Graff, Michael

Working Paper

The quantity theory of money in historical perspective

KOF Working Papers, No. 196

Provided in Cooperation with:
KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Graff, Michael (2008) : The quantity theory of money in historical perspective, KOF Working Papers, No. 196, ETH Zurich, KOF Swiss Economic Institute, Zurich, http://dx.doi.org/10.3929/ethz-a-005582276

This Version is available at:
http://hdl.handle.net/10419/50322

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
The Quantity Theory of Money in Historical Perspective

Michael Graff
The Quantity Theory of Money in Historical Perspective

Brisbane and Zurich, April 2008

Abstract
The paper reconstructs the origins of the quantity theory of money and its applications. Against the background of the history of money, it is shown that the theory was flexible enough to adapt to institutional change and thus succeeded in maintaining its relevance. To this day, it is useful as an analytical framework. Although, due to Goodhart's Law, it now has only limited potential to guide monetary policy and was consequently abandoned by most central banks, an empirical analysis drawing on a panel data set covering more than hundred countries from 1991 to the present confirms that the theory still holds: a positive correlation between the excess growth rate of the stock of money and the rate of inflation cannot be rejected. Yet, while the correlation holds for the whole sample, proportionality is driven by a small number of influential observations with very high inflation.

JEL-Classification:
B10 History of Economic Thought through 1925, General
E41 Demand for Money
E58 Central Banks and Their Policies

Keywords:
Quantity theory of money, demand for money, monetary targeting
The Quantity Theory of Money in Historical Perspective

1 Introduction

The quantity theory of money (hereafter: "quantity theory") is possibly the only economic theory dating back from before the establishment of classical national economics around the end of the eighteenth century, which still features prominently in economics, being called upon as a theoretical tool and a basis of economic policies and strategies.

The relevance of the quantity theory reveals itself in historical perspective. It emerged in times of then unfamiliar price fluctuations in mediaeval Europe. The ancestry of its classical version, referring to changes in the money supply, can be traced to the early modern era.

In the 1970s, when prices in the capitalist world rose across the board, Milton Friedman's reformulation of the quantity theory (FRIEDMAN 1956) achieved highest standing amongst economists. Apart from providing the theoretical basis of monetarism, with significant influence on academic discussion, it became highly influential in shaping the conduct of monetary policy in central banks around the globe.

The theory's recent heyday did not last long. In the US, monetary targeting was abandoned in the early 1980s, and in the English-speaking world, monetary aggregates are nowadays generally considered inappropriate as intermediate targets. However, other countries, particularly in Central Europe, were not so quick to dismiss the theory. The German Bundesbank communicated the control of the money supply as its theoretical foundation for the conduct of monetary policy for over a quarter of a century to the very end of the DM era, i.e. until the establishment of the European Monetary Union (EMU) in 1999, and the Swiss National Bank (SNB) adhered to monetary targeting as long as the Bundesbank. Furthermore, at the time when the SNB abandoned monetary targeting in 1999, the European Central Bank (ECB), as a legacy from the Bundesbank, adopted the quantity theory as one of its two official pillars for monetary policy. A reason for this was certainly to transfer some of the reputation from the old to the new Frankfurt monetary authority.

A few years later, in May 2003, the ECB, in what appeared to be a rather low-key press release, significantly downgraded the theory as a practical guidance for its conduct of monetary policy. Nevertheless, the ECB's monetary analysis persists up to this date.

Arguably, on the operational level, monetary policy in Frankfurt never truly targeted the quantity of money, but rather inter-bank interest rates. Indeed, the fact that the Bundesbank's track shows significant and repeated deviations from its official monetary targets, suggest that it may have referred to the quantity theory and the money supply calculus on a communications level, while it actually tried to target inflation. Nevertheless, the change in emphasis that the ECB devotes to its two pillars is now leaving no doubt that also in Frankfurt

---

1 For an authoritative representation of Bundesbank's monetary policy since 1994, see ISSING (1997).
3 Titled "The ECB's monetary policy strategy", the 8 May 2003 press release states: "… the Governing Council of the ECB has undertaken a thorough evaluation of the ECB's monetary policy strategy … [T]he monetary analysis mainly serves as a means of cross-checking, from a medium to long-term perspective, the short to medium-term indications coming from economic analysis. To underscore the longer-term nature of the reference value for monetary growth as a benchmark for the assessment of monetary developments, the Governing Council also decided to no longer conduct a review of the reference value on an annual basis."
4 See VON HAGEN ET AL. (1999). The same is claimed regarding the SNB (MISIKIN 1999).
the money supply is de facto no longer an intermediate target. Thus, all major central banks now, at least in practical terms, have shelved the quantity theory.\footnote{See e.g. BERGER ET AL. (2006), who analyse the ECB's statements and press conferences from 1999 to 2004 and find that the relative amount of space devoted to monetary analysis has decreased; implying that role of money has indeed been downgraded since 1999.}

While the theory's practical influence has recently faded, it is still a subject of passionate discussions by both followers and sceptics. Indeed, as this paper will argue, it would be premature to delegate the quantity theory to the history of economic thought.

Economic history, and in particular the history of money, is crucial for a critical assessment of the theory and its relevance for economic policy. This paper will hence track the quantity theory in historical perspective. In particular, we shall first identify major steps in the history of money (section 2). On this basis, we shall reconstruct the theory as a reflection of monetary conditions (section 3). Then, we shall conduct an empirical analysis to assess in how far the relationship between money and inflation that follows from the quantity theory still holds today (section 4). On this basis, we shall conclude with a discussion of the present status of the theory and its implication for monetary policy (section 5).

2 The theory's background: a short history of money

2.1 The origins

Early standardised means of transaction (pre-monetary currency) were \textit{marketable}\footnote{See VON MENGES'S (1892) seminal contribution.} goods, where the value of money was identical to the commodity value of the circulating means of exchange. Due to practical requirements, such as homogeneity, durability and, last but not least, divisibility, metals and bullion soon formed the basis of early commodity money.

Since metal or bullion does not provide any obvious, intrinsic units or cuts, at this stage, the problem of standard weighting arose. The solution was to define standardised units of weight for bullion that would be used as currency, and to cut it to size accordingly. In the second century BC, bullion with \textit{private} stamps appears, which would have been used for monetary purposes (EINZIG 1949: 226). Around the end of the eighth century BC, the first \textit{royal} stamps confirming the purity or weight emerge in Asia Minor. Although these early forerunners to coinage were rather clumsy lumps, the sovereign's stamps confirming their weight implies that they can indeed be considered as coins in economics terms (BURNS 1927: 39 ff., DAVIES 1994: 60 ff.).

Coinage quickly spread through the Eastern Mediterranean as well as through the Middle East. Technical improvements soon followed, such as the disc shape, which continues in use today, embossing instead of casting, and while the first coins carried stamps on one side only, later the entire surface of the coin was richly embossed, making it increasingly difficult to clip or rasp bullion off the coin's edges, or to commit outright forgery.

On this basis, the new medium of exchange gained acceptance (BURNS 1927: 54 ff.) While it may have been customary to weigh the coins in the early days, a trusted seal from a chartered mint would contribute to eliminate the need for weighing for many, if not most, transactions (BURNS 1927: 58 ff.). Thus, the prototype of the full-bodied coinage, minted under official charter came about.\footnote{The next notable improvement in the production of coinage, the screw press, followed in early modern times.}
Notably, full-bodied metal currency has remained the basis of the conception of 'money' right into this century. Coins below their intrinsic value were rare, and generally just those of low value. For some thousand years, money was bullion (Burns 1927: 284 ff.).

Based on the antique design, this monetary system lasted for some two and a half thousand years. Apart from a minuscule fraction that the mint would charge for its services – usually far less than one per cent – money that consisted of full-bodied coins was a standard that referred to arbitrarily defined quantities of bullion. Underweight coins would either be rejected or circulate at a discount, and undervalued coinage would vanish from circulation as soon as its face value dropped below the commodity value (plus cost of melting and selling the metal). Price stability was guaranteed by the intrinsic scarcity of gold and silver. As the use of bullion for monetary in addition to ornamental or technical purposes created additional demand that would raise the price of bullion in terms of other commodities, the scarcity of bullion was further enhanced, which helped to keep the prices of other goods low. This contributed to the trust in the purchasing power of bullion as well as in the monetary system based on the latter.

Yet, with money came inflation, a phenomenon that is accordingly known since antiquity. The earliest systematic debasements of coinage go back to the Greek city-states. Performed as reductions of the share of gold or silver in the alloy to be minted, or by hiding cores of low value metal within the coin, debasements were initially carried out with caution and probably only in wartime, so that they would not be permanent threat to acceptance of the money as a means of payments (Burns 1927: 339 ff., Schwartz 1973: 244). The noticeable increase in price levels in ancient Greece was hence not brought about by debasement but by the influx of large amounts of Persian gold after the conquest of Persia under Alexander the Great (Schwartz 1973: 244 ff.). In Rome, when the Punic wars were stressing the state finances, debasement was more common, and indeed, the first long-term inflation in history is due to systematic coin debasement in the Roman Empire, when in the course of roughly a century with average inflation rates of up to four per cent, the price level increased by a factor of fifty (Burns 1927: 407 ff., Friedman 1994: 190).

