A Framework for the Analysis of Counterfeiting Behaviour

Curtis Wright
5155259
Econ 898

October 26, 2010

1 Introduction

The incentive to counterfeit money has existed wherever and whenever money has been used, but modern technology has allowed for a new generation of currency counterfeiting to emerge. While technological and institutional changes have brought about popular new ways of completing transactions without money, such as debit and credit cards, cash is still a very popular option, especially for smaller transactions. Estimates in the RCMP Criminal Intelligence report Counterfeit Currency in Canada (2007) suggest that 63% of all transactions under $25 in Canada use cash as a medium of exchange, while debit cards are the second most popular option, being used 21% of the time. For larger purchases, in the $25 to $100 range, debit cards are the most popular option, being used for 39% of transactions while cash is the second most popular choice, accounting for 28% of transactions (RCMP Criminal Intelligence, 2007). As long as cash remains such a popular mechanism of financial intermediation, the threat of counterfeiting will exist. Counterfeiting behaviour, both domestically and abroad, is not restricted to the creation of fake currency: goods and intellectual property have increasingly become the target of modern criminals. The spread of counterfeiting has led to suggestions that counterfeiting is "the world's fastest growing crime wave" (Phillips, 2005). Since money is such a popular tool for exchange, it is important to analyze the economics of counterfeiting behaviour and identify the effects of counterfeiting in a monetary economy.

The RCMP's (2007) empirical analysis of counterfeiting behaviour indicates that the use of fake currency peaked in 2004, when the total face value of counterfeit money either passed or seized in Canada was 13 million dollars. Passed money refers to fake money that is successfully used in a transaction, while seized money is fake currency that is identified as such and confiscated. By 2006, the total face value of counterfeit money passed and seized in Canada had fallen to 4 million dollars (RCMP Criminal Intelligence, 2007). The rapid decrease has been attributed to the introduction of new bank notes, which contain several features designed to deter counterfeiting. These features include a holographic stripe, watermark portrait, windowed colour shifting thread, and a see-through number.

There are a number of reasons to be concerned about the counterfeiting of currency. The most obvious issues are the costs to sellers who have accepted the fake bills, and the potential for reduced confidence in a country's currency. There are, however, other concerns associated with counterfeiting activities. The RCMP Criminal Intelligence reports further suggest that there are links between counterfeiting behaviour and other forms of fraud, such as identity fraud. The skill set required to perform counterfeiting would presumably lend itself to the forging of documents required for other types of fraudulent activities. The report also suggests that there are established ties between counterfeiters and organized crime. This is likely the result of the significant capital, both human and financial, required to produce the highest quality counterfeit bills. The report also indicates that counterfeiting behaviour is linked to terrorist financing, which would be another reason to take the threat of counterfeiting seriously. It is worth noting that since currency travels across borders and can be held in foreign countries, counterfeit money may not be produced domestically. In the 1970's and 1980's, Canada was even considered to be a hub for the counterfeiting of American currency (RCMP Criminal Intelligence, 2007). The US Treasury Department's report The Use and Counterfeiting of United States Currency Abroad (2007) suggest that the ratio of counterfeit American bills to real bills is roughly the same abroad as it is domestically.

Technological changes have contributed significantly to the feasibility of establishing counterfeiting operations (RCMP Criminal Intelligence, 2007). The two dimensions of counterfeiting quality are visual appearance and tactile properties. The increased availability and quality of inkjet printers and scanners have contributed to the improvement of visual appearance of counterfeit bills, and future technological improvements threaten to continue this trend. While creating an accurate visual reproduction has traditionally been the primary goal of counterfeiters, there has been a recent trend towards reproducing the tactile properties of Canadian currency. Specialty inks and foils have become available that allow counterfeiters to produce bills that not only look the part, but have many of the same non-visual properties of a genuine Canadian bill. Judson and Porter (2010) suggest that the quality of the non-visual properties of a counterfeit note are generally most important to the success of passing fake bills, since cashiers and tellers tend to rely on the physical feel of money to identify fake notes. While the $20 bill has traditionally been the most popular denomination for counterfeiting, another recent trend has been the production of high quality one and two dollar coins in Canada. This suggests that the style of counterfeiting in Canada may be changing as a result of improved counterfeit deterrence measures on bills (RCMP Criminal Intelligence, 2007).

The model presented in this paper provides a theoretical framework for the analysis of currency counterfeiting behaviour. By using agents with a distribution of productivities, and a directed search environment with a distribution of terms of trade, the model is able to analyze the questions of who counterfeits, where they spend their fake money, and how it affects firms. The model is designed to mimic the real world where people who counterfeit have to decide whether to "dump" the fake money on small inexpensive purchases or attempt to make larger, more risky purchases. It also determines who counterfeits based on the relative costs and benefits of labour market participation and counterfeiting.

2 Literature Review

The economic analysis of criminal behaviour was broached by Gary Becker in 1968 with the publication of Crime and Punishment: An Economic Approach. This was a theoretical formalization of the costs imposed by illegal activities, and the policies used to prevent it. Becker focused on the derivation of optimal policy, both public and private, to maximize social welfare in a model with criminal behaviour. This required balancing the trade-off between the costs to society of criminal behaviour; and the costs of apprehension, prosecution, and punishment. While the model presented in this paper is not specifically designed for the analysis of optimal legal policy, the formalization of an incentive structure for rational criminal behaviour is inspired by Becker's famous work.

