax, dealer discounts and rebates’ (US Department of Commerce, 1998: 14). Even during benchmark years, in which the Census Bureau provides the relevant information, there is still a margin of error based on the unreliability of self-reported data and other factors. While these are certainly the best methods available for estimating the inputs into the GDP calculation, the final value is clearly that – an estimation. Even small mismeasurements can lead to important consequences. Studies of policymakers responding to economic signals have shown that small errors lead to large volatility in their decision-making, which consequently destabilizes the real economy. Even minor increases in the accuracy of a statistic can reduce this volatility effect (Bomfim, 2001).

While economists normally have to rely on estimates and sampling for large-scale research and on fallible direct observation for small-scale research, virtual worlds represent the potential to improve on both aspects. Virtual worlds can be large-scale, with all the data observed and recorded automatically. Whereas macroeconomists usually have to employ sampling methods to estimate economic aggregates in the real world, a complete database from an existing virtual world obviates such methods, since all transactions are recorded centrally. Similarly, the calculation is simplified since the economy examined here has no investment or government spending, thus eliminating these inputs into the economy and leaving consumption as the only contribution to economic production. Just as people's purchases of food, electronics and cars count towards a consumption tabulation, the players' purchases of food, clothing, accessories, furnishings, weapons, materials, collectibles and scrolls count towards in-game gross economic consumption. All the data are collected as a result of players taking actions. There is no need for estimation, sampling or even contacting the players involved, since the recording of data happens automatically without the players' knowledge. It follows that demand conditions cannot be a factor (Webb et al., 1966) and there can be no 'Hawthorne effects' (Cook and Campbell, 1979). Practically speaking, this is an unobtrusive methodology which could not have existed before.

The research method itself warrants reflection: for the first time, researchers were given access to the data that game companies regularly collect as their games operate. Relative to real-world methods where, for example, an interviewer directly asks questions of a respondent, these methods seem unobtrusive. Although unobtrusive, the methods still involve the observation of individual behavior and their invisibility perhaps calls for careful treatment. To what extent is this form of data collection also a collaboration with game owners in the management of their players? How does one team's special access to inside data interact with the norm of making data available to all who wish to see it? In this as in other areas, virtual worlds research requires an alert sense of research ethics. Here, the guiding principle is the same as in previous work: to protect the research subject, and the chief protections are twofold. First, the data were anonymous in that real-world identities were never connected to characters within the game; data were based on account IDs that could not be connected to real people. Second, both the game operator and the research team signed legal agreements prohibiting the use of the results for punitive or commercial uses, i.e. the game company also was bound from using the results to single out any of the accounts involved.

Although there have been attempts to conduct macroeconomic measurement within virtual worlds, researchers have been stymied by the simple problem of access. Without permission to access the private datastores of commercial enterprises, studies have had to reply on partial sampling (Nash and Schneyer, 2004) or to remain purely theoretical (Simpson, 1999). The most-cited work in this area is Castronova (2001), who had to rely on secondhand sources and to use estimation.[3]

This article, by contrast, relies on a census of transaction data from within *EQ2*. These data were used to answer research questions and hypotheses that use the quantity theory of money to test the more general concept of mapping human behaviors on a large scale from the real to the virtual.

RESEARCH QUESTIONS AND HYPOTHESES

The first, most exploratory, question involves whether virtual economies resemble real ones *prima facie*. If virtual worlds are at all similar to real ones, then the core elements of their economies should look and work like real ones. The most basic test of this is whether products in virtual worlds can be broken down into categories that look like real-world categories (e.g. clothing, durable goods, etc.) rather than nonsensical ones that bear no relationship to the real (e.g. pixie dust):

> RQ1: What items are sold in the economy of a virtual world?

The economic concept behind this question lies in the basic definition of an economy. It is a matter of first principles that an economy has three functions: production of goods, allocation of resources to production and distribution of produced goods to different people (Hall and Lieberman, 2003). Moreover, if there has been some kind of mapping from the real to the virtual, one should expect that these three functions are being exercised on virtual goods, whose uses and affordances map directly onto those of similarly-named goods in the real world:

> H1: *EQ2*'s economy is based on items with Earth-like names and uses, such as 'food', 'clothing', and 'housing'.

While item categories represent necessary conditions for an economy to exist, they are not as important as measures such as GDP and price level. However, if those items were economically empty, any aggregates based on them would fluctuate wildly and without apparent reason. If, for example, people do not map their notion of time value (understood as the amount of money they must be paid to sacrifice an hour of their time to an alternative activity, e.g. work) from the real to the virtual, then the value of time devoted to making a suit of magic armor might be effectively zero, even though it takes a person 10 hours of play time to complete the task. In the real world, that kind of time requirement would result in a positive price for the item. Moreover, that price would not change dramatically over time, since its underlying determinant – the individual's willingness to give time to other tasks in return for compensation – is a core psychological feature. However, if time means nothing in the virtual world, then the price of that item (indeed, all items) would have no fixed point of reference and would vary erratically with the whims and play strategies of producers and consumers. Conversely, strong mapping would suggest that macroeconomic aggregates are anchored to the fundamental behavioral processes and mental judgments of the players, and therefore would maintain some kind of stability. Whether any mapping can exist would be ascertained by observing aggregates such as GDP and price level. As a hypothesis, a strong mapping would be indicated by the basic magnitude of virtual economic production measured in terms of real-world money. Therefore:

> RQ2: Is it possible to calculate nominal and real GDP and price levels in a virtual economy?