2.2 The Dark Ages and the Commercial Revolution

The fall of the Western Roman Empire in 476 heralded the start of the European 'Dark Ages', stretching on to the eleventh century. During this period of organisational decay and stagnation, the ancient monetary standards largely disappeared and non-Byzantine Europe experienced progressive demonetisation. Coinage crumbled away and money vanished from circulation. Although Roman currency still served as a measure of value, there was no intrinsic difference between coins, jewellery, or other valuables. Credit was largely limited to consumption loans in times of hardship and denominated in quantity of goods rather than in monetary units. Accordingly, investment had to rely on internal finance, and an occasional surplus would be hoarded (Cipolla 1956: Chapter 1, 1993: 160, Davies 1994: 111 ff.).

Isolated resumptions of European coinage are documented for the eighth century, and by the tenth century, in Northern Italy new coinage reappeared in large numbers, resulting in effective remonetisation. This marks the beginning of the 'Commercial Revolution', which from the eleventh to the thirteenth century shaped the prototype of the modern financial system, including the settlement of payments, banking and insurance. The driving factor for this surge of innovation was the international trade in goods that had started to recommence on a large scale, with Northern Italy as its hub. Merchants in Genoa, Venice, Florence and the

---

8 Where minting continued, coins were 'barbaric' or 'impure' (Taeuber 1933: 66), so that larger transactions were carried out through exchange of goods for gold or other goods, and not coined money.
cities of Lombardy and Florence created an unprecedented need for finance and liquidity as well as constant demand for financial services such as currency exchange and naval insurance.

With the return to the regular minting of money based on full-bodied coins, the related problems do not fail to re-emerge. In addition to wear and tear, the clipping of coins (by the public) and periodic debasements (by the authority) regularly contributed to reducing the bullion content of the coinage, which soon or later inevitably would result in devaluation (LOPEZ 1976, SUSSMAN and ZEIRA 2003). The history of money, which for the most part is the history of full-bodied money, is hence mostly a chronology of debasements and subsequent monetary reforms to restore the old standard or to establish a new one with modified mint parity. The significance of this chapter of economic history consists in illustrating that the ancient monetary system quite effectively ensured that money as a unit of measurement would correspond to a well-defined quantity of bullion. Significant deviations from the official parity would soon or later be detected and not tolerated, hence resulting in monetary reform. Inevitably, the authority at some stage would feel that its coffers were unduly depleted and again be tempted to debase the coinage, so that the cycle of debasement and reform would continue.

Though coined money as a rule would conform rather closely to the mint parity, the risk to receive underweight coins was certainly too high to be neglected. Accordingly, for large payments settled in cash, the requirement to verify purity and weight persisted. This inconvenience triggered the path-breaking innovation of credit money (LOPEZ 1976: 71 ff.).

To reduce the number of payments that would finally have to be settled in cash, merchants of Northern Italy developed a number of new financial instruments. The most straightforward was to grant customers trade credit for later settlement. By softening the liquidity constraint of the cash economy, this contributed to make investment and consumption decision less dependent on current liquidity (CIPOLLA 1993: 160 ff.).

Another – arguably more important – innovation was the private cash account ('giro di partita') at the new 'banchi di scritta' or 'banchi del giro'. Apart from the peace of mind from knowing that coins or bullion was stored securely in the coffers of a bank, this innovation soon delivered a new kind of payments service, as payments from one customer of the bank to another – usually members of the local merchant community – could be performed as transfers between accounts. Initially both payer and payee had to appear at the bank in person, but soon, a written order by the payer, stating the payee and the sum to be transferred, would be sufficient. Thus, the earliest known cheque was issued in 1368.9 This innovation made the settlement of payments safer and at the same time markedly reduced transaction cost.

Since the early cheques represented intrapersonal transfers of an existing claim on coins or bullion in the coffers of the bank, they did not increase the quantity of money, at least as long as the banks held 100 per cent reserves, as their statutes required. It would not take long, however, until some banks silently began to breach the rules and offer standing overdraft facilities to their most trustworthy customers, which marks the beginning of fractional reserve banking (CIPOLLA 1993: 181, SPUFFORD 1988: 256 ff.). The most important implication for money and banking was that fractional reserve banking allowed expanding the means of payments – i.e. the quantity of money – beyond the rigid limits previously set by the quantity of bullion.

During the Commercial Revolution, bank money, important as it was to the merchant community, occupied a relatively small niche. Hence, despite its potential, the effect that this

---

9 Note, however, that elsewhere cheques did not appear before the seventeenth century (PARKER 1979: 344).
innovation had on the quantity of money initially remained very limited. In fact, it was not before the twentieth century that it became common for payments affecting the wider public.

Another important financial innovation was the mediaeval letter of exchange, an instruction to a merchant or a banker in another city to pay a specified amount to a third party. Since before this, inter-city or international trade had required risky shipments of coins or bullion, the letter of exchange was highly significant in supporting the expansion of trade. Apart from this, the letter of exchange helped to evade the ecclesiastical usury regulations of that time. Instead of charging explicit interest, it could disguise it as 'discount'. For transactions between areas of different coinage, the exchange rate provided an opportunity to effectively charging interest without arousing the potentially dangerous attention of the church.

Though the Commercial Revolution took place in the centres of trade, i.e. the town, money and credit soon and increasingly affected also the rural population of mediaeval Europe. Initially this was because the growing population in the towns would buy victuals from the surrounding countryside for money, with the result that using money became common amongst rural people. Moreover, while serfdom prevailed, from the tenth century, an increasing part of the rural population was employed as wage-labourers (CAMERON 1993: 50). Moreover, mediaeval lords of feudal properties were increasingly being paid in money instead of feudal service or payment in kind. They too hence started to measure profit in money (DUBY 1978: 132 ff., LOPEZ 1976: 155 ff.). Thus, throughout the medieval society, economic activity was directed more and more at the realisation of money revenue, and prototypes of the capitalistic enterprise emerged (WEBER 1924: 86, see also CIPOLLA 1993: 94). In addition to this, the way of life of the aristocracy further promoted monetisation and credit. On the one hand, there were the expenses related to taking up public office. On the other hand, gambling was endemic and almost obligatory in aristocratic circles and resulted in huge gains and losses. As the bourgeoisie copied the habits of the aristocracy, gambling spread beyond the nobility (DEWALD 1993). As a result of these developments, within a few centuries, the largely cash-free society of the 'Dark Ages' was transformed into an economy that was monetised 'from head to toe' (SPUFFORD 1988: 378). Yet, though the Commercial Revolution successfully developed credit money and clearing soon became common practice,

10 In WEBER'S judgement (1924: 229), the mediaeval letter of exchange had the same irreplaceable function as a means of making payments over distances that the cheque had in Weber's time. Yet, the letter of exchange was not accessible to anyone, since it depended on close relations with trustworthy and liquid business partners. Merchants would most likely refer to networks consisting of relatives, which were common amongst Sephardic Jews or Calvinist families (GLAMANN 1979: 328).

11 See (BERNARD 1978: 201 ff.). A canonical prohibition against interest for clerics was enacted in 325; from 806 it was extended to lay people. A breach could lead to condemnation by the church. A gradual relaxation took place from 1516, when money lending was tolerated on charitable grounds if the interest rate was modest (BORN 1994: 5 ff.). The protestant rejection of interest prohibition can be traced back to Calvin in 1754 (BEUTIN and KELLENBENZ 1973: 133). The Napoleonic Code civil interest payments were explicitly allowed, which at the time could have had a catalyst effect, at least in catholic Southern Europe (CAMERON 1993: 213). Although, as indicated above, ecclesiastical prohibition of usury could be circumvented, the normative power of mediaeval religious authority should not be underestimated (PARKER 1979: 343). The fact that credit at interest spread throughout business circles owes much to the fact that agents and traders had to have an adventurous spirit and to be willing to take risks in view of the precarious nature of land and sea travel. Moreover, due to their nomadic way of life outside of the ordered mediaeval world, it was probably easier for merchants than most other people to distance themselves from the prevailing values and mores. However, a deliberate, rational, anti-clerical rejection of interest prohibition was not able to evolve until well into the time of the Enlightenment; see KELLENBENZ (1979: 150).

12 Yet, this transformation proceeded slowly and initially only in a limited number of places; in Italy and Russia, its completion took until the nineteenth century (MINCHINTON 1976: 59).
liquidity was ultimately restricted and depended on access to gold and silver: Unbalanced accounts – at the bank or with a trading partner – would finally require settlement in full-bodied coins or bullion.

From a quantity theory perspective, an important fact from the economic history mediaeval Europe is that apart from all innovations brought about by the Commercial Revolution, the stock of money grew rather smoothly, without sudden increases or decreases that would have led contemporary observers to speculate about possible links between the quantity of money and the level of prices.