This paper provides a framework for the analysis of counterfeiting behaviour, and explains both who counterfeits and how they use their counterfeit money. The choices associated with engaging in illegal behaviour are discussed by Stigler (1970), who identifies the location and scale of crimes as variables that a criminal can control to maximize income expectations. These choices are the reflected in the model presented in this paper. Modelling the scale of a crime is relatively straightforward in the context of counterfeiting, but the choice of location is modelled somewhat indirectly. Rather than considering location in a geographic sense, the following model offers agents the choice of markets to enter, using a directed search framework. The directed search is used by Menzio, Shi, and Sun (2009) to build a monetary framework with non-degenerate distributions. The difference between the directed search used in their model and the standard search models is that buyers know the terms of trade prior to entering a market, rather than negotiating them after an appropriate match is made. In the model presented, the distribution of money holdings is determined by a combination of exogenous abilities, labour market decisions, and counterfeit decisions. The directed search structure borrowed from Menzio et al. approximates the real world options available to counterfeiters, where they can choose to make small, relatively easy purchases, or large, more difficult purchases.

The counterfeiting of money itself has been modelled theoretically in various ways for different purposes. Green and Weber (1996) present a monetary search model with counterfeiting and two types of money, one of which can be counterfeited and one of which cannot. They use this model to predict the effect of the introduction of new $100 dollar bills into the United States' currency pool. Green and Weber's model incorporates old money, which can be counterfeited and new money, which cannot be counterfeited. The search model includes government agents whose specific purpose is to detect and confiscate counterfeit money. The search model used by Green and Weber features identical, infinitely lived agents who produce unique goods and seek to consume the goods produced by other agents. Transactions are initiated by random matching with other agents. The government agents are exogenously introduced into the search environment, and serve two tasks, both to confiscate counterfeit money and to replace old, counterfeitable money with new, non-counterfeitable money. This design is meant to mimic the real world monetary environment of the United States, where old $100 dollar bills are still accepted, and only removed from circulation over time. The detection mechanism is different from the detection mechanism of this paper, where there is an exogenous detection technology at the point of transaction, but it only successfully detects counterfeit money a fraction of the time. The counterfeiting technology in Green and Weber's model is used when a trader has his money confiscated by a government agent. The agent can then choose whether or not to pay an exogenous cost in order to replace the indivisible unit of counterfeit money and continue trading.

Green and Weber's model predicts that there are three potential equilibria depending on model parameters. The first equilibrium has counterfeiting behaviour remaining at a level consistent with the old-money equilibrium. The introduction of new money has no effect on counterfeiting behaviour. The second equilibrium has both old and new style money immediately go out of circulation when new money is introduced. This is a result of the repetitive nature of the model, and agents only accepting money that they think will be accepted in the future. The third equilibrium has both genuine and fake old money slowly growing out of circulation over time. The speed at which old money is removed from circulation depends on the aggressiveness of government policy for deploying government agents. The results of this model are interesting because they provide a monetary search microfoundation for counterfeiting behaviour. Green and Weber's paper provides interesting insight into the potential for theoretical modeling of counterfeiting as a tool for policy analysis. The two policy mechanisms introduced for the deterrence of counterfeiting (new money and government agents) are both based on real world counterfeit deterrence strategies, and it is an interesting reconciliation of monetary search theory and real world policy. While the model presented below is set up significantly differently from Green and Weber's model, they both provide insight into the equilibrium analysis of counterfeiting. The most interesting difference between the model below and Green and Weber's approach may be the mechanism for counterfeit detection. The use of government agents in a search environment is a novel and realistic way of modelling the detection of fake money. The following model relies solely on point of purchase detection technology, and it may be most appropriate, though significantly more difficult, to consider some combination of the two detection mechanisms.

Williamson (2002) considers counterfeiting behaviour in a completely different context, examining the effect of counterfeiting on the relative costs and benefits of public and private money. Williamson suggests that one of the costs of private money is that it is relatively easy to counterfeit, which should be considered an additional cost to the use of private money.

To analyze the impact of counterfeiting in a monetary economy with fiat and private money, Williamson considers an economy with two sectors: a search sector and a banking sector. Agents are randomly drawn into one of those two sectors each period. An agent in the banking sector can obtain a private bank note in exchange for goods, and the private bank note is backed by an investment, so it can be redeemed for goods at any time. A fraction of agents are initially endowed with a unit of fiat money, which is intrinsically worthless. The search sector is a simple monetary search model, where agents can produce, trade for, and consume goods, subject to a preference shock that determines whether or not they want to consume in any given period. To incorporate the possibility of counterfeiting, Williamson assumes that private money can be counterfeited at a cost, but that fiat money has an infinite cost to counterfeit.

Williamson finds that when counterfeiting is not included in the model, privately backed money provides a benefit to social welfare through financial intermediation. When counterfeiting is introduced, the model suggests that this threat can lead to the collapse of monetary exchange. This suggests that the threat of counterfeiting can, in theory, necessitate the prohibition of private money. While the analysis of private versus fiat money is certainly not a new issue, including the threat of counterfeiting is a novel twist that produces interesting results, and reinforces the importance of considering the threat of counterfeiting in monetary models.

Studying the effects of currency counterfeiting is generally performed in a search theoretic monetary model in order to provide a microfoundation for money. This concept has been advanced by several authors, including Kultti (1996), Nosal and Wallace (2007), Li and Rocheteau (2008), and Quercioli and Smith (2009).

Kultti (1996) used a search model with counterfeiting and fiat money to investigate welfare effects of counterfeiting. Kultti built a model where during each period agents can choose to either produce goods or produce counterfeit money, both at a cost. There is some probability that agents holding counterfeit money will be caught and punished. The degree of punishment is determined by the government. Agents trade in a random match market with no barter and do not consume their own production goods. Fiat money and counterfeit money are used as mediums of exchange, and agents can verify the nature of their money holdings.