How these levels should be assessed is another question. If the virtual money remains in the virtual world, it can be argued that this money is useful to economists for theory testing and useful psychologically to players, but not actually relevant for the real world, where virtual currency theoretically holds no value. However, because players place a value on that currency and because there is in fact a time value for their virtual assets, it has become possible to trade real currency for virtual currency and vice versa (Dibbell, 2006). Many players would prefer to buy an in-game item for real-world dollars rather than earn it through (often laborious) play sessions. As a result, virtual currency has a market and an exchange rate, frequently brokered through third-party web-based services such as Internet Gaming Entertainment (IGE, www.ige.com). The exchange rate from those services can be used to translate virtual values into their equivalent real-world values. Thus this study hypothesizes:

> H2: GDP in a virtual economy is non-trivial in terms of real-world money.

A third question is whether the interactions between macroeconomic aggregates map back from the real to the virtual. As the quantity theory of money suggests, price levels should change according to some basic laws of motion. Since virtual economies are relatively simple, these laws ought to hold here. Therefore, the specific changes predicted by the quantity theory of money can be tested:

> H3: If money supply rises in a virtual economy, all else being equal, the price level should rise.

> H4: If population rises in a virtual economy, all else being equal, the price level should fall.

While the economic measures are of interest in their own right, the broader concern here is the concept of behavioral mapping and what causes it. Some early virtual world research has found human behaviors that map directly from the real to the virtual, such as players' preferences for personal space or eye contact (Bailenson et al., 2005; Yee and Bailenson, 2007). However, these were all interpersonal effects, not macro-level behaviors. Macroeconomic behaviors provide a test of Lessig's (1999) 'code is law' concept: that the mere structure of a virtual world is enough to create behavior in a predictable way; if all else is equal and populations are randomly distributed, behaviors should be a predictable outcome of the world's structure. In other words, if Lessig (1999) is correct and code truly is a direct driver of aggregate behaviors, one world's patterns should be directly replicated when an exact copy of that world is made and populated with the same kinds of people. Thus:

> RQ4: Can the mere structure of a virtual world create predictable economic behaviors?

METHOD

Study selection

The research questions and hypotheses were addressed by examining computer-collected data on human behaviors that occurred within the virtual fantasy world *EQ2*. The current study focused on *EQ2* due to its popularity, its representativeness of mainstream MMORPGs and the unique access provided by the game operator. *EQ2* represents the mainstream of the MMORPG market – fantasy role-playing games – which accounts for 85 percent of all MMORPG subscriptions (White, 2008). Because no other operator has released similar data, there is no way to verify the ability to generalize the results of this study across other virtual worlds. We note, however, that most fantasy MMORPGs share a variety of features germane to this particular investigation, including population-centric servers, market economies, crafting, player trades, grouping, questing and the like.

Data collection

After the research team agreed to protect the proprietary aspects of the game's operation, data were supplied on magnetic tape drives by the game operator and were comprised of approximately three terabytes of log data. Each entry was a time-stamped string comprising account information, action taken and location. These strings were converted via a script into a more standard matrix form for database hosting and access. The data for this particular project selected a four-month interval of time for study. This amount of data was at the boundary of the manageable, as large-scale data analysis requires a three-to-one ratio of space to data to allow for indexing and operations. In other words, the three terabytes of log data required a 12 terabyte system for analysis.

Data analysis

Once collected, the data were converted to a readable state within an Oracle database and hosted on a secure supercomputing array at a large-scale academic supercomputing facility operated in conjunction with a major midwestern research university. The data were examined for patterns, most frequently by collapsing entries into aggregates based on periods of time (e.g. the volume all of the currency added to players' inventories for sequential months). Database queries were run with basic Structured Query Language (SQL) commands.

The analysis below tracked all of the economic transactions on two servers, Antonia Bayle and Nagafen, from January to May 2006. This time window allowed for a series of months to be examined for trends and as noted below, also allowed for a unique natural experiment. Between the two servers, there were a total of 314 million transactions in the time window.

There are four major transaction types recorded in the economy log: item production, item consumption, merchant transactions and consignment trade. The first three categories are not true transactions in an economic sense because economic transactions only happen between real people; rather, these are cases of the production and consumption of goods. Nonetheless, these basic categories have large implications for the measurement of GDP. If a player uses the crafting system to create an item, as in category 1, that certainly seems to be production. Yet in the real-world application of GDP, a produced good is not counted until it is sold to someone. If a corn farmer grows and eats some of his/her corn, that corn does not count in GDP, even though it was produced and has value. In practice, GDP is the value of all goods produced and sold in an economy. Thus, an effort to test the mapping of GDP into the virtual requires a focus solely on goods that are sold from one player to another: i.e. genuine trade.

Genuine trade in *EQ2* happens through the consignment system. Consignment trade constitutes 7.17 percent of the total transactions and is similar to broker trading in a real economy. A player can go to a special consignment merchant non-player character and place items with that non-player character. The consignment non-player character then lists the item for sale. Other players may go to any consignment non-player character and browse all available items. A player may buy one item from a consignment non-player character and the money (minus a fee) goes to the selling player. For each consignment trade, there are three records in the economy log: the item in the transaction, the amount of currency spent and the amount of currency received. The difference between the two amounts is the consignment fee.

The information provided by consignment sale records was used to identify transactions according to items traded and money expended. Across all servers, an average of 2 million items are sold each month, for 5 million gold pieces of expenditure (one gold is worth 10,000 copper). Because each type of item has a unique name, the quantity of every type of item sold and total amount of coins spent on that item type for every month on both servers were aggregated. The result was a datafile listing all of the sales of all of the identified items, aggregated by amount sold and total expenditures, by server and month. This datafile was used to test RQ1, H1 and RQ3. In order to answer the basic research question about item categories, the items were examined for descriptors and categorized in accordance with basic economic principles of production.