Yet, when the bubonic plague swept over Europe from 1347 and within a few years reduced the population by an estimated third (PIRENNE 1986: 186), this had an impact on prices that did not escape the attention of a thoughtful observer. (We shall return to this below.)

2.3 Modern times

In late mediaeval Europe, the demand for money started to outstrip supply, so that in relative terms, currency became scarcer than in the preceding centuries. The mining and refining of bullion failed to keep pace with the increase in economic activity that was settled in cash. In addition, money in circulation was depleted through normal wear and tear as well as through hoarding, which, in times of unrest often meant that coins were buried in depots and abandoned when their possessors lost their lives.

Furthermore, the world economy at that stage required a constant flow of gold and silver to the East. While Oriental goods were sought-after in Europe, in the East there was little interest in European commodities. Hence, the European balance of trade with the East was constantly negative and bullion has to be shipped to the East for settlement (DAY 1978, PARKER 1979).

The 'Great Bullion Famine' that resulted from these developments formed the empirical basis for European Mercantilism, and particularly the doctrine of 'bullionism' (DAY 1978: 49). Indeed, for an economy in which the means of payments was predominantly restricted to full-bodied coins, bullionism was a well-founded strategy to counter a downward pressure on prices and hence to avert deflation and repression. However, with a production of bullion that was minor relative to the stock inherited from the past, so that the supply of money was practically exogenous, bullionism, once adopted by all relevant European Countries, was effectively a zero-sum game. Consequently, the economic policy of mercantilism could offer no remedy to the scarcity of money and the 'hunger' for bullion in Europe as a whole. The only promising strategy for Europe was to find substitutes for bullion that were not subjected to the same inherent scarcity and yet suitable to serve as money.13

Hence, at the dusk of the Middle Ages, the European monetary system was lacking innovation that would deliver substitutes for its bullion currency, which by now had become both common in usage and scarce in supply. Yet, it took centuries for the solution to this problem – currency that circulated below a parity corresponding to its commodity price – to gain widespread acceptance.14 In fact, the final dissociation of money from bullion did not

13 We note in passing that, given the knowledge of the nature of chemical elements at that time, alchemy was a rational attempt to fight the 'bullion famine', although – as we know now – it was doomed to fail.

14 According to PARKER (1979: 352), the financial innovations that occurred during the early centuries of modern times modified the elasticity of the money supply, allowing an increase in the ratio of money in circulation relative to the supply of bullion by 25 per cent in the most successful country (England). The advance of credit money was hence painfully slow.
take place before the 1970s, when the convertibility of the US dollar into gold was suspended.15

Apart from institutional inertia, the most important retarding factor in the spread of credit money was that as a result of what has been called the 'Age of Discovery', the European expansion. It was driven by the hunger for bullion as much as – if not more than – by the desire to spread the gospel and succeeded to establish a continuous and sizeable flow of bullion from the New World to the Old, thus increasing the supply of bullion based money and alleviating the scarcity of currency. Thus, "[t]he famine of precious metal that had strangled the European Economy during the Middle Ages was over" (CIPOLLA 1993: 214), implying that monetary innovation became far less urgent.

In fact, scarcity of money was followed by abundance, and hence, towards the end of the sixteenth century, people could witness the beginning of a new kind of inflation, subsequently called the 'Price Revolution', that lasted some 150 years and provided the quantity theory with a sound empirical basis. Europe's stock of bullion increased modestly between 1500 and 1580, then rapidly between 1580 and 1620. Before silver mining collapsed and shipments from America fell strongly during the first 20 years of the seventeenth century,16 the influx of precious metal from the New World had elevated the price level in Europe by some 300 to 400 per cent (CAMERON 1993: 107, CIPOLLA 1993: 215). This historical account of an obvious correlation between the stock of money and the price level for centuries constituted the empirical basis for the quantity theory.

Yet, the relief to the European 'bullion famine' though influx from the New World brought, was temporary and hence not a way to put an end to the scarcity of money for good. When the shipments started to decline, innovation was back on the agenda. By that time, it was not only triggered by the scarcity of money but also by the inconvenience of full-bodied coins for the settlement of large payments, as the early prototypes of credit money that had been developed in Italy during the Commercial Revolution had not yet diffused from their niches.

This now started to change. When the mediaeval trade with the East, taking the Silk Road to the Levantine Coast, then by ship across the Mediterranean to the trade centres of Northern Italy, was replaced by the Africa route, new centres of trade and commerce emerged in the European ports of the Atlantic. By the seventeenth century, the financial instruments developed during the Commercial Revolution had been adopted in the Northwest (SOMBART 1916, Vol. II, DE ROOVER 1942).17 From that time, until 1914, the world's centres of money,

---

15 Yet, let us recall that economic history provides a host of examples of monetary decay and disintegration, illustrating that organisational and institutional progress is reversible, as during regressions to a barter economy or a commodity money standard. Frequently cited in the economic literature are the 'cigarette standards' in Sicily, Germany and in the Pacific after 1944, see e.g. Radford (1945).

16 See GLAMANN (1979: 337). According to PARKER's (1979) calculations, the balance of influx and outflow of precious metals for Europe in early modern times (up to 1730) was as follows: the stocks in 1500 were around 3500 tons of Gold and 37500 tons of silver. From internal sources (mainly the mines of Saxony and the Tyrol), up to 70 tons per year were extracted. Apart from this, around 1500, half a ton of gold flowed from Portuguese West Africa to Europe. However, these amounts vanish into nothing compared with the influx from the New World. In 1571 silver was discovered in Potosi (now Bolivia), in1546 in San Luis Potosi (Mexico) and in Guanajuato (Mexico). Moreover, in 1571 mercury was discovered in Huencavelica (Peru), which dramatically reduced the cost of silver production (CIPOLLA 1993). The official records – which do not account for smuggling and privateering – show that as a result, from 1500 to 1650, at least 17,000 tons of silver arrived in Spain. After the 1693 gold discovery in Minas Gerais (Brazil), gold started flowing from Brazil to Portugal. Although there still was a negative balance of trade with the Levant and the East, due to which up to 80 tons of silver per year had to be exported, the overall balance of bullion flows was clearly positive.

17 For descriptions of the technical and organisational innovations in money and banking in mediaeval Europe and during the dawn of modern times, see VAN DER WEE (1977), PARKER (1979) and GLAMANN (1979).
banking and finance were found in North Western Europe; first in Antwerp, then in Amsterdam and Hamburg, and, after the Napoleonic wars, in London (Glamann 1979: 323 ff.) Now financial development gained momentum, and various local variants of secured and unsecured credit were added to the Italian prototypes, including different types of negotiable debt instruments. In particular, the mediaeval letter of exchange underwent important modifications and its successor, the bill of exchange, became a negotiable instrument for credit and payment. An essential prerequisite for negotiability was that personal trust in the ability and willingness of the issuer to meet his obligations had to be substituted for by some more formal – institutional – guarantee. The solution found was bill discounting. For this purpose, in the larger centres of trade and commerce municipal discount houses were chartered. The first of these was in 1609 the Amsterdamsche Wisselbank, and here as elsewhere, credit and payment for regional as well as international trade was considerably facilitated.

Yet, the bill of exchange never became a local means of payments. In Amsterdam, this function was eventually taken by the Wisselbank's guilder (florin), which was credited to personal accounts for discounted bills as well as against deposits of full-bodied coins or bullion. The Wisselbank's guilder soon became the dominant currency for trade and commerce in Amsterdam and far beyond. 18 The opportunity to settle payments by transfers denominated in a widely accepted bank deposit currency was much safer and more convenient than shipments of coinage or bullion, particularly in the face of the 'dreadful coinage mayhem' of the sixteenth and seventeenth centuries (Sombart 1916, Vol. I: 424).

As others, Adam Smith was impressed by convenience that the Wisselbank meant to trade and commerce (Smith 1879, Book 4, Chapter 3). 19 Yet, he also realised that the Wisselbank did not restrict its activity to bill discounting and the issue of bank guilder against deposits of full-bodied coinage or bullion. In fact, although the statutes of the Wisselbank required the bank's guilder to be backed by 100 per cent reserves, in practice, credit-worthy customers could transfer guilder in excess of their deposits. In other words, the Wisselbank engaged in fractional reserve banking and made personal loans.

As the beginnings of fractional reserve banking can be traced back to the Commercial Revolution and to the early goldsmith bankers of London, the Wisselbank's practice was not strictly an innovation. But now fractional reserve banking was performed by a public bank that was supervised by the authorities, and at a much larger scale than ever before. Hence, economic historians tend to identify the Wisselbank as the starting point of credit money. 20 In the seventeenth century, more than twenty public banks were chartered, modelled on the Amsterdam example, e.g. in Barcelona (1609), Hamburg (1619), Delft (1621), Nuremberg (1621), Rotterdam (1635), Stockholm (1656) as well as 1694 in London. Although the creation of money through fractional reserve banking was not covered by any of these bank's charters, it became soon accepted practice. 21 Bank money had thus become a reality.