The equilibrium conditions of Kultti's model yield several interesting results. Kultti finds that there can exist an equilibrium in which both real and fake money circulate in the economy. This is consistent with empirical observations of counterfeiting behaviour, but the existence of said equilibrium depends on certain conditions. For fake money to circulate in the economy, there must be some positive probability of counterfeit money holders going undetected. If the holders of counterfeit money have no opportunity to avoid punishment then the fake money will not be accepted and will not circulate. The degree to which counterfeit holders are punished also affects the circulation of counterfeits. Low punishment will offset high probability of apprehension, and encourage the circulation of fake money. Circulation of counterfeits in Kultti's model is also dependant on the time preference of agents. The demand for counterfeit money in Kultti's model is the result of agents being relatively impatient. If the time preference of agents is low enough, agents will patiently wait for real money rather than take the risk of obtaining fake money. The welfare effects of counterfeiting in Kultti's model are also determined by several variables. If the amount of genuine money circulating in the economy is low enough, counterfeit money can improve the efficiency of the exchange economy. This however requires punishment be appropriate to prevent the flooding of the economy with fake money.

Kultti's model focuses on the introduction of counterfeiting behaviour to a search theoretic economy, isolating the conditions required for a counterfeiting equilibrium, and the welfare effects of this behaviour. The model to be presented has a similar concept of agents choosing to counterfeit or obtain real money in preparation for future transactions. The exogenous probability of detection and nature of punishment is also similar, but the following model focuses on the result of future transactions rather than the welfare effects of financial intermediation via counterfeit money.

Nosal and Wallace (2007) use a monetary matching model with imperfect information to analyze the effects of counterfeiting in a simple exchange economy. They present a model in which agents can choose to counterfeit currency at a cost, or produce a specialized good for trade. The counterfeit money in their model is not subject to the threat of being caught by law enforcement, but all counterfeit money disappears after one period. There is a fixed stock of real money, which is distributed among some fraction of the total agents, as well as any counterfeits produced during that period. In each period agents are randomly paired, and when a single coincidence of wants occurs between a money holder (genuine or fake) and a seller, a signal is provided to both agents. With some probability, the signal reveals the true nature of the money to both agents, otherwise it is uninformative. The buyer can then make an offer to the seller, and if accepted, the transaction occurs.

The result of Nosal and Wallace's model is that in there is no steady state equilibrium with counterfeiting. If the cost of counterfeiting is low enough, then there is no monetary equilibrium. This is the result of the model's assumptions, specifically that counterfeit money disappears after one period. Another reconciliation of the model's predictions and empirical evidence offered is based on the observation that counterfeiting behaviour is often sporadic, and sometimes has led to reduced trade of the currency in being counterfeited. This would suggest that real life counterfeiting would not necessarily be best thought of as part of a steady-state monetary equilibrium. The model also has interesting policy implications, since increasing the cost of counterfeiting or the probability of recognizing counterfeits will increase the set of parameters in which a monetary equilibrium exists. This implies that standard counterfeit-deterrence strategies will be beneficial.

The key differences between the model built by Nosal and Wallace and the model to be presented are the structure of the market, the number of periods, and the nature of the agents. Nosal and Wallace use a random matching market structure in each period to establish a monetary framework, while this paper uses alternating centralized and directed search markets. Nosal and Wallace also use agents that are identical and live across many periods, while this model uses agents with different abilities that only live for two periods. The two models arrive at very different conclusions as a result of differences to the model environment. Nosal and Wallace use a monetary framework for the analysis of counterfeiting, and predict that there is no stable steady state counterfeiting equilibrium, while the model presented in this paper is not a monetary framework per se, but a simple environment that allows for the analysis of counterfeiting behaviour in equilibrium.

The lack of a monetary equilibrium with counterfeiting in a search-theoretic exchange economy is addressed by Li and Rocheteau (2008). They use a model environment that is very similar to that of Nosal and Wallace (2007), but find different results. One key deviation from Nosal and Wallace's model made by Li and Rocheteau is the assumption that all counterfeit money is confiscated at the end of each period. Li and Rocheteau assume that some fraction of total counterfeit money is confiscated at the end of each period, but not all. This assumption leads to each period beginning with some agents holding real money, some holding fake money, and some without money. The other key assumption changed is allowing sellers to set terms of trade in some matches, as opposed to buyers making take it or leave it offers in Nosal and Wallace's model.

By changing either of these assumptions, Li and Rocheteau are able to build a model where the existence of counterfeiting behaviour does not eliminate the possibility of a monetary equilibrium. Li and Rocheteau find that welfare and the value of real money are affected by the threat of counterfeiting, regardless of whether or not counterfeiting actually occurs in equilibrium. This allows for the analysis of policy implications and suggests that counterfeit deterrence efforts may increase welfare regardless of the presence of counterfeiting. The difference between the equilibrium outcomes of Nosal and Wallace's model compared to Li and Rocheteau's model illustrates the effect of subtle changes to the model's assumptions, and identifies some of the key assumptions required to establish a monetary equilibrium with the threat of counterfeiting.

While the model set up of Li and Rocheteau is very similar to that of Nosal and Wallace, the resulting equilibria provide the framework for a monetary equilibrium with counterfeiting. The model to be presented also attempts to provide a theoretical framework for the analysis of counterfeiting behaviour, but uses a model that is built very differently. Li and Rocheteau's model is able to provide a monetary framework with counterfeiting, or the threat of counterfeiting, in equilibrium. Their model allows for the analysis of counterfeiting's effect on general welfare, while the model presented in this paper attempts to identify the details of illegal behaviour, including who counterfeits, and where they spend their fake money.

Another explanation of counterfeiting behaviour is presented by Quercioli and Smith (2009), who build a strategic model consisting of "good guys" and "bad guys". The bad guys are counterfeiters, and good guys are agents who wish to transact. Good guys choose how much effort to spend screening bills to monitor for counterfeits. More effort yields better screening, but if fakes go undetected, they can be passed on anonymously through a random matching exchange economy. Bad guys choose whether or not to counterfeit, the value of bills to produce, and what quality of bills to produce. Better quality bills are more difficult to verify and therefore more likely to pass. The model has two games being played out, the first consists of bad guys choosing whether or not to enter at the cost of legal repercussion, and choosing what quality of counterfeits to produce. The second game consists of good guys in a random matching exchange economy unknowingly passing on counterfeit goods or money to other good guys, while choosing how much effort to dedicate to verifying incoming goods or money.