In order to answer H2, an accurate exchange rate for *EQ2* gold and the US dollar during the time window was sought. The most prominent gold seller in this period was the transactions firm IGE, which posted gold for sale in various lots on various *EQ2* servers throughout the study period. A review of the IGE webpages revealed gold-for-dollar prices that varied slightly depending on the lot size (1000 gold, 500 gold, etc.) and server. On the server Antonia Bayle, for delivery of gold to players on the 'Good' side of the map, the average amount of gold available per dollar was 24.85 from January to March. Delivery to the 'Evil' side of the game could be obtained at a rate of 24.72 gold per dollar. Because this time window overlapped the sampling window, a figure of 25 gold to the dollar was adopted as the basic real-world exchange rate for *EQ2* virtual gold.

H3 and H4, testing applications of the quantity theory of money, required measures of the money supply and the population. The money supply was tracked with two variables in the logs. For every economic transaction, the 'coin_added' and 'coin_removed' variables record the flow of virtual money. When players are putting money into their bank accounts or withdrawing it, there is no net effect on the total money supply. Similarly, when player A buys from player B, the net effect on money supply is zero. However, in every other transaction a change in coin_added or coin_removed meant a genuine increase or decrease in the money supply. With consignments and bank transactions removed, these two variables were aggregated by month. The difference between total coin_added and total coin_removed represents the change in money supply. None of the logs contained information about the amount of money currently in the system. Thus, it was not possible to measure money supply, only the changes within it.

The *EQ2* logs do not supply information on the number of active characters in the game, requiring a rule for counting what characters were 'active'. It was assumed that an active character either had to earn at least one experience point, or to conduct at least one economic activity. In other words, a character had to have at least one record in the experience log or the economy log in order to be considered 'active'. Following this logic, the number of unique active characters was counted for each month for each server, providing a population count.

RQ4 asked whether the mere structure of a virtual world would be enough to predict aggregate economic behavior. As demand for a MMORPG ebbs and flows, it is common for game operators to introduce or remove servers in order to maintain the proper population range per server. Therefore, the introduction of a server represents the opportunity to test whether an empty world will be populated and behave in the aggregate like an existing one. This is a natural experiment that is unlikely to occur in the real world; it would be the equivalent of introducing an empty, exact copy of the physical USA and then populating it to capacity in a month in order to examine whether the infrastructures, laws and systems embedded would generate the same behaviors. In a sense, a new server launch is a new draw from the population of all possible servers, in Monte Carlo style, but done with thousands of real people flowing into an unreal world. For the analysis here, the Antonia Bayle server had been operating since the launch of *EQ2* in February 2004 and so represents the baseline condition. The Nagafen server opened in February 2006 and players were given the opportunity to transfer their characters there to alleviate population pressures elsewhere. Therefore, the test of the research question was whether the new population of Nagafen – presumably made up of no different kinds of people than that of Antonia Bayle – would generate the same GDP and price levels as Antonia Bayle. An affirmative answer to the research question would require the Nagafen GDP and price level to start at zero, but then approach and eventually match Antonia Bayle's.

RESULTS

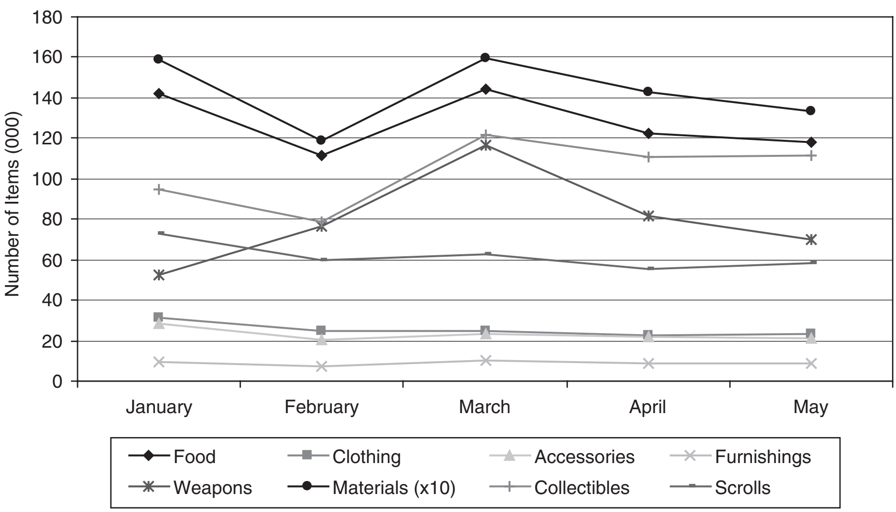
RQ1 and H1 involved the categories and types of goods or items sold in a fantasy economy. As determined by the criteria above, there were eight item categories in the transactions:

- food (e.g. bread, alcohol);
- clothing (e.g. armor, shields);
- accessories (e.g. accessories, bags);
- furnishings (e.g. decorations and furniture);
- weapons (e.g. arrows, swords);
- materials (e.g. leather, herbs);
- collectibles (valuables); and
- scrolls (e.g. books, tomes, recipes).