---

18 In 1720, when the peak was reached, almost 300 local and foreign merchants held bank guilder accounts with the Wisselbank (Houtman de Smedt and Van der Wee 1993: 131).

19 See also Sombart (1916, Vol. I: 424 ff.), who cites a considerable number of observers, reflecting how contemporary merchants greeted the improvement that the new money meant to trade and commerce.

20 See e.g. Galbraith (1995: 8), according to whom the founding of the Wisselbank was nothing short of "a step that joins the history of money to the history of banking".

21 A number of economists, such as Marx (1894), Sombart (1927), de Roover (1942) and Houtman de Smedt and Van der Wee (1993), maintain that the Wisselbank strictly upheld its statutes, i.e. ensured that its bank money was backed 100 per cent. However, there is no dispute that the public banks that followed the Wisselbank's example soon established the practice of fractional reserve banking.
Another type of credit money that emerged in the modern era is the bank note, or paper money, an impersonal means of payments, which is characterised by the fact that, whatever is its face value, its commodity price is practically zero. In Europe, paper money was first issued in 1661 by the Swedish Riksbank, followed by the Bank of England in 1694 and the Bank of Scotland in 1695. From 1690, paper money was issued in New England, and in the early eighteenth century in France as well as in various Italian and German States.

From a quantity theory perspective, the numerous innovations in money, finance and the payments system during this period can be summarised as a gradual substitution of credit for commodity as a means of payments. This process had started during the Commercial Revolution, but it gained momentum in modern times. Still, until recently, credit money was confined to a relatively narrow range of economic activities, and bank money was not common in everyday life until well into the twentieth century.

The move towards credit money freed the monetary system from the material restrictions of the past. However, the downside of removing the natural and technical constraints that the antique system of full-bodied money minted from bullion had laid on the quantity of money is that without these restraints, the potential for inflation is infinite. And indeed, while moderate inflations and tinkering with coinage have been recorded since the first coins were minted, hyperinflation has exclusively affected credit money. Due to this imminent danger, opposition to uncovered paper money ('fiat money') has been fierce well into the twentieth century, both among the public and among those responsible for the stability of the monetary system.

For the nineteenth century, Peel's Bank Act of 1844 is revealing in illustrating the dilemma. It declared the notes of England as legal tender, which could be perceived as a decreed 'fiat money'. However, at the same time the Act included the obligation that two thirds of the notes in circulation had to be covered by gold in the Bank's coffers. This effectively allowed supplying liquidity to the economy in excess of the stock of bullion. But with mandatory reserves of two thirds, the public could trust that the notes were indeed convertible into gold, at least in the absence of circumstances that might trigger a run on the coffers of the Bank. Other European countries soon followed this example, which indeed proved reasonable flexible in supplying money and liquidity for the fast growing economies of that time.

With the onset of hostilities in 1914, the situation to cause a flight from paper into gold had arrived and consequently, convertibility into gold was suspended in all belligerent nations. During the interwar period, many countries attempted to or indeed restored their pre-war gold parities. However, these moves eventually proved futile when the gold standard was again largely abandoned during the Great Depression. After World War II, a modified gold standard was introduced as the basis for currencies taking part in the international monetary regime that had been approved at Bretton Woods in 1944. The US dollar was fixed to gold at a

---

22 See e.g. BORN (1972) and RITTANN (1975: 483 ff.). The oldest known paper money circulated in China from the seventh century. However, at the end of the Ming Dynasty in the seventeenth century, it perished.

23 Initially, governments issued uncovered paper money in times of crisis, war or revolution. The first European issue of government paper money occurred in pre-revolutionary France; where after the revolution the notorious 'assignats' were emitted. Other early examples are the paper monies in North America during the War of Independence, in England during the Napoleonic Wars, and the greenbacks from the US Civil War. As FRIEDMAN (1994: 45) remarks, these emergency measures without exception quickly resulted in inflation and ultimately disruption of the paper currency, after which stability had to be restored under a bullion standard.

24 See (BORN 1976: 20 ff.). Robert Peel was Prime Minister of Great Britain from 1834–35 and from 1841–46.

25 At the Bretton Woods conference in July 1944, representatives from 45 countries, dominated by the anti-Hitler coalition, agreed to a plan to restore currency convertibility.
parity of 1/35 ounces and, since no other currency was convertible into gold, it became the 'anchor currency' of the Bretton Woods system, where the other participating currencies were pegged to the dollar within a relatively narrow band of ±1 per cent. Consequently, as soon as a currency was made convertible into US dollar, it was de facto indirectly pegged to and convertible into gold. Hence, a link from paper to gold, albeit indirectly via the US dollar, was re-established for another three decades. The end of the Bretton Woods system came in 1971, when the US government suspended convertibility of the dollar into gold. Consequently, not only the dollar, but all currencies that were pegged to it, suddenly were stripped of the provision – as hypothetical as it may have been – to redeem paper into gold at a fixed rate.

The link between money and bullion, which had been established thousands of years ago, was suddenly cut and world's currencies turned into uncovered paper money. Understandably, the public – as far as people were interested in the nature of money – was concerned. At the same time, central banks were urgently looking for an alternative to 'anchor' potentially worthless 'fiat money' to a solid base. The quantity theory of money, at that time widely discussed and forcefully promoted by the influential monetarist school, delivered this device.

3 The quantity theory
3.1 The equation of exchange

Before we turn to the epistemology of the quantity theory, it will be helpful to lay out its twentieth century formalisation, the equation of exchange. Let $M$ denote the amount of money in the hands of the public, $T$ the economic transactions paid for with $M$ within a given period, $P$ the price level of $T$, and $V$ the velocity of the circulation of $M$. With these definitions, we can write down the following identity

$$M V = P T,$$

which is the textbook version of the equation of exchange associated with Irving Fisher. The right hand of the equation denotes the volume of transactions paid for, quantified in terms of their prices. With $M$ defined as the stock of means of payments at hand to settle the volume of payments given by $P T$, $V$ is the number of times per period that the stock of money has to pass from payer to payee to satisfy the equality condition, which is another way to look at the concept of the velocity of money.

$M$, $V$, $P$ and $T$ all have empirical correlates that in principle can be quantified. In practice, however, measurement of any of the four variables is fraught with formidable difficulties. Regarding $M$, for periods before the advent of bank money, the stock of bullion and full-bodied coins may serve as a useful approximation. After liquid bank liabilities have become the major means of payments, measurement is not as straightforward, which is illustrated by the fact that applied economics nowadays refers to various definitions such as M1, M2, M3 (occasionally M4), where a higher ordinal number indicates a wider aggregate. The textbook definition for M1 is legal tender (notes and coins) plus demand deposits with banks in the hands of the public. Aggregates with higher ordinal numbers consecutively add bank liabilities of lesser liquidity, implying that money's core function as a means of payments, which is the focus of M1, is amended by its quality to serve as a means for storage of wealth and financial investment. Hence, M1 is most in line with $M$ in the equation of exchange. Yet, not even the narrow aggregate M1 consists of homogenous elements, and in fact, Irving Fisher's original equation of exchange

$$M V + M' V' = P T$$

(2)
takes account of this, where "M signifies the quantity of money in circulation; V, its velocity of circulation, or rate of turnover per annum; M', the volume of deposits subject to check; and V', its velocity of circulation, or rate of turnover per annum." The distinction between two categories of means of payments, with the possibility of different velocities, adds realism to the equation, and monetary statistics nowadays allow distinguishing between cash and demand deposits, so that, for the time being, let us consider the measurement issues regarding $M$ as manageable. However, the same does not hold for $T$, which includes intermediary products and services. Since the commonly available measure of aggregate economic activity like gross domestic product (GDP, or $Y$) exclude intermediary products and services, $Y < T$. Referring to $Y$, for which useful approximations are usually available, rather than to $T$, hence requires us to define a modified identity given by

$$M V^Y = P Y^r,$$

(3)

which is the GDP version of the equation of exchange (1). Equation (3) is the norm today. $Y^r$ denotes real GDP and $V^Y$ the velocity of money related to GDP. Since $Y < T$ by definition, from (1) and (3) we can derive $V^Y = V (Y^r/T) < V$.

At first glance, the difference between the transaction and GDP versions seems minor, but it should be noted that in (1) and (2), money is strictly referred to as a means of transactions, whereas in the GDP version, $V^Y$ cannot readily be interpreted as the velocity of money. Yet, as long as the fraction $Y^r/T$ is constant, the growth rate of $V^Y$ will be the same as that of $V$, and with estimates of $Y^r/T$, we could actually compute $V^Y$ from $V$. In other words, $V^Y$ is essentially representing the same concept as $V$, implying that it is characterised by the same benefits or drawbacks. We shall return to this later.

3.2 The epistemology: important steps in the evolution of the quantity theory

3.2.1 Mediaeval thought foreshadowing the theory of money.

Scholasticism (the most prominent proponent, Thomas Aquinas, died in 1274) tried to reconcile secular aspects with normative theological considerations. Important in our context is the 'just price doctrine', which on the one hand postulates the exchange of equivalents, but on the other insists that individual incomes should secure the means of living corresponding to the social status (including the means to be able to give to charity). Prices were hence seen as phenomenon that should reflect market forces as well as normative aspects.