Quercioli and Smith's (2009) model is characterized by a Nash equilibrium that identifies the quality of counterfeits produced, the effort dedicated to verification, and the quantity of counterfeits produced. There exists a minimum value of notes or goods at which counterfeiting will take place, since the value must be great enough to cover expected legal costs. The model predicts interesting relationships between the value of the note or good being counterfeited, and the quality and verification effort. Near the lowest value counterfeited, Quercioli and Smith find that the verification cost and quality of counterfeits disappear. As the value increases, counterfeiters have more to gain, and counterfeiting quality increases. Verification effort also increases, preventing counterfeiting from being any more profitable at higher values then at lower values. This means that the fraction of total counterfeits passing is lower at higher values of currency or goods.

The model to be presented is significantly different from the model constructed by Quercioli and Smith, because the agents in Quercioli and Smith's model are exogenously appointed as either good guys or bad guys. The model to be presented has a distribution of agents that can choose whether or not to engage in counterfeiting behaviour based on the idiosyncratic costs and benefits of work and counterfeiting. There are, however, multiple elements of Quercioli and Smith's model that could be incorporated into a model similar to the model to be presented. The first is the counterfeiter's choice of quality. This is a decision that affects the probability of a note being passed, and is a more realistic choice than the simple counterfeiting technology presented in this paper, which produces identical quality fakes for all counterfeiters. In the same spirit, the endogenous choice of verification effort by non-counterfeiters would provide a realistic choice variable for the sellers in this model, who only have access to an exogenously supplied verification technology. Finally, the fact that Quercioli and Smith's model extends over multiple periods leads to the "hot potato" game of passing on counterfeit money. This cannot happen in the model presented in this paper because the world ends, but if the model was extended, it is a realistic outcome of a counterfeiting equilibrium.

3 Model Environment

The model consists of a measure of agents, normalized to 1, living in a discrete time world. Each agent is born, lives for two periods, then dies. In the first period agents do not consume, but they can accumulate money and counterfeiting capabilities. In the second period, agents are retired, and cannot work. They can however, consume using money, either real or fake, to purchase goods.

3.1 First Period

In the first period agents are endowed with one unit of time, which they can divide between two tasks: working and counterfeiting. If an agent chooses to work they enter a Walrasian market and supply l_i units of labour, where $0 \leq e \leq 1$. Production by agents in the Walrasian market is sold to the rest of the world, so there is no consumption in period 1. Any positive profits are redistributed to employees. Agents in the labour market are paid in real money, but agents who supply labour incur a disutility from working of $h\left(l\right)$. $h(l)$ is twice differentiable and has the following properties: $h'(l)0 > 0$, $h''(l) > 0$,$h(0) = 0$, and $h'(1) = \infty$. Agents can also choose to counterfeit, which will incur an exogenous fixed cost c as the agent must pay for the setup of equipment. Once the fixed cost has been paid, the agent owns a counterfeiting operation. This counterfeiting operation will be able carried forward into the second period, allowing the agent to produce any amount of counterfeit money with zero marginal cost. Real money is also carried without cost and interest free into the second period.

3.2 Second Period

In the second period, agents obtain utility by consuming goods, facing a utility function $u\left(c\right)$. $U(c)$ is twice differentiable and has the following properties: $U'(c) > 0$, $U''(c) < 0$, and $U'(1) =$ a large number. Consumption goods are purchased from the rest of the world, here is only one type of good, and it is perfectly divisible. There is no centralized or clearing market for goods in the second period, so buyers and sellers must seek each other out in order to perform a transaction. The model uses a directed search structure, where terms of trade are established prior to agents engaging in the search process. Rather than a single set of terms of trade, there exists a continuum of sub-markets, characterized by the quantity of money exchanged x, the quantity of goods exchanged q, the buyer's matching probability function $b(x, q)$ and the seller's matching probability function $s(x, q)$. The matching probabilities that buyers and sellers face in any given sub-market is a function of the relative number of buyers and sellers. If N_s and N_b are the total number of sellers and buyers in a sub-market, and $M(N_s, N_b)$ is the number of matches in a market, then $b(x, q) = M(N_s, N_b)/N_b$ and $s(x, q) = M(N_s, N_b)/N_s$. The matching technology can be rearranged and expressed as $s(x, q) = \mu(b(x, q))$.

Based on the characteristics of each sub-market, buyers can choose which market they enter, and sellers can decide how many trading posts to set up. Buyers can only enter one sub-market in the second period. Once a buyer has committed to a sub-market, they will be matched with a seller with probability $b(x, q)$. If matched, a transaction can occur with previously established terms of trade (x,q). If not matched, the agent does not get to consume. Buyers who have worked some positive amount in the first period will have carried money into the second period, which can be used in the second period transactions. Agents who invested in a counterfeiting operation in the first period have carried this operation into the second period. Agents with a counterfeiting operation established can print any amount of fake money with zero marginal cost. This fake money can also be used in second period transactions, and agents who choose to counterfeit will print however much money they need for a transaction in the sub-market that they choose to enter. There is a counterfeit detection technology that monitors each transaction, and is identical for each transaction. If a buyer is successfully matched with a trading post, and attempts to use any positive amount of counterfeit money in a transaction, there is probability θ that they will be caught. If caught, the punishment is confiscation, meaning that the agent does not get to consume this period. If the buyer is matched, uses counterfeit money, and with probability $(1-\theta)$ is not caught, the buyer gets to consume quantity q. A matched buyer who uses all real money also gets to consume quantity q. It is also worth noting that this implies that an agent who chooses to establish a counterfeiting operation in the first period will not work in the labour market at all, since the agent can produce enough counterfeit money to enter any sub-market in the second period. If an agent decides to counterfeit, work would simply impose a disutility of labour without improving one's expected outcome in the second period.