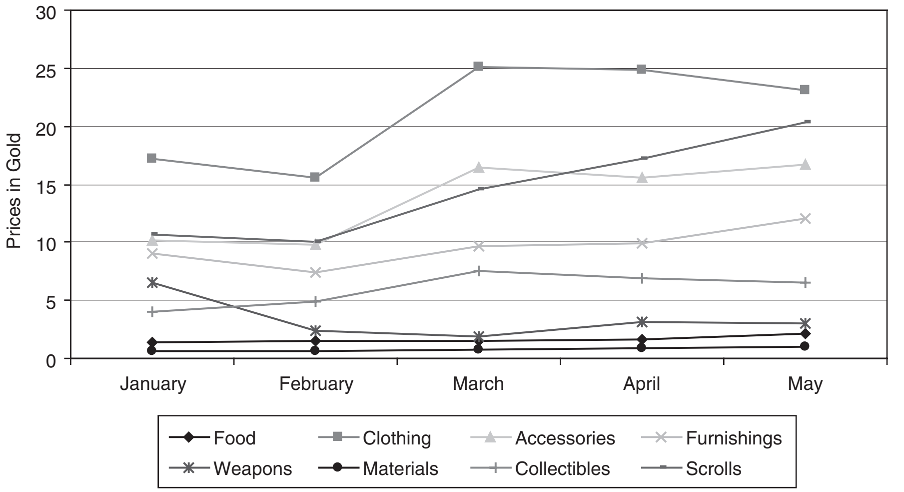
As suggested by H1, these were all real-world descriptors and categories. Figures 1 and 2 show the aggregate sales of these items and the aggregate expenditures for January to May 2006 on Antonia Bayle. These figures are comparable to those in a normal, real-world economy. Thus, one necessary condition for a real-to-virtual aggregate economy mapping has been satisfied.

The next set of questions asserted that normal macroeconomic measures could be calculated in a virtual economy. Using the standard definitions of these measures, the item sales above were aggregated into all-economy data. Figures 3 and Figure 4 show these macroeconomic aggregates.

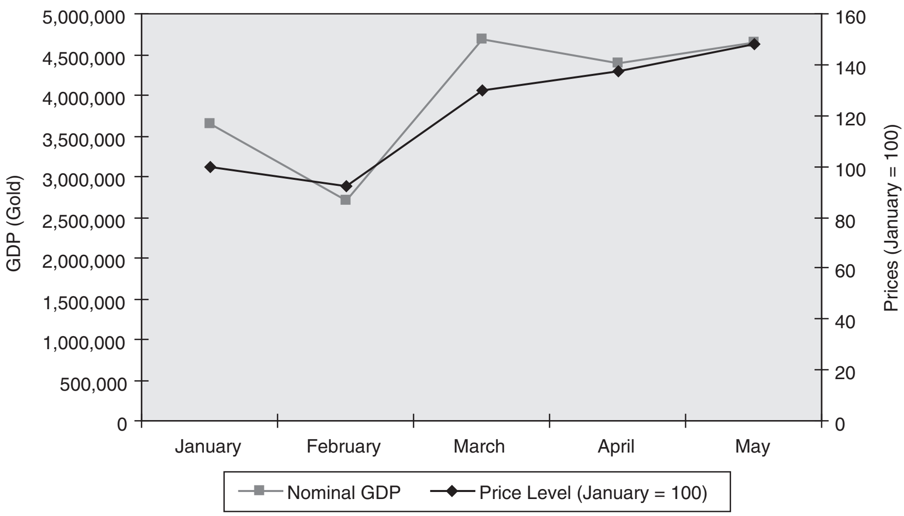
Figure 3 shows how the GDP and the price level varied on Antonia Bayle during the study period. The GDP is the aggregate of all the items sales



• Figure 1 Items sold on Antonia Bayle

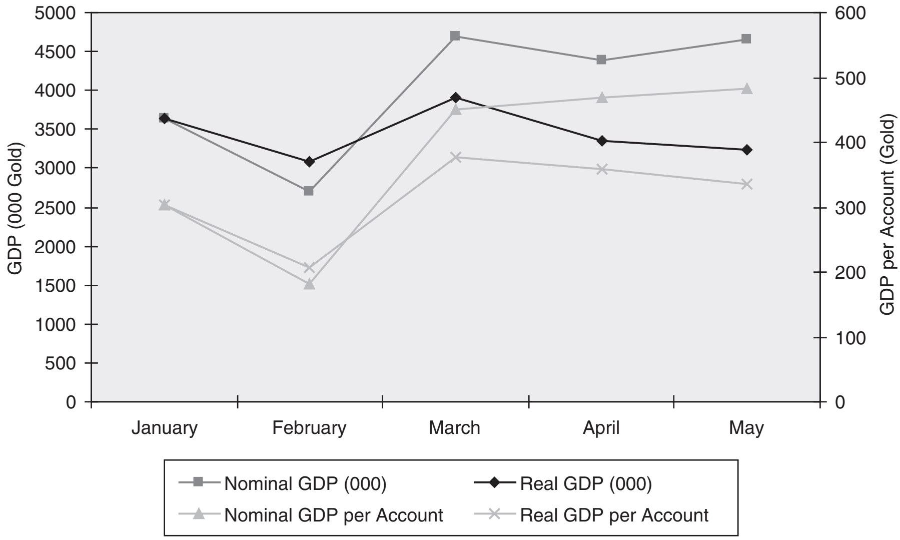


• Figure 2 Item Prices on Antonia Bayle



• Figure 3 GDP and prices on Antonia Bayle

depicted in Figures 1 and 2. It rises from about 3.5 million gold in January to 4.5 million gold in May. Then the items sold in January on Antonia Bayle are taken as the 'market basket' and the price index is calculated as the relative cost of those same items in February, March, April and May. We see that this price level (which by definition starts at 100 in January) rises across the period to 148 in May. In Figure 4, real GDP is calculated as the value of produced items using January prices. Contrasted to nominal GDP, real GDP did not rise from 3.5 million gold to 4.5 million gold; rather it remained at roughly