The first documented attempts at a positive theory of money came from Nicolai de ORESMIUS (1325–1382). ORESMIUS was the bishop of Lisieux, an astronomer, mathematician, and last but not least, advisor to Charles V on financial matters. In his tracts on economics, he mainly argued against coinage debasement (even if ordered by Charles V). In today's terminology, ORESMIUS' opinion was that reducing the bullion content of coins that maintained the same face value had led to a corresponding increase in prices, so that in the end, the king had not gained anything by the debasement. For this, see GORDON (1987), and others. According to MUNDELL'S (1998: 3) judgment, ORESMIUS' "De Moneta" (written during the reign of Charles V, when no less than 86 coin devaluations took place) is the

\[\text{References}\]

\[\text{Note:}\]

26 FISHER (1911: 38). Note that Fisher did not consider demand deposits (which are not directly convertible into bullion) as proper money. Yet, inclusion of $M'$ into the equation of exchange proves that he considered it a fully functional means of payment. In 1911, the ancient identity 'money = bullion' still reigned in economics.

27 Setting prices clearly different from the 'just' price was a case for the secular jurisdiction, so that the just price doctrine certainly had normative power.

28 For this, see GORDON (1987), and others. According to MUNDELL'S (1998: 3) judgment, ORESMIUS' "De Moneta" (written during the reign of Charles V, when no less than 86 coin devaluations took place) is the
Natura, Jure, et Mutationibus Monetarum" of 1355 is probably the first statement of the quantity theory. Now it is important to notice that this publication came three years after Europe's first plague epidemic (1347–1352). This was an epidemic of horrendous impact. It is estimated that in these five years, the population of Europe was reduced by a third (PiRENNE 1986: 186). The resulting decline of aggregate economic activity was – in today's terminology – a dramatic negative supply shock. Accordingly, the price increases that attracted ORESMIUS' attention were only partly due to coinage debasement; the supply shock was probably the major factor. Though it may seem an anachronism to discuss earlier contributions in terms of the equation of exchange, the effect of the plague on the price level can easily be captured by the GDP version of the equation of exchange (3) with time subscripts

\[ M_t V_t^Y = P_t Y_t. \] (3a)

Let \( t-1 \) denote the time before the outbreak of the plague, so that \( Y_t < Y_{t-1} \). Now, in a currency system where money is bullion, the stock of money \( M \) is practically constant in the short term. Hence, without a dramatic decline in \( V \), it follows that \( P_t > P_{t-1} \). While we can speculate that some hidden money was permanently lost when people fell victim to the plague, and that the survivors tended to hoard money more then previously (leading to a decline in \( V \)), the decline of \( Y \) to appears to have outweighed the declines of \( M \) and \( V \). The price increases that ORESMIUS observed were then caused by a sharp decline in transactions that had to be settled in cash in an economy that had inherited a stock of money from a far more numerous population. Of course, this does not contradict the debasement story. As can be seen from the equation of exchange, without a decline of \( V \), both an increase in the nominal supply of money \( M \) and a decrease of \( Y \) will result in a higher price level \( P \).

### 3.2.2 Jean Bodin

The classical quantity theory is referred back to Jean BODIN (1530–1596) "Réponse aux paradoxes de Monsieur de Malestroict touchant l'enchérissement de toutes choses" (1566).

MALESTROICT had blamed widespread practice of coin clipping, i.e. debasing of coins by the public, for the general acceleration in prices "de toutes choses", which had become noticeable by the middle of the 16th century. BODIN took a different stance. He argued that the price level had risen together with the stock of money, or more precisely, with the amount of bullion available for monetary purposes. BODIN witnessed the results of fifteenth and sixteenth century shipments of bullion from the Americas to Europe and the resulting 'Price Revolution', and was able to draw the right conclusion about the link between the two. The classical quantity theory of money is thus a reflection of the history of money in the early modern era. It preceded the advent of classical economics by two centuries.

### 3.2.3 Classical theory

The classic's version of the quantity theory is the proportionality theorem, which postulates that the price level reacts proportional to changes to the stock of money. For the classics, the quantity theory was a major device to overcome the doctrine of mercantilism, according to which the money supply, i.e. the stock of bullion within state borders, is the deciding determinant of economic prosperity. Accordingly, the proportionality theorem can be

---

first traditional version of GRESHAM'S (1519-1579) Law, and the most important work on the theory of money in the period prior to the 16th century.
documented by numerous citations from the classical literature. For our purpose, it will be sufficient to refer to a few examples. David Hume (1711–1776) explains:

"If we consider any one kingdom by itself, 'tis evident, that the greater or less the plenty of money is of no consequence; since the prices of commodities are always proportion'd to the plenty of money ..." (1752: 127)

After having presented the proportionality theorem as self-evident, Hume explains that the acquisition of bullion would be to the detriment of foreign trade, as the domestic export goods would become more expensive. He concludes:

"In my opinion, 'tis only in this interval or intermediary situation, betwixt the acquisition of money and rise of prices, that the increasing quantity of gold and silver is favourable to industry." (op. cit.: 133)

With these reflections, known today as the 'price specie flow mechanism', Hume formulated the theory of the balance of trade in an international bullion (or gold) standard, where the flow of 'specie' would tend to equilibrate the balance of trade.

Some 50 years later, in the 'bullionist controversy', David Ricardo (1772–1823), perhaps the most influential classical theoretician, was a devoted proponent of the quantity theory with a strictly bullionist concept of money. Consequently, he advocated a strict correspondence between the issue of banknotes and the bullion in the issuing banks' vaults (Ricardo 1809).

Finally, let us cite the last great advocate of classical economics, John Stuart Mill. It is, as he writes, self-evident

"that the mere introduction of a particular mode of exchanging things for one another, by first exchanging a thing for money, and then exchanging the money for something else, makes no difference in the essential character of transactions. It is not with money that things are really purchased. … There cannot, in short, be intrinsically a more insignificant thing, in the economy of society, than money; except in the character of a contrivance for sparing time and labour." (1848: 8 f.)

"If the whole money in circulation was doubled, prices would be doubled. If it was only increased one-fourth, prices would rise one-fourth." (op. cit.: 15)

"So that the value of money, other things being the same, varies inversely as its quantity: every increase of quantity lowering the value, and every diminution raising it, in a ratio exactly equivalent." (op. cit.: 16)

More formally, the classical quantity theory's proportionality theorem can be derived from the equation of exchange (3) as follows:

\[ M \cdot V \cdot Y = P \cdot Y \Rightarrow P = (V \cdot Y / Y) \cdot M \]

\[ \Rightarrow P = f (M). \]  

(4)

(4')

In the history of economic thought, the classical quantity theory is a theory of the general price level, as formalised in equation (4'). Assuming an exogenously given level of potential output and full employment, in the short term, \( Y \) can be treated as a constant. Assuming further that the income velocity of money \( V \) is exogenous and constant, equation (4') can be specified as \( P = a \cdot M \), where \( a \) is the proportionality constant. Notice that under these assumptions, money is 'neutral' in the sense that real output is completely independent of the stock or changes to the stock of money. Money affects the price level, and nothing else.29

---

29 However, the so-called dichotomy between the real and the monetary sectors poses a theoretical problem; namely how a change in the money supply would lead to a change in the price level. As Patinkin (1965: 162 ff.) has shown, the missing link between money and the price level is the real balance effect. This rests on the assumption that the desired real balances \( M/P \) are constant (see Friedman, 1969; Patinkin, 1965) Thus, when the money supply increases, people find their money balances above preferred levels, and will
To derive a theory of the price level (4') with empirical content from the identity stated by the equation of exchange (3), the following conditions must be met:

- \( M \) is an exogenous variable and independent of \( P, Y \) and \( VY \). (Otherwise, there is no conclusive causality.)
- \( VY \) is an exogenous variable; in particular, it is independent of \( P, M, \) and \( Y \). (Otherwise, the link between money and prices is indeterminate.)
- \( Y \) is independent of \( M \), i.e. neutrality of money holds. (To the degree to which real output is a positive function of the money supply, extra money would raise output but not prices.)

### 3.3 The quantity theory as theory of the demand for money

The equation of exchange can serve to derive a simple money demand function. Let us define \( k \equiv 1/VY \) and \( Y \equiv PYr \). Equation (3) can than be rearranged into the 'Cambridge Equation'

\[
M = k Y \quad (5)
\]

\[
\Rightarrow M^d = f(Y) \quad (5')
\]

The Cambridge Equation is a money demand function, a behavioural equation, where \( M \) is interpreted as the demand for money \( M^d \) and the 'Cambridge constant' \( k \) shows the proportion of real income that individuals want to hold in liquid form. As \( k \) is the inverse of \( V \), stability of \( VY \) in (3) implies a stable money demand function (5').