Sellers in the second market are large, identical firms owned by the rest of the world. These large firms maximize profits by establishing trading posts in sub-markets based on their characteristics. Firms can establish as many or as few trading posts as they want in a sub-market. Firms must pay a fixed cost k to establish a trading post, and if matched will get the opportunity to sell. A trading post that is matched with a buyer must produce the goods to be sold, incurring marginal cost of production $\varphi(q)$, and will sell quantity q of goods for quantity x of money. After being matched but before the seller produces the goods, the buyer's money is verified. If any amount of money is fake, there is probability θ that the counterfeiter will be identified. If identified, the transaction is cancelled, so the trading post does not incur any marginal cost of production but also does not get to trade that period. In the case that some amount of counterfeit money is involved in the transaction, and it is not detected, the trading post produces the goods, but is left with some amount of worthless fake money, which does not count as revenue.

4 Theory

The firm's problem in period 1 is independent of the firm's problem in period two, and is presented first. The consumer's problem is solved using backwards induction, so logical presentation of these individual optimization problems is in reverse chronological order, beginning with the second period.

4.1 Firm's Problem: First Period

In the first period firms produce a good using labour as the only factor of production. This product is sold to the rest of the world at an exogenous world market price P, which is normalized to 1.

The firm's profit function in the first period is expressed as:

$$\max_{l_i} \pi = \int \{a_i l_i - w_i l_i\} \, d_i \tag{1}$$

Where a_i is the productivity of type i labour, l_i is the demand for type i labour, and w_i is the market wage for type i labour. Prices are normalized to $P = 1$. The first order condition of the firm's problem with respect to their labour demand decision:

$$a_i - w_i = 0 \tag{2}$$

Thus the equilibrium wage rate in the first market is $w_i = a_i$.

4.2 Firm's Problem: Second Period

In the second period, agents search decentralized markets in a directed fashion. In particular, there is a continuum of sub-markets, each of which is characterized by its terms of trade (x, q) and its matching probabilities (b, s). Here x is the amount of money traded and q is the quantity traded in a match. The variable b represents the matching probability for a buyer and s for a seller. As will be shown later, these matching probabilities are endogenous functions of the terms of trade, i.e., $b(x, q)$

and $s(x,q)$. Large firms from the rest of the world can choose to enter any sub-market by establishing a trading post in that sub-market. The number of trading posts that a firm sets up in any sub-market is N, and the firm incurs a fixed cost k representing the start-up cost of each trading post. If a trading post is matched with a buyer and a trade occurs, the firm incurs marginal production costs $\varphi(q)$. All costs are in units of wages paid to employees in the rest of the world w, and any profits are redistributed to employees as dividends. Successful matches result in a transaction, but only successful matches with buyers using real money result in revenue. The firm incurs the marginal cost of production regardless of whether the buyer is using real money or if the buyer uses counterfeit money that goes undetected.

The profits of a large firm in the second period are expressed as:

$$\max_N \int \{xs(x,q)\,(1-f(x,q)) - w\,[k + \varphi(x,q)s\,(x,q)\,(1 - f\,(x,q) + f\,(x,q)\,(1-\Omega))]\}\,dN \tag{3}$$

In equation (3), the revenue of each trading post in each sub-market is $xs\,(x,q)\,(1-f\,(x,q))$ since $(1-f(x,q))$ represents the fraction of real money users in the market, and thus the probability of receiving legitimate money. The costs of each trading post in each sub-market is $w\,[k+\varphi\,(x,q)\,s\,(x,q)\,(1-f\,(x,q)+f\,(x,q)\,(1-\Omega))]$ because $(x,q)\,s\,(x,q)\,(1-f\,(x,q)+f\,(x,q)\,(1-\Omega))$ represents the fraction of buyers using real money or using fake money and not getting caught. Wages are normalized to $w=1$.

Free entry leads to zero profits in equilibrium. The zero profit condition implies:

$$xs\,(x,q)\,(1-f\,(x,q)) - [k+\varphi\,(x,q)\,s\,(x,q)\,(1-f\,(x,q)+f\,(x,q)\,(1-\Omega))] = 0 \tag{4}$$

From equation (4) the seller's matching function for each market with terms of trade q can be derived:

$$s(x,q) = \frac{k}{m(1-f(x,q)) - \varphi(x,q)(1-f(q)+f(q)(1-\Omega))} \tag{5}$$

The matching technology of the decentralized market is defined as $s\,(x,q) =$

$\mu[b(x,q)]$, so the buyers matching function for each market with term of trade (x, q) can be derived:

$$b(x,q) = \mu^{-1}\left[\frac{k}{m(1-f(x,q)) - \varphi(x,q)(1-f(q)+f(q)(1-\Omega))}\right] \tag{6}$$

These two equations define the likelihood of a buyer or seller being matched with a counterpart for trade, given that they enter the market defined by selling quantity q for money x.

The quantity of goods that a firm is willing to supply in a given market (x, q) can also be derived from equation (5) as a function of the seller's matching probability $s(x,q)$, and counterfeit properties of the market Ω, $f(q)$.

The firm's supply function is expressed as:

$$q = \varphi^{-1}\left[\frac{x(1-f(x,q)) - \frac{k}{s(x,q)}}{1-f(x,q)+f(x,q)(1-\Omega)}\right] \tag{7}$$

4.3 Consumer's Problem: Second Period Counterfeiter

A consumer who has established a counterfeiting operation in the first period enters the second period's decentralized market with the ability to produce as much fake money as necessary for a transaction in any sub-market. This means that in practice all counterfeiters face the same optimization problem in the second period, regardless of productivity index i since they have identical preferences and can all print as much money as they need. The counterfeiter's choice in the second period is which sub-market to enter based on the term of trade q and matching probability $b(q)$. There is however, a problem with the simplest form of this setup. Since all counterfeiters are identical, in equilibrium they would all choose to enter the same optimal sub-market, which would lead to an unstable equilibrium as demonstrated below in proof (1). This problem is overcome by allowing for counterfeiters to assign a probability $P(q)$ with which they will enter any given market rather than forcing them to commit to entering a specific market.