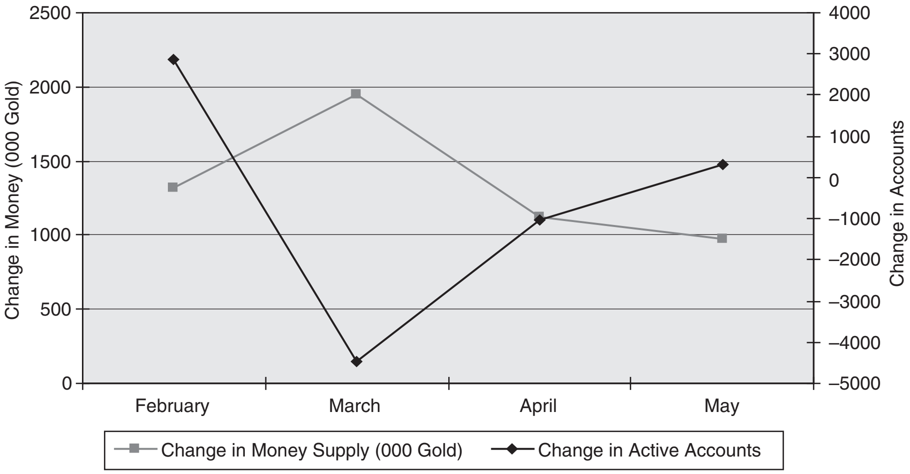


• Figure 4 GDP on Antonia Bayle

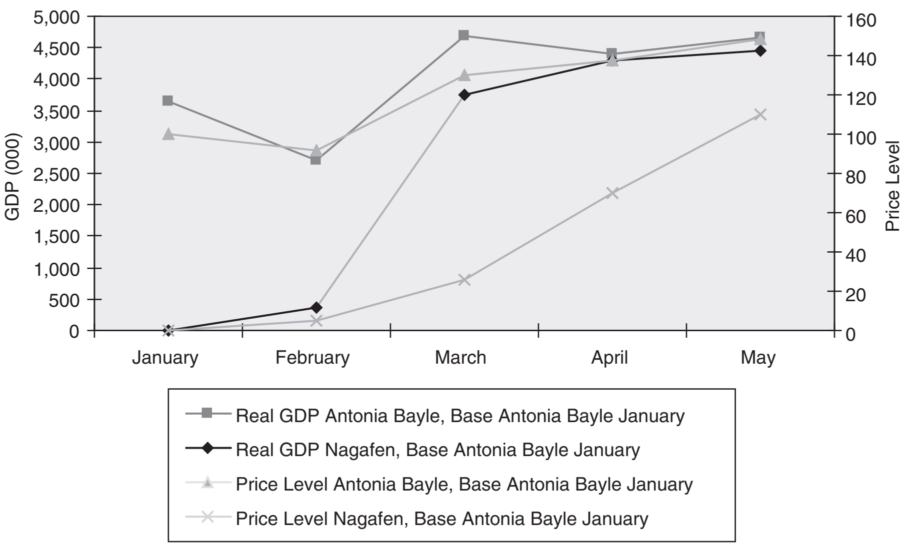
the 3.5 million gold level throughout. These findings are internally consistent: if the nominal GDP is rising and the price level is rising, real GDP should grow less quickly (or not at all) relative to nominal GDP. That is, the increase in nominal GDP was largely an increase in the price at which goods traded, not an increase in the amounts produced and sold.

Figure 5 provides background information about the macroeconomic events recorded in Figures 3 and 4, allowing an examination of the QTM's operation in the virtual economy (H3 and H4). According to the theory, price levels rise only when there is an increase in money supply or a decrease in money demand. Figure 5 shows two statistics necessary to determine whether this happens in the virtual economy: currency flowing in and out of the world, and the number of active accounts on the server (population), which is believed to be closely related to the number of economic transactions. Figure 5 shows how these datastreams, population and the money supply changed month by month from January to May. In March the money supply rose significantly; it also rose in February, April and May, but not as much. Meanwhile, the population rose in February, fell greatly in March, fell a bit more in April and stayed constant in May. These changes suggest that the QTM does operate normally in virtual economies, supporting H3 and H4.

RQ4 examined the power of the world's structure in creating aggregate behaviors. The data from Antonia Bayle show the control condition or the baseline state of an ongoing world, while the new data from Nagafen represent the sudden population of an empty world. As Figure 6 illustrates,



• Figure 5 Change in money supply and population on Antonia Bayle



• Figure 6 GDP and prices, new and mature servers

the Nagafen real GDP and the price levels begin at zero and then quickly approximate the patterns of the baseline world, approaching them both in total size and trend. This suggests that mere structure does have a substantial (nearly total) predictive power on aggregate behavior.

DISCUSSION AND CONCLUSION

This article has proposed an empirical test of whether aggregate economic behavior maps from the real to the virtual. By first creating real-world categories and metrics for consumption and population, the mapping principle was tested through the quantity theory of money. As expected, the existence of categories and relatively stable GDPs presented the necessary conditions for quantity theory of money tests. These tests then revealed that aggregate economic behavior in *EQ2* follows what one would expect to see in the real world. The final test also generated strong evidence that human behaviors in virtual worlds are strongly a product of the system in which they are placed. The mere structure of parallel spaces with parallel rules was enough to offer nearly perfect predictions of these aggregates. This discussion will consider each of the research questions and hypotheses, the theoretical implications of the results and the implications of the research method.

First, it proved simple and intuitive to examine the data and determine the uses of the various items listed for sale. Those uses mapped fairly strongly onto real-world categories: 'Furniture' items were solid, durable things that players kept in their virtual houses; 'Food' items were things that characters ate and drank; 'Weapons' were implements for fighting; 'Armor' items were things that characters wore – 'clothing'. In all, there were no items in the *EQ2* economy which had no apparent real-world analog in terms of their uses and affordances. This was a key condition for true mapping to work: if the real world had only apples and the virtual had only oranges, then the long-term usefulness of mapping would be small, i.e. one would only be able to create predictive models specific to one virtual space, with no possibility of either the virtual or the real predicting one another. For this world, at least, this was not the case.