### 3.3.1 The monetarist liquidity preference theory.

The success of the monetarist liquidity preference theory among economic theorists and practitioners around 1970 reflects the political and economic situation of these days. 'Vulgar Keynesianism' with particular emphasis on deficit spending to combat underemployment, had taken root as orthodoxy, whereas monetary policy, at best, was allowed an auxiliary role. At that time, it became apparent that inflation picked up along with unemployment, which went against the predictions of the theory. Moreover, with the end of the Bretton Woods system, monetary policy suddenly had to take on new functions, so a revised theoretical foundation the monetary policy became the order of the day.

In these circumstances, Friedman's version of quantity theory was readily accepted as the appropriate paradigm, leading to new directions in macroeconomics. What made Friedman's so attractive? It offered clear guidance to monetary policy by directly addressing the main weaknesses of the classical quantity theory, namely ignorance of the velocity of money and its determinants. Let us hence look at his seminal work in some detail. Friedman (1956) approached the demand for money like the demand for any good would be analysed in a microeconomics textbook. \( M^d \) is hence a function \( f() \) of

---

30 On the aggregate level, stability of (5') is based on the premise that \( k \) reflects the amount of money \( M \) that is required to process the volume of monetary transactions associated with \( Y \), when payment habits as well as the technical and institutional features of the payments system are considered as given.
the price of the relevant good; here the price level \( P \), where a high price level implies a low 'price of money' so that \( f_P > 0 \);

- the prices of substitutive goods; here interpreted as the real return of other financial claims, such as yields from fixed-interest bonds \( r^b \), shares \( r^e \) as well as the yield from holding real assets, where \( f^b, b < 0 \) and \( f^e, e < 0 \). The yield from holding real assets rather than financial claims corresponds to the rate of inflation, so that \( f^b(P/dP/dt) < 0 \);

- the budget constraint; here represented by FRIEDMAN’S famous permanent income \( Y \), where \( f_Y > 0 \);

- the degree of liquidity of an individual’s total assets; here Friedman defines \( w \) as the ratio of non-human assets to human capital. As human capital is less liquid as non-human assets, it follows that \( f_w < 0 \);

- preferences; here FRIEDMAN writes down \( u \), but remains vague, so that this should be regarded as a catch-all term to make sure that the money demand function lists all arguments.

FRIEDMAN’S money demand function hence reads as follows:

\[
M_d = f(P, r^b, r^e, (1/P)(dP/dt), Y, w, u). \tag{6}
\]

Now, as monetarism assumes the people are free from money illusion, the nominal demand for money \( M \) is linearly homogeneous in nominal variables (here: \( P \) and \( Y \)), so that

\[
M_d/Y = f(P/Y, r^b, r^e, (1/P)(dP/dt), 1, w, u) \tag{7}
\]

\[
\Longleftrightarrow M_d = f(P/Y, r^b, r^e, (1/P)(dP/dt), 1, w, u) Y. \tag{8}
\]

A comparison between equation (8) and the Cambridge Equation (5) reveals that the function \( f(P/Y, r^b, r^e, (1/P)(dP/dt), 1, w, u) \) is as an elaboration of the Cambridge \( k \).

Now, if \( f \) can empirically be handled as a stable function of traceable macroeconomic variables, FRIEDMAN’S neo-quantity theory can be called upon as the basis for an economic policy that is focussing on control of the money supply: Solving (7) for \( Y \) and referring to the equilibrium condition \( M_d = M^e \), we get

\[
Y = M^e / [f(P/Y, r^b, r^e, (1/P)(dP/dt), 1, w, u)]. \tag{9}
\]

With a stable money demand function \( f \) and the money supply \( M^e \) exogenously set by the monetary authority, nominal output \( Y \) can be directly controlled. Hence, from the monetarist point of view, the money supply is, or should be, the central variable of economic policy.

3.4 Quantity theoretical inflation theory

Yet another approach to derive a positive theory from the equation of exchange is to transform it into growth rates and to solve for the growth rate of the price level, i.e. inflation. Taking logarithms and time derivates of the equation (3) following \( t \) result in

\[
d(ln M)/dt + d(ln V^Y)/dt = d(ln P)/dt + d(ln Y^e)/dt \tag{10}
\]

\[
(1/M)(dM/dt) + (1/V^Y)(dV^Y/dt) = (1/P)(dP/dt) + (1/Y^e)(dY^e/dt) \tag{10'}
\]

\[
g(M) + g(V^Y) = g(P) + g(Y^e) \tag{10''}
\]

\[
\pi = g(M) + g(V^Y) - g(Y^e) \tag{10'''}
\]
where \( g(X) \) represents the growth rate of a variable \( X \) and \( \pi \equiv g(P) \). As can be inferred from equation (10''), if the velocity of money is a constant, hence \( g(Y^V) = 0 \), any growth rate of the money supply exceeding that of real output, leads to inflation. This is the basis of perhaps the most famous, and most often quoted of Milton Friedman's statements:

"inflation is always and everywhere a monetary phenomenon in the sense that it is and can be produced only by a more rapid increase in the quantity of money than in output." (Friedman 1994: 49)

Empirically the quantity theoretical explanation of inflation can refer to a firmly established long-term correlation between growth of the money supply and increases in the price level.\(^{31}\) Apart from this, proponents of the theory point to evidence that this correlation not only holds in longitudinal studies but also for cross sections of countries.\(^{32}\)

It is in order to note that the observed correlation between \( \pi \) and \( g(M) \) does not prove that causality runs from \( g(M) \) to \( g(M) \); \( \pi \) could also be the cause and \( g(M) \) the consequence, especially in hyper-inflationary situations. Moreover, this inflation theory has to assume super-neutrality of money, i.e. the independence of the growth rate \( g(Y^r) \) from \( g(M) \). The question of the super-neutrality of money in the inflation process has not as yet been empirically clarified.\(^{33}\) But if \( g(Y^r) \) reacts positively to a change in \( g(M) \), the correlation between \( g(M) \) and \( \pi \) is not as narrow as the equation of exchange suggests.

3.5 Application: The Frankfurt monetary targets, 1974 to the present

From 1974 until the end of its responsibility for monetary policy in Germany in the end of 1998, the German Bundesbank referred to a monetary target that can readily be derived by solving equation (12'') for \( g(M) \), which results in

\[
g(M) = \pi + g(Y^r) - g(V^Y). \tag{12}
\]

Let us take as an example the Bundesbank's last monetary target, i.e. for 1998. To arrive at a number for \( g(M) \), it referred to a forecast for the real GDP growth rate of 2.0 per cent, a "medium-term price target" (implicitly an inflation target) of 1.5 to 2.0 per cent, and a decline of 1.0 per cent in the velocity of money.\(^{34}\) This gives a target growth rate of 4.75 per cent. While the Bundesbank had initially communicated a point target, it later switched to a target band that was three percentage points wide. For 1998, this would have resulted in a target ranging from 3.25 to 6.25 per cent for the growth rate of M3, which was the Bundesbank's preferred monetary aggregate in the later years. In 1998, the belief in the accuracy of this arithmetic, that had previously been followed exactly, obviously had slightly faded, so that the Bundesbank left the 1997 target corridor of 3.0 to 6.0 unchanged instead of adjusting it down 25 percentage points, as equation (12) had given 5.0 per cent for 1997.

---

32 Schwartz (1973: 267) e.g. reports a correlation between \( \pi \) and \( g(M/Y) \) from 1952–1969 in a sample of 40 countries as high as 0.94. Notice however, that recent studies find that this result in mainly driven by a limited number of high inflation countries. We shall come back to this point below.
33 Traditionally, economics would not usually assume super-neutrality of money, as individuals easily err about the rate of inflation. Contrary to this view, 'Neoclassical Macroeconomics' today frequently refers to rational expectations of the expected rate of inflation, and money is therefore not only neutral, but also super-neutral; see McCallum (1990).
34 The latter conjecture is consistent with Friedman's 'luxury goods hypothesis', according to which the preference to hold money is increasing disproportionately with income. Based on a meta-analysis of some 500 studies on money demand, Knell and Stix (2005) conclude that money is indeed a superior good. First estimates for the euro area (Brand et al. 2002) arrive substantially at the same conclusion, namely a trend towards moderate decline in the velocity of M3 around 0.5 to 1.0 per cent.
The ECB, which took up its responsibilities in 1999, initially closely stuck to the practice of the Bundesbank and performed yearly calculations according to equation (12). When it downgraded its second pillar – 'monetary analysis' – in 2003, it decided to abandon the yearly review of the reference value. Since then, the calculus that results in its reference value of \( g(M) = 4.5 \) per cent refers to \( g(Y) = 2.25 \) per cent and \( g(Y)^{\ast} = -0.75 \) per cent. According to equation (12), the ECB's implicit inflation target is thus centred around 1.5 per cent.

While the ECB continues to declare monetary analysis as its second pillar, it certainly did not restrict the growth of monetary aggregates in line with the reference value of 4.5% that results from the quantity theory. Apart from a trough in the beginning of 2001, the annual growth rates of both M1 and M3 were always –and at times considerably – above this target.\(^{35}\) Clearly, the ECB does not conduct a rule-based monetary policy according to equation (12). Monetary targeting has been abandoned in Frankfurt too.