The counterfeiter's problem can be expressed as:

$$V_f = \max_{P(x,q)} \int P(x,q) \left\{ b(x,q)[\Omega U(0) + (1-\Omega)U(q)] + [1-b(q)]U(0) \right\} dq \quad (8)$$

$$s.t. \int P(x,q)dq = 1$$

The Lagrangian for this optimization problem is:

$$\mathcal{L} = \int P(x,q) \left\{ b(x,q)[\Omega U(0) + (1-\Omega)U(q)] + [1-b(q)]U(0) \right\} dq + \lambda[1 - \int P(x,q)dq] \quad (9)$$

The first order condition with respect to $P(x,q)$

$$b(x,q)[\Omega U(0) + (1-\Omega)U(q)] + [1-b(x,q)]U(0) - \lambda \begin{cases} = 0, & \text{if } P(x,q) \in [0,1] \\ < 0, & \text{if } P(x,q) = 0 \\ > 0, & \text{if } P(x,q) = 1 \end{cases} \quad (10)$$

The Lagrangian multiplier λ here represents the shadow value of assigning additional probability to entering a market. The value of λ therefore is strictly positive, since there is always benefit to increased chances of trading. The situation where $P(x,q) = 0$ is the least interesting of the three possibilities, since it implies that counterfeiters will assign a probability of zero to entering those markets. The situation where $P(x,q) = 1$ implies that all counterfeiters are committed to entering the same market with probability 1. This situation cannot support an equilibrium, as demonstrated in proof (1).

The situation where $P(x,q) \in [0,1]$ is the most interesting since it implies that the counterfeiters assign some probability to entering that market, and some probability to entering other markets. In this situation, there can exist an equilibrium with counterfeiters and money holders coexisting in the decentralized markets.

The relevant first order condition is:

$$b(x,q)[\Omega U(0) + (1-\Omega)U(q)] + [1-b(x,q)]U(0) - \lambda = 0 \quad (11)$$

$$b(x,q) = \frac{\lambda - U(0)}{(1-\Omega)[U(q) - U(0)]} \quad (12)$$

From equation (12), it is apparent that the buyers matching probability $b(x, q)$ is only a function of q, and not of x. This is an important property of the model because it demonstrates that the distribution of markets can be indexed by one term, either x or q. The distribution of markets can be thought of as a one dimensional line in price space, as opposed to a two dimensional distribution across x and q. From this point forward sub-markets will be indexed only by quantity, although it is important to remember that this index fully identifies the terms of trade for the sub-markets.

4.4 Consumer's Problem: Second Period Money Holder

The consumer in the second period who has chosen not to counterfeit will hold legitimate money earned in the first period's labour market. Entering the decentralized market, the consumer can choose which sub-market to enter based on the terms of trade and matching probabilities. The consumer's value function of money entering the second market is:

$$V(m) = \max_{x,q} b(x, q)U(q) + [1 - b(x, q)]U(0) \tag{13}$$

$$\text{s.t. } x \leq m$$

The consumer's first order condition with respect to q is

$$\frac{\partial b(x, q)}{\partial q}[U(q) - U(0)] + b(x, q)U'(q) = 0 \tag{14}$$

4.5 Consumer's Problem: 1st Period

The consumer's problem in the first period is to determine their allocation of resources across tasks. The agent can rent their labour to firms from the rest of the world in a Walrasian market, earning them real money which is carried into the second period. Agents earn marginal wage w, and incur disutility of labour $h(l_i)$ for each unit of labour supplied. Agents can also establish a counterfeiting operation, incurring fixed cost c but allowing them to produce as much counterfeit money in the second period as required to transact in any sub-market of their choosing. It is important to notice that since counterfeiting allows for zero marginal cost money production, agents will never choose to work and counterfeit. The agent's problem is therefore expressed as:

$$\max_{\epsilon_i, l} \epsilon_i[V_f - c] + (1 - \epsilon_i)[V(m) - h(l_i)] \tag{15}$$

Note that $m = w_i l_i$, and from equation (2) $w_i = a_i$

The first order condition with respect to labour supply l_i:

$$V'(w_i l_i) w_i = U'(l_i) \tag{16}$$

This implies of course in equilibrium the marginal benefits of labour supply are equal to the marginal costs.

The first order condition with respect to the work/counterfeit decision ϵ is:

$$[V_f - c] - [V(m) - h(l_i)] \begin{cases} = 0, & \text{then } \epsilon_i \in [0, 1] \\ < 0, & \text{then } \epsilon_i = 0 \\ > 0, & \text{then } \epsilon_i = 1 \end{cases} \tag{17}$$

This implies that when the net benefits of working and the net benefits of counterfeiting can be equal, agents will be indifferent between counterfeiting and working, and will counterfeit with some probability $\epsilon_i \in [0, 1]$. When net benefits of counterfeiting are strictly greater than net benefits to working, $\epsilon_i = 1$ and these agents will always counterfeit. If net benefits of counterfeiting are strictly less than net benefits to working, $\epsilon_i = 0$ and these agents will never counterfeit.

This result implies that there exists some exogenous productivity parameter a^*, which is a cut-off between counterfeiting and working. If $a^* \in [\underline{a}, \overline{a}]$, then the equilibrium will be characterized by agents with $a_i < a^*$ who will counterfeit and agents with $a_i > a^*$ who will work. This result is shown in proof (2).