Second, the indicators of aggregate production and prices followed directly from the item categories. It was possible to calculate GDP and price levels and track their change through time. These changes suggested that while GDP did not go to zero or completely explode, there were fairly abrupt and unusual changes in this figure across time. The fluctuations were more dramatic than are seen normally in a real economy. In the real world, it is not typical for nominal GDP to rise by more than 50 percent in a month, as happened on Antonia Bayle from February to March. Specifically, the data showed that much of this jump in nominal GDP was caused by a rapid inflation of prices. On this server, in just five months there was a 50 percent increase in the price level: that is, 50 percent inflation. What appeared to be a dramatic increase in the amount of real goods that *EQ2* players were producing turned out to be mostly an increase in the prices that these goods were commanding in the marketplace. This is an unusual pattern, if not unheard of (Paarlberg, 1992), but monetary economists do not characterize it as a 'normal' economic situation.

Taken altogether, this is evidence that virtual economies are not perfect analogs of real economies at the aggregate level. They seem to be less stable.

This conclusion is tempered by a number of factors which can be explored in future research. The rapid changes in GDP may not be a function of the fact that the economy is virtual, but because it is being managed in a certain way. Certainly, evidence of rapid changes in the money supply from Figure 5 points out the source of the fluctuations. Thus, the practical application for mapping from virtual to real (or vice versa) might be most useful when comparing the virtual to volatile real-world conditions. After all, the implicit real-world benchmark in this study has been a developed post-industrial economy. Perhaps virtual economies are very precise analog for other kinds of real-world economies, such as frontier, developing or black market economies.

In addition to the mapping question, it was possible to test whether this economy was substantively important in real-money terms, i.e. does virtual currency matter in any real context? On real-world markets such as eBay (www.ebay.com) and IGE, people buy and sell virtual currency in lots of many thousands. Therefore, it was possible to observe an 'exchange rate' between virtual currency and real currency. According to the doctrine of shadow pricing, one may use this kind of exchange rate to put a real-money value on transactions denominated in virtual currency (Castronova, 2001). The exchange rates operating during the study period suggest that the real GDP of Antonia Bayle was approximately between $120,000 and $150,000 per month. In this same period, Antonia Bayle had about 11,000 active accounts. This would imply that the real GDP per account was between $11 and $14. This is a monthly figure; annually, this would come to between and $130 and $164. According to the World Bank (2008), only the poorest countries of the world have a GDP per capita this low, e.g. Liberia ($150), Congo ($140) and Burundi ($110) (World Bank, 2008).

A previous estimate of virtual world per capita GDP was much higher than this: approximately $2000 (Castronova, 2001). What explains the discrepancy? First, the data in this article are unquestionably more accurate. Castronova's first paper relied heavily on small samples taken from publicly available data. This analysis, by contrast, utilizes complete transactions data from internal sources. Yet accuracy alone is probably insufficient to explain such a large difference in numbers.

The main difference is in the conceptual approach to GDP. In this article, the hypotheses demand an attempt to replicate exactly real-world practices in measuring GDP. GDP in the real world is measured as the value of final goods sold in an economy. The data here measure the value of the goods sold by one player to another through *EQ2's* consignment system. However, while this approach is necessary to test questions about mapping, it has a severe conceptual problem: GDP is supposed to be a measure of the aggregate value of production in an economy. In modern, advanced economies, it is true that almost everything produced is also sold. In virtual worlds, by contrast, firsthand reports suggest that a great part of valuable production is consumed directly by the producer (Taylor, 2006).

Thus in the real world one can observe almost all production by observing all the new things that are sold, but applying this approach in a virtual world results in substantial undercounting of the true value of the productive economy. This is a case where the mapping hypothesis is violated not by the behavior of virtual world users per se, but by the inapplicability of a real-world observational practice to that behavior. In future research, it might be possible to modify real-world economic data-gathering so as to better measure the creation of value inside virtual worlds.

An additional reason that might be advanced to explain why some of the production in the virtual world was not captured through the methods utilized in this study, is that some player-to-player trading happens directly, face-to-virtual-face as it were. However, the Sony Online Entertainment data indicate that only 0.66 percent of all transactions involved this type of activity. Because macroeconomic statistics are based on transactions involving currency rather than item-for-item transactions, this small number of barter trades was excluded from the study, as were gifts. Nonetheless, in some environments the gift and barter economy may make up a substantial portion of trade, and future research should remain open to measuring this. Given that MMORPG players often develop strong social relationships with one another (Steinkuehler and Williams, 2006), there is reason to believe that sometimes direct trading may play a strong role, if not here. Regardless, the present article's estimate of GDP is undoubtedly more accurate than an arm's-length study might be, and thus may be thought of as an accurate lower bound on the true value. The estimates in this article are conservative, meaning that the value of virtual world production is at least $130 per person per year and probably much more. The fact that this study even raises the question of how to handle goods consumed by their producer is another mild indicator of mapping; real-world statisticians must wrestle with this problem from time to time.

Furthermore, in the real world there is much debate about the role of structure in determining human outcomes. On the one hand, the fact that two different servers seemed to converge on the same macroeconomic conditions may be evidence in this regard. On the other hand, macro-level similarities may not be caused by the game's code – it is 'law' in Lessig's formulation – but rather by deeper similarities in human behavior. Or perhaps the result was caused neither by the passive operation of code nor the passive adaptation of humans, but rather as an active policy of directed management by the game owners. To date there have been no studies of the management of economies from inside games. It seems likely that game companies use a mix of active and passive management: part code, part sweat. Discovering the cause of similarities across shards is an avenue for future work.