4 Is the quantity theory obsolete? An empirical assessment

If the quantity theory holds, as long as the velocity of money remains stable, the rate of inflation should equal the difference between the growth rates of money and real GDP. Is this still an adequate tool to explain the relationship between money and inflation in recent times? To confront this hypothesis with the data, we refer to a panel comprising yearly data from 1991 to 2005 across 105 countries, so that equation (12) is amended by time and country subscripts:

\[
\pi_{it} = g(VY_{it}) + g(VY_{it}) + g(M_{it}) - g(Y_{it}).
\]  

Moreover, we specify our reduced form as a fixed-effects model, allowing for year-specific as well as country-specific effects, including particularities of \( g(VY_{it}) \), for which we do not have any direct measure. The regression that follows from this is hence

\[
\pi_{it} = \beta_{i} + \beta_{t} + \beta_{1} g(M/Y_{it}) + \epsilon_{it},
\]

or

\[
\pi_{it} = \beta_{0} + \beta_{1} g(M/Y_{it}) + \epsilon_{it},
\]

where equation (14) represents the fixed effects model and equations (14') the same without fixed effects. Our data are taken from the World Bank's World Development Indicators (WDI), which include adequate specifications of all variables needed to estimate this model.\(^ {36}\)

The endogenous variable \( \pi \) is represented by inflation, defined as the annual growth rate of the GDP deflator in per cent, and excess money growth \( g(M/Y) \) is the difference between the annual growth rate of M2 in per cent of money and the annual percentage growth rate of real GDP, quantified at market prices based on constant local currencies.\(^ {37}\)

The sample selection is based on data availability. The WDI comprise yearly series, starting in 1971. At the time of conducting this analysis, the last data point in the WDI was 2005. As can be expected for large samples of countries, the series are affected by missing

\(^{35}\) The latest figure is for November 2007, when the annual growth rate of M3 was as high 12.6 per cent.

\(^{36}\) Our data are taken from the online version (http://media.worldbank.org/secure/wdi/qquery.htm).

\(^{37}\) More specifically, the World Development Indicators definition of GDP is "the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources" and M2 (money and quasi money) is "the sum of currency outside banks, demand deposits other than those of the central government, and the time, savings, and foreign currency deposits of resident sectors other than the central government ... it corresponds to lines 34 and 35 in the International Monetary Fund's (IMF) International Financial Statistics (IFS). The change in the money supply is measured as the difference in end-of-year totals relative to the level of M2 in the preceding year."
values, where a balanced sample with complete time series for all countries will include successively more countries when we choose a shorter time period. Regarding the variables required to estimate equation (14), the number of countries with complete monotonically increases from \( n = 80 \) with 1971 as starting date to a maximum of \( n = 105 \) with 1991 as the first year. Fortunately, four our purpose, which is to assess the performance of the quantity theory when confronted with recent data, the missing values for the 1970s and 1980s are not particularly troublesome, so that we choose broad coverage rather than longer time series. With yearly observations from 1991 to 2005 and 105 countries, we hence refer to a balanced panel that covers 1575 observations for \( \pi \) and \( g(M/Y) \).

Before we turn to our analysis, let us take a brief look at the core variable: inflation. The evolution of inflation rates from 1971 to 2005 as reflected by mean median inflation across the increasing sample of countries covered by the WDI is shown in figure 1. The most striking finding from this figure is the very high mean inflation rates above 150 percent in the second half of the 1980s, which result from hyper-inflation in a number of countries. Somewhat harder to detect, as hyper-inflation dominates the picture, is the general acceleration of inflation around the mid-seventies, which in world-wide perspective lasted well up to the mid-1990s.

*Figure 1: World inflation since 1971, median inflation rate across countries, 80 \( \leq n \leq 105 \)*
Figure 2 focuses on the years from 1991, which constitute our sample period. Higher average than median inflation is still indicating skewness in the distribution, but this is now much less pronounced than during the earlier years. From 1995 on, a new regime emerges where both mean and median inflation are moderate and evolving without any clear trend.

The data points reflected in figure 2 are now regressed on excess money growth \(g(M/Y)\) as specified in the reduced form. The results are given in table 1. Column 1 is the fixed country and year effects model from equation (14), column 2 a fixed country effects model, column 3 a country fixed year effects model, column 4 a model without fixed effects but with a linear time trend, and column 5 the OLS specification from equation (14').

<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\beta_i) (df = 104)</td>
<td>(F = 0.90)</td>
<td>(F = 0.95)</td>
<td>(F = 0.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(p = 0.76)</td>
<td>(p = 0.64)</td>
<td>(p = 0.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\beta_t) (df = 14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(F = 0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(p = 0.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time trend</td>
<td></td>
<td></td>
<td>(-0.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>((-1.21))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\beta_0)</td>
<td></td>
<td></td>
<td></td>
<td>(3.56)</td>
<td>((-1.75))</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>((0.71))</td>
<td>((-0.74))</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>(1.013)</td>
<td>(1.017)</td>
<td>(1.000)</td>
<td>(1.002)</td>
<td>(1.004)</td>
</tr>
<tr>
<td></td>
<td>((40.9))</td>
<td>((44.6))</td>
<td>((45.0))</td>
<td>((45.2))</td>
<td>((45.6))</td>
</tr>
<tr>
<td>(\beta_1 \neq 1)</td>
<td>(p = 0.61)</td>
<td>(p = 0.49)</td>
<td>(p = 0.99)</td>
<td>(p = 0.95)</td>
<td>(p = 0.85)</td>
</tr>
<tr>
<td>(N)</td>
<td>1575</td>
<td>1575</td>
<td>1575</td>
<td>1575</td>
<td>1575</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.60</td>
<td>0.60</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
</tr>
<tr>
<td>Adjusted (R^2)</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
<td>0.57</td>
</tr>
</tbody>
</table>
As the coefficient of interest, $\beta_1$, is practically unaffected by changes to the specification, the result so far appear robust too. Notably, $\beta_1$ is very close to unity in any specification from table 1, and as the p-values in the row below indicate, there is no support for the hypothesis $\beta_1 \neq 1$. Obviously, these results are perfectly consistent with the proportionality theorem.

This seems too good to be true. Therefore, let us consider whether these results are driven by a limited number of influential observations. Different heuristics and ad hoc procedures are suggested for this. A method to ensure inter-subjectivity is to look at the studentised residuals (STRES), which are asymptotically t-distributed. Taking a low significance level of $p = 0.1$, the tabulation in KLEINBAUM ET AL. (1988: 661) gives a critical value of $|\text{STRES}| \geq 3.58$, which implies that six observations should be dropped from the fixed effects regression.38 The results are given in table 2.

Compared to the reference model, we observe an increase in $R^2$ from 0.60 to 0.71 and a decline in the point estimate of $\beta_1$ from 1.01 to 0.66, which is now significantly lower than unity. Moreover, the fixed effects are now jointly significant, implying that apart from excess money growth, both world inflation relating to given years, as well as country specific inflation, are important factors in explaining the inflation rate in a particular country $i$ at time $t$.

<table>
<thead>
<tr>
<th>Method</th>
<th>N</th>
<th>$\beta_1$</th>
<th>t</th>
<th>$p$</th>
<th>$\beta_1 \neq 1$</th>
<th>Fixed effects (df ≤ 118)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>Fixed effects</td>
<td>1575</td>
<td>1.01</td>
<td>40.9</td>
<td>0.61</td>
<td>0.90</td>
<td>0.76</td>
<td>0.60</td>
</tr>
<tr>
<td>$</td>
<td>\text{STRES}</td>
<td>&lt; 3.58$</td>
<td>1569</td>
<td>0.66</td>
<td>41.5</td>
<td>&lt; 0.01</td>
<td>2.21</td>
</tr>
<tr>
<td>$</td>
<td>\text{STRES}</td>
<td>&lt; 1.96$</td>
<td>1551</td>
<td>0.72</td>
<td>36.8</td>
<td>&lt; 0.01</td>
<td>1.79</td>
</tr>
<tr>
<td>Bounded influence</td>
<td>1575</td>
<td>0.74</td>
<td>44.0</td>
<td>&lt; 0.01</td>
<td>3.18</td>
<td>&lt; 0.01</td>
<td>0.75</td>
</tr>
</tbody>
</table>

PINDYCK and RUBINFELD (1991: 170) favour the stricter $2\sigma$-criterion. They suggest to regard an observation as outlier when, in absolute terms, the studentised residuals exceed 1.96. In our case, this concerns an additional 18 observations.39 As table 2 shows, eliminating these observations as well does not in qualitative terms change the results compared to the KLEINBAUM ET AL. proposal. As with the 3.58 cut-off value, the proportionality theorem does not hold any longer, and year as well as country specific effects are shown to be important.

38 The six outliers are Nicaragua 1991 (33.41), Guinea 1991 (14.35), Brazil 1993 (12.36), 1994 (11.31) and 1992 (6.42) as well as Guinea 2000 (4.27), with their corresponding $|\text{STRES}|$ shown in brackets.