5 Equilibrium

The equilibrium consists of value functions $V(m)$, V_f; matching probabilities $b(x,q)$, $s(x,q)$; terms of trade (x,q); firm's loss probability $f(x,q)$; counterfeiting decision ϵ_i; entering probability $P(q)$; labour supply l_i; and wages w_i. Furthermore, these elements must satisfy requirements: (i) Given $\varphi(q)$, k, $f(q)$ and the matching function $s(x,q)$, a seller's choice solves the firm's problems; (ii) Given w_i and the function $b(x,q) = \mu^{-1}(s(x,q))$, a buyer's optimal choices solves the money holder's problem $V(m)$; (iii) Given w_i, and matching function $b(x,q) = \mu^{-1}(s(x,q))$, a counterfeiter's optimal choices solves V_f; (iv) Given a_i and c, the agent's choice ϵ_i maximizes utility (v) Free entry in both first and second period means that the expected profit of a firm in the first period and a trading post in any active sub-market in the second period is zero; (vi) Counterfeiting behaviour is present, or for some agent i, $\epsilon_i > 0$.

The first property to be analyzed is the effect of counterfeiters on firms in individual sub-markets. This requires defining the function $f(q)$, or the fraction of consumers in a market using fake money. Using the fact that agents are normalized to one and productivities are uniformly distributed with cumulative distribution function $G(a)$ and corresponding probability distribution function g, we can define $f(q)$:

$$f(q) = \frac{G(a^*)P(q)}{G(a^*)P(q) + g} \tag{18}$$

From equation (18), the first order condition with respect to q, $f'(q)$ is strictly positive if $P'(q)$ is positive. This result means that firms entering into high quantity markets will expect to face a greater fraction of counterfeiters.

To analyze the behaviour of agents entering the second decentralized market, one first turns to equation (10). This equation allows one to examine the relationship between terms of trade and matching probabilities of a chosen sub-market. From equation (10), one considers the case with a counterfeiting equilibrium, or the case where the first order condition is equal to zero. Solving for the buyer's matching function gives:

$$b(q) = -\frac{\lambda - U(0)}{\Omega U(0) + (1-\Omega)U(q) - U(0)} \tag{19}$$

Taking the first order condition of equation (19) and rearranging yields:

$$b'(q) = -\frac{\lambda - U(0)U'(q)}{[(1-\Omega)(U(q) - U(0)]^2} \tag{20}$$

Every individual term on the right hand side of equation (20) is positive, so the total right hand side of equation (20) is strictly negative, indicating that $\frac{db}{dq} < 0$. This result demonstrates that high quantity sub-markets offer a lower matching probability.

If the distribution of sub-markets is such that $q'(m) > 0$, or that sub-markets charging more money give more goods, then one can analyze how the distribution of agents using real money will self select into different sub-markets. From $b(m) = b[q(m)]$ the chain rule gives:

$$\frac{db}{dm} = \frac{db}{dq} \times \frac{dq}{dm} \tag{21}$$

Since $\frac{db}{dq} < 0$ and $\frac{dq}{dm} > 0$, it is shown that $\frac{db}{dm} < 0$

The result of $\frac{db}{dm} < 0$ is that when agents who enter the second period holding more money are choosing to enter markets with a lower probability of a match. These are high quantity markets, where firms are reluctant to establish trading posts because of the greater marginal costs of production and risk of counterfeiting.

To determine who counterfeits and who does not, equation (15) can be used to determine the effect of exogenous productivity parameter a_i on counterfeiting behaviour. To determine $\frac{dV}{da_i} > 0$, the chain rule is applied:

$$\frac{dV}{da_i} = \frac{dV}{dm} \times \frac{dm}{da_i} \tag{22}$$

First, solving the sign of $\frac{dV}{dm}$ requires the consumer's first order condition, equation (14). The First order condition with respect to q is:

$$\frac{\partial b(m,q)}{\partial q}[U(q) - U(0)] + b(m,q)U'(q) = 0 \tag{23}$$

Applying the envelope theorem gives:

$$\frac{dV(m)}{dm} = \frac{\partial b(m,q)}{\partial m}[U(q) - U(0)] > 0 \tag{24}$$

From equation (24) $\frac{dV(m)}{dm}$ is strictly positive. This is an intuitive result, showing that agents who hold money entering the second period prefer to hold more money than less.

First period consumers budget constraint $m = w_i l_i$ shows that $\frac{dm}{dw_i} > 0$. Since $\frac{d}{dw_i} > 0$ and $\frac{dV}{dm} > 0$, by equation (22) $\frac{dV}{dw_i} > 0$. Taking this result and using the first order condition from the firm's first period problem, equation (2), it becomes evident that $\frac{dV}{da_i} > 0$. This result simply shows that those who choose to use money in the second period are better off if they were more productive in the first period.

Finally, combining this result with the first order condition from the consumer's first period decision, equation (17), shows that $\frac{de_i}{da_i} \leq 0$. This is an important result, showing that in equilibrium, less productive agents are more likely to counterfeit than their more productive counterparts. In proof (2), it is demonstrated that there exists a productivity parameter a^* such that agents more productive than this level will always counterfeit, and agents that are less productive than a^* will always counterfeit.

6 Conclusions

This model presented provides a framework for the analysis of counterfeiting behaviour in equilibrium, and attempts to approximate the decisions that counterfeiters must make in the real world. The model suggests that there is some reservation level of productivity at which agents will counterfeit otherwise they will work in the labour market. The actual level is a function of several model parameters. The model predicts that sub-markets will arrange in a way such that markets selling higher quantities of goods have a lower matching probability. Agents who enter the second period with more money will choose to enter the higher quantity, lower probability sub-markets. The effect of counterfeiters is found to be increasing in sub-market quantity, which implies greater costs to sellers in those sub-markets.

A potential direction for future research would be to generalize the model into a microfoundation over multiple periods. It would also be interesting to add a banking sector which would affect the verification of money, since counterfeit money is generally removed from circulation at banks. The exogenous nature of the verification technology is also somewhat problematic. An appropriate extension would be to introduce a system of costly verification whereby individual sellers could choose the amount of resources to devote to money verification.

7 References

- Becker, G. (1968): Crime and Punishment: An Economic Approach, *The Journal of Political Economy, 7*(2), 169-217.