These data also provided an opportunity to replicate and extend predictions made by the quantity theory of money in the case of virtual worlds. This theory received strong support, as evidenced by the movements of GDP and prices. Consider the data for March 2006: there was a sharp rise in the money supply accompanied by a sharp decline in the population. The population may be taken as a rough indicator of the demand for economic transactions and hence the demand for money; therefore, March can be seen as a time when money supply rose and money demand fell. The quantity theory of money, as summarized in H3 and H4, states that such conditions ought to cause large increases in the price level, which happened right away. Referring back to Figure 3, the price level in March rose to 130 from 92 in the month before. That is an inflation rate of 41 percent, the largest in the time period under study. Thus, while 41 percent is a very large inflation rate, indeed one so large as to raise a question about the strength of economic mapping, it is actually a very standard response to changes in the fundamentals which themselves happen to be large. The GDP and the price level are actually robust economic indicators: they responded normally to volatile population changes and an unstable monetary situation.

This result suggests that macroeconomic mapping does occur, despite the rapid change observed in GDP and the price level. In other words, although the volatility was atypical, the reactions were anything but. This adds to the robustness of the QTM at the same time as it suggests that virtual worlds are possible testbeds for mapped behaviors – in volatile situations, but possibly also for more stable ones. This suggests an intriguing avenue for future research. Just as simulations offer a low-cost means of testing human behaviors, reactions and interactions, so too might virtual spaces offer a low-cost means of testing aggregate behaviors. While policymakers wrestle with the expected and unexpected consequences of their choices, a virtual-world test of some policies could be undertaken with substantially lower costs. To name but one example, if a central bank suggests freezing all assets in a country but is unsure what the impact on prices will be, it would be substantially less risky to test the proposition in a virtual economy.

In sum, this article concludes that the conditions may be virtual, but the players in them are very real and apparently quite rational in the aggregate – at least in their economics. As noted earlier, the uniqueness of these data makes it impossible to demonstrate their ability to be generalized. Given that most virtual worlds offer the same basic structures and operating mechanics, the authors do not suspect that there would be large differences; however, this remains untested. Certainly, more nuanced investigations will be more subject to the small differences that may exist between worlds. Future research operating in the paradigm of large-scale, virtual-world data analysis can replicate and test these assumptions. We suggest that a robust framework and theory for testing the mapping principle could help to guide researchers across many domains in testing human behaviors. There are clearly validity hurdles to be cleared and best practices to be developed, yet the potential for low-cost, low-risk and high-impact work is evident.

Notes

1 Introductory texts in economics typically state that increases in GDP and stability of the price level are the primary objectives of macroeconomic policy. For example Hall and Lieberman (2003) begin Chapter 4 ('What Macroeconomics Tries to Explain') by saying that economists and society at large agree on the basic macroeconomic goals. The two foremost of these goals are 'rapid economic growth' and 'stable prices'. They then introduce GDP changes as the proper measure of growth and changes in the price level as the proper index of price stability.

2 Throughout the article the focus is on the internal economy of *EQ2*. Any fees paid by players to the game company are not part of this internal economy, thus they do not play a role in the definition of GDP. If the focus was on the incomes of players, then these fees would be taken into account.

3 A student at the University of Helsinki (Lehdonvirta, 2005) has developed a series of macroeconomic indicators explicitly for use in virtual economies. Our approach differs in that we are not attempting to modify existing indicators and principles so that they fit virtual economies. Rather, we are taking real-world protocols as they are and attempting to apply them, in order to see how well they work.

References

Bailenson, J., A. Beall, J. Blascovich, J. Loomis and M. Turk (2005) 'Transformed Social Interaction, Augmented Gaze and Social Influence in Immersive Virtual Environments', *Human Communication Research* 31(4): 511–37.

Bomfim, A. (2001) 'Measurement Error in General Equilibrium: The Aggregate Effects of Noisy Economic Indicators', *Journal of Monetary Economics* 48(3): 585–603.

Brown, C., R. Solow, A. Ando and J. Kareken (1963) 'Lags in Fiscal and Monetary Policy', in C. Brown (ed.) *Stabilization Policies: Commission on Money and Credit*, pp. 1–14. Englewood Cliffs, NJ: Prentice-Hall.

Castronova, E. (2001) 'The Rise and Fall of a Policy Rule: Monetarism at the St. Louis Fed, 1968–86', *Federal Reserve Bank of St. Louis Review* January/February: 1–24.

Castronova, E., M.W. Bell, M. Carlton, R. Cornell, J.J. Cummings, W. Emigh, M. Falk, M. Fatten, P. LaFourest, N. Mishler, J. Reynard, S. Robbins, T. Ross, W. Ryan and R. Starks (2008) 'A Test of the Law of Demand in a Virtual World: Exploring the Petri Dish Approach to Social Science', CESifo Working Paper No. 2355, Social Science Research Network.

Cook, T.D. and D.T. Campbell (1979) *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin.

Cox, J. (1997) 'On Testing the Utility Hypothesis', *Economic Journal* 107(443): 1054–78.

Cunningham, T. (1992) 'Some Real Evidence on the Real Bills Doctrine versus the Quantity Theory', *Economic Inquiry* 30(2): 371–83.

Dibbell, J. (2006) *Play Money*. New York: Basic Books.

Fixler, D. and B. Grimm (2008) 'The Reliability of the GDP and GDI Estimates', *Survey of Current Business*, February, URL (consulted 29 March 2009): http://www.bea.gov/scb/pdf/2008/02%20February/0208_reliable.pdf

Hafer, R. and D. Wheelock (2001) 'The Rise and Fall of a Policy Rule: Monetarism at the St. Louis Fed, 1968–1986', *Review* 83(1): 1–24.

Hall, R. and M. Lieberman (2003) *Macroeconomics: Principles and Applications* (2nd edn). Mason, OH: Thomson Learning.

Jankowski, N. (2007) 'Exploring e-Science: An Introduction', *Journal of Computer-Mediated Communication* 12(2): 549–62.

Kazdin, A. (1977) *The Token Economy: A Review and Evaluation*. New York: Plenum.

Landefeld, J., E. Seskin and B. Fraumeni (2008) 'Taking the Pulse of the Economy: Measuring GDP', *Journal of Economic Perspectives* 22(2): 193–216.

Lankenau, S. (2001) 'Smoke 'Em if You Got 'Em: Cigarette Black Markets in US Prisons and Jails', *The Prison Journal* 81(2): 142–61.

Lehdonvirta, V. (2005) 'Virtual Economics: Applying Economics to the Study of Game Worlds', paper presented at the 2005 Conference on Future Play (Future Play 2005), Michigan State University, Lansing, MI.

Lessig, L. (1999) *Code and Other Laws of Cyberspace*. New York: Basic Books.

Nash, J. and E. Schneyer (2004) *Virtual Economies: An In-depth Look at the Virtual World of Final Fantasy XI: Online*. Philadelphia, PA: Wharton School of Business.

Paarlberg, D. (1992) *An Analysis and History of Inflation*. New York: Praeger.

Radford, A. (1945) 'The Economic Organization in a POW Camp', *Economica* 12(48): 189–201.

Samuelson, P. and R. Solow (1960) 'Analytical Aspects of Anti-inflation Policy', *Proceedings of the American Economic Review* 50(2): 177–94.

Sargeant, T. and N. Wallace (1981) *The Real Bills Doctrine vs. the Quantity Theory: A Reconsideration*. Minneapolis, MN: Federal Reserve Bank of Minneapolis.

Simpson, Z. (1999) 'The In-game Economics of Ultima Online', paper presented at the Computer Game Developer's Conference, San Jose, CA, 8–12 March.

Steinkuehler, C. and D. Williams (2006) 'Where Everybody Knows Your (Screen) Name: Online Games as "Third Places"', *Journal of Computer-Mediated Communication* 11(4), URL (consulted 29 March 2009): http://jcmc.indiana.edu/vol11/issue4/steinkuehler.html

Taylor, T.L. (2006) *Play between Worlds: Exploring Online Game Culture*. Cambridge, MA: MIT Press.

US Department of Commerce (1998) 'Updated Summary NIPA Methodologies', *Survey of Current Business*, September, URL (consulted 29 March 2009): http://www.bea.gov/scb/account_articles/national/0998niw/maintext.htm

Webb, E., D. Campbell, R. Schwartz and L. Sechrest (1966) *Unobtrusive Measures: Non-reactive Research in the Social Sciences*. Chicago, IL: Rand McNally and Company.

White, P. (2008) 'MMOGData: Charts', URL (consulted 29 March 2009): http://mmogdata.voig.com/Charts.html

Williams, D., N. Yee and S. Caplan (2008) 'Who Plays, How Much and Why? A Behavioral Player Census of a Virtual World', *Journal of Computer-Mediated Communication* 13(4): 993–1018.

World Bank (2008) 'Gross National Income Per Capita 2007: Atlas Method and PPP', in *World Development Indicators Database*, URL (consulted 29 March 2009): http://siteresources.worldbank.org/DATASTATISTICS/Resources/GNIPC.pdf

Yee, N. and J. Bailenson (2007) 'The Proteus Effect: Self Transformations in Virtual Reality', *Human Communication Research* 33(3): 271–90.

EDWARD CASTRONOVA is an associate professor of telecommunications at Indiana University. His research involves virtual worlds and society.
Address: Department of Telecommunications, Indiana University, 1229 E. 7th Street, Bloomington, IN 47405, USA. [email: castro@indiana.edu]

DMITRI WILLIAMS is an assistant professor at the Annenberg School for Communication, University of Southern California. His research interests include community and networked games.
Address: Annenberg School for Communication, University of Southern California, 734 W. Adams Boulevard, Los Angeles, CA 90089, USA. [email: dcwillia@usc.edu]

CUIHUA SHEN is a doctoral candidate in Annenberg School for Communication, University of Southern California. Her research focuses on virtual communities that are created for information sharing, collaborative production and entertainment.
Address: Annenberg School for Communication, University of Southern California, 3502 Watt Way, Los Angeles, CA 90089-0281, USA. [email: cuihuash@usc.edu]

RABINDRA RATAN is a doctoral candidate at the Annenberg School for Communication, University of Southern California. His research focuses on self-presence, avatar presentation, serious games and other psychological and social facets of virtual worlds and video games. [email: raratan@gmail.com]

LI XIONG is a doctoral candidate in the Annenberg School for Communication at the University of Southern California. His research focuses on the sociotechnical interactions in the engineering and use of new communication technologies. [email: lixiong@usc.edu]

YUN HUANG is a research scientist and postdoctoral research fellow in the Science of Networks in Communities (SONIC) research group, Department of Industrial Engineering and Management Sciences, Northwestern University. His research explores the dynamic and evolution of individual behavior and interactions in digital-enabled environments.
Address: Department of Industrial Engineering and Management Sciences, Northwestern University, 2145 Sheridan RD, TECH C210, Evanston, IL 60208, USA.
[email: yun@northwestern.edu]

BRIAN KEEGAN is a doctoral candidate in the Media, Technology and Society Program at Northwestern University. His research interests include online deviance, group norms and computer-mediated decision-making.
Address: Northwestern University School of Communication, Room 2-168, 2240 Campus Drive, Evanston, IL 60208 [email: bkeegan@gmail.com]