- Green, E., and Weber, W. (1996): Will the New $100 Bill Decrease Counterfeiting?, *Federal Reserve Bank of Minneapolis Quarterly Review, 20*(3), 3-10.

- Judson, R., and Porter R. (2010): Estimating the Volume of Counterfeit U.S. Currency in Circulation Worldwide: Data and Extrapolation, *Federal Reserve Bank of Chicago Financial Markets Group Policy Discussion Paper Series*, 2010-2.

- Li, Y. and Rocheteau, G. (2008): On the Threat of Counterfeiting, *Federal Reserve Bank of Cleveland Working Papers 08-09*

- Menzio, G., S. Shi and H. Sun (2009): A Monetary Theory with Non-Degenerate Distributions, *manuscript in preperation.*

- Nosal E., and Wallace N., (2007): A Model of (the Threat of) Counterfeiting, *Journal of Monetary Economics, 54* 994-1001

- Phillips (2005): *Knockoff: The Deadly Trade in Counterfeit Goods.* Kogan Page Limited, London.

- Quercioli, E., and Smith, L. (2009): *The Economics of Counterfeiting.* Retrieved October 20, 2010 from Social Science Research Network http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1325892

- Royal Canadian Mounted Police (2007): *Counterfeit Currency in Canada.* Retrieved October 20, 2010 from http://www.rcmp - grc.gc.ca/pubs/ci - rc/cf/index - eng.html

- Stigler, G (1970): The Optimum Enforcement of Laws,*The Journal of Political Economy, 78*(3), 526-536.

- Unites States Treasury Department (2003): *The Use and Counterfeiting of United States Currency Abroad, Part 2.* Retrieved on-line October 20, 2010 from http://www.federalreserve.gov/boarddocs/RptCongress/counterfeit2003.pdf.

- Williamson, S. (2002): Private Money and Counterfeiting, *Federal Reserve Bank of Richmond Economic Quarterly Volume*, 88/3.

8 Appendix

8.1 Proof 1

If $P(q) = 1$, or if all counterfeiters were to commit to entering a particular sub-market, it would not exist in equilibrium. First, consider some $\tilde{q}$ such that $P(\tilde{q}) = 1$. This means that the sub-market indexed by $\tilde{q}$ be the market that counterfeiters have determined to be optimal. Recall that all counterfeiters face the same value function, equation (8), and therefore they all enter the same sub-market. Now, consider the properties of market $\tilde{q}$. Let $E = \int \epsilon_i di$ be the total number of counterfeiters, and therefore the number of counterfeiters in sub-market $\tilde{q}$. Let $G(\tilde{q})$ = Total number of money users in sub-market $\tilde{q}$. Then $f(\tilde{q}) = \frac{E}{E+G}$ is the fraction of counterfeiters in sub-market $\tilde{q}$

Recall from equation (3) that $\pi(\tilde{q}) = xs(q)(1-f(q))-w[k+\varphi(q)s(q)(1-f(q)+f(q)(1-\Omega))]$, and therefore, $\frac{d\pi}{df} < 0$. So, if there is $\pi(f(\tilde{q}), \tilde{q})$, and a continuum of sub-markets in the second period, then there exists a market characterized by terms of trade $\tilde{q} \pm \beta$ where β is some very small number. When $P(\tilde{q}) = 1$, $f(\tilde{q}) = \frac{E}{E+G} > 0$ and $f(f(\tilde{q} \pm \beta) = 0$. This implies that $\pi(\tilde{q}) < \pi(\tilde{q} \pm \beta)$.

The result of all counterfeiters entering sub-market $\tilde{q}$ would be that $N(\tilde{q}) = 0$, since all firms would have an incentive to set up trading posts in sub-market $\tilde{q} \pm \beta$ rather than $\tilde{q}$, where profits would be strictly greater. This means that there are no trading posts in sub-market $\tilde{q}$, matching probabilities are zero, and the sub-market effectively ceases to exist. The fact that for any value of $\tilde{q}$ that no sub-market $\tilde{q}$ will exist with $P(\tilde{q}) = 1$, means that all counterfeiters cannot be in the same market in equilibrium. QED.

8.2 Proof 2

There exists an exogenous productivity parameter a^* that behaves as a reservation wage. That is, agents more productive $a_i > a^*$ will work, while agents less productive $a_i < a^*$ will counterfeit. This reservation productivity level is only relevant to the model if it is calibrated in such a way that $\underline{a} < a^* < \overline{a}$. Consider the first order condition of the consumer's first period decision, equation (17),

$$[V_f - c] - [V(m) - h(l_i)] \begin{cases} = 0, & \text{then } \epsilon_i \in [0,1] \\ < 0, & \text{then } \epsilon_i = 0 \\ > 0, & \text{then } \epsilon_i = 1 \end{cases} \tag{25}$$

From equation (2) and equation (25), we can re-write equation (25) as: $[V_f - c] - [V(a_i l_i^*) - h(l_i^*)]$. Agents are indifferent between counterfeiting and working when this term is equal to zero, or $[V_f - c] = [V(a_i l_i^*) - h(l_i^*)]$. Rearranging and solving for the a_i value that allows this equality to hold gives the unique a^* that characterizes indifference between working and counterfeiting, and therefore functions as a reservation productivity. Endogenous productivity parameter a_i can be characterized in one of three ways.

$$a_i \begin{cases} = a^*, & \text{then } \epsilon_i \in [0,1] \\ < a^*, & \text{then } \epsilon_i = 0 \\ > a^*, & \text{then } \epsilon_i = 1 \end{cases} \tag{26}$$

The interpretation of this result is that if $a_i < a^*$ then $\epsilon' = 1$ and agents will always counterfeit, and, if $a_i > a^*$ then $\epsilon' = 0$ and agents will never counterfeit, and if $a_i = a^*$ agents are indifferent between counterfeiting and working. Therefore, a^* is the reservation wage. QED.