39 The additional 18 outliers are Zimbabwe 2005 (3.04), Nicaragua 1996 (2.58), 1997 (2.56) and 1995 (2.36), Guinea 1999 (2.35), Nicaragua 1998 (2.31), Suriname 1994 (2.27), Nicaragua 2002 (2.23), Brazil 1991 (2.21) as well as Nicaragua 2004 (2.19), 2003 (2.18), 1999 (2.16), 1993 (2.12), 1994 (2.10), 2000 (2.05), 2001 (2.05), 2005 (2.04) and 1992 (2.04), where as above, the $|\text{STRES}|$ are shown in brackets.
An alternative to eliminating influential observations is to limit their influence. To this end, one can refer to the DFFIT statistics, which measures the change in the predicted value of the endogenous variable when the $i$-th observation is removed from the sample. On this basis, and with critical values from Maddala (1992), we run a full sample weighted least squares 'bounded influence' regression where the weights are $w_i = 0.34/|\text{DFFIT}_i|$ when $|\text{DFFIT}| \geq 0.34$ and $w_i = 1$ otherwise.\footnote{This procedure results in weight below unity for 1059 of the 1575 observation in the sample. Unlike the $|\text{STRES}|$ criteria that affected only a fraction of the sample, the 'bounded influence' criterion hence relates to more than two thirds on the observations in our panel.} The result is shown in the last row of table 2. Again, compared to the Kleinbaum et al. procedure, the results are qualitatively unchanged. This implies that, based on these sensitivity analyses, we clearly have to reject the proportionality theorem, as $\beta_1 = 1$ does not hold as soon as the six most pronounced outliers are removed from the sample. Nevertheless, we should not abandon the quantity theory altogether. The sensitivity analyses show that $\beta_1 < 1$, but they also show unambiguously that $\beta_1 > 0$, which conforms to the prediction that excess money supply creates inflation, though not in strict proportion.

We saw a pronounced decline in $\beta_1$ as we dropped some high inflation observations from the sample. Is this a clue to some structural regularity? To investigate whether this is the case, we now run a number of regressions where we progressively reduce the sample by eliminating all observations that exceed a certain inflation cut-off value. To determine the appropriate cut-off values, let us look the observations that reflect the highest inflation rates in our panel. All observations with inflation exceeding 100 per cent are put together in table 3, along with the other measures that enter into equation (12).

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>$\pi$</th>
<th>$g(Y)$</th>
<th>$g(M)$</th>
<th>$g(M) - g(Y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicaragua</td>
<td>1991</td>
<td>4523.7</td>
<td>-0.2</td>
<td>1519.6</td>
<td>1519.8</td>
</tr>
<tr>
<td>Brazil</td>
<td>1994</td>
<td>2239.1</td>
<td>5.9</td>
<td>1279.7</td>
<td>1273.8</td>
</tr>
<tr>
<td>Brazil</td>
<td>1993</td>
<td>1996.6</td>
<td>4.9</td>
<td>2854.0</td>
<td>2849.1</td>
</tr>
<tr>
<td>Brazil</td>
<td>1992</td>
<td>968.5</td>
<td>-0.5</td>
<td>1551.9</td>
<td>1552.4</td>
</tr>
<tr>
<td>Suriname</td>
<td>1994</td>
<td>439.0</td>
<td>3.3</td>
<td>204.7</td>
<td>201.5</td>
</tr>
<tr>
<td>Brazil</td>
<td>1991</td>
<td>415.3</td>
<td>1.3</td>
<td>636.7</td>
<td>635.4</td>
</tr>
<tr>
<td>Peru</td>
<td>1991</td>
<td>379.9</td>
<td>2.2</td>
<td>230.6</td>
<td>228.4</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2004</td>
<td>350.2</td>
<td>-4.2</td>
<td>229.3</td>
<td>233.5</td>
</tr>
<tr>
<td>Suriname</td>
<td>1995</td>
<td>273.9</td>
<td>1.1</td>
<td>181.5</td>
<td>180.4</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2003</td>
<td>262.8</td>
<td>-10.4</td>
<td>430.1</td>
<td>440.5</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2005</td>
<td>240.3</td>
<td>-7.1</td>
<td>532.7</td>
<td>539.8</td>
</tr>
<tr>
<td>Romania</td>
<td>1993</td>
<td>227.4</td>
<td>1.5</td>
<td>143.3</td>
<td>141.8</td>
</tr>
<tr>
<td>Romania</td>
<td>1992</td>
<td>200.1</td>
<td>-8.8</td>
<td>75.4</td>
<td>84.2</td>
</tr>
<tr>
<td>Romania</td>
<td>1991</td>
<td>194.9</td>
<td>-12.9</td>
<td>102.2</td>
<td>115.1</td>
</tr>
<tr>
<td>Suriname</td>
<td>1993</td>
<td>159.0</td>
<td>-7.3</td>
<td>65.7</td>
<td>72.9</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2002</td>
<td>150.5</td>
<td>-4.4</td>
<td>191.7</td>
<td>196.1</td>
</tr>
<tr>
<td>Romania</td>
<td>1997</td>
<td>147.3</td>
<td>-6.1</td>
<td>105.0</td>
<td>111.1</td>
</tr>
<tr>
<td>Sudan</td>
<td>1996</td>
<td>146.8</td>
<td>5.9</td>
<td>65.6</td>
<td>59.7</td>
</tr>
<tr>
<td>Romania</td>
<td>1994</td>
<td>138.9</td>
<td>4.0</td>
<td>138.1</td>
<td>134.1</td>
</tr>
<tr>
<td>Argentina</td>
<td>1991</td>
<td>133.0</td>
<td>12.7</td>
<td>141.3</td>
<td>128.7</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>1991</td>
<td>128.8</td>
<td>2.4</td>
<td>76.2</td>
<td>73.9</td>
</tr>
<tr>
<td>Guyana</td>
<td>1991</td>
<td>126.7</td>
<td>6.0</td>
<td>73.4</td>
<td>67.4</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>2001</td>
<td>121.7</td>
<td>-2.7</td>
<td>128.5</td>
<td>131.2</td>
</tr>
</tbody>
</table>
The highest yearly inflation rate reported in the WDI for the 1991–2005 period was as high as 4524 per cent (Nicaragua in 1991), followed by 2239, 1997 and 969 per cent (Brazil in 1994, 1993 and 1992). All remaining very high growth rates are below 500 per cent, which is still impressive, but way behind the four highest rates that roughly fell into a range of 1000 to 4500 per cent. Clearly, these observations can exercise a lot of leverage and hence cause misleading results if pooled together with less extreme cases. Therefore, we shall set the cut-off values in a way that ensures that the extreme observations are eliminated from the sample one by one. Table 3 suggests cut-off values at $\pi \leq 4000\%$, $\pi \leq 3000\%$, $\pi \leq 2000\%$, $\pi \leq 1000\%$ and $\pi \leq 500\%$, and after this we shall lower the threshold in stepwise by 1 percentage point.

With these cut-off values, we repeatedly run the fixed effects regression according to equation (14), while the observations with the highest inflation rates are one by one removed from the sample, which will allow us to trace the effect of these observations on the pooled sample estimations. Selected results from this exercise are summarised in table 4.
Table 4

<table>
<thead>
<tr>
<th>Inflation cut-off (π ≤ x)</th>
<th>N</th>
<th>β₁</th>
<th>t</th>
<th>p</th>
<th>β₁ &gt; 0</th>
<th>p</th>
<th>β₁ ≠ 1</th>
<th>Fixed effects (df ≤ 118)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1575</td>
<td>1.01</td>
<td>40.9</td>
<td>&lt; 0.01</td>
<td>0.61</td>
<td>0.9</td>
<td>0.76</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>3000%</td>
<td>1574</td>
<td>0.68</td>
<td>52.4</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>1.3</td>
<td>&lt; 0.01</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>2000%</td>
<td>1573</td>
<td>0.58</td>
<td>70.6</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>2.2</td>
<td>&lt; 0.01</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>1000%</td>
<td>1572</td>
<td>0.41</td>
<td>36.9</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>3.3</td>
<td>&lt; 0.01</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>500%</td>
<td>1571</td>
<td>0.26</td>
<td>19.9</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>4.6</td>
<td>&lt; 0.01</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>100%</td>
<td>1547</td>
<td>0.05</td>
<td>7.5</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>10.7</td>
<td>&lt; 0.01</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>1500</td>
<td>0.03</td>
<td>6.6</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>10.2</td>
<td>&lt; 0.01</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>1335</td>
<td>0.02</td>
<td>5.2</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>7.6</td>
<td>&lt; 0.01</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>15%</td>
<td>1256</td>
<td>0.04</td>
<td>5.9</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>6.3</td>
<td>&lt; 0.01</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>14%</td>
<td>1230</td>
<td>0.03</td>
<td>5.8</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>6.1</td>
<td>&lt; 0.01</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>13%</td>
<td>1201</td>
<td>0.03</td>
<td>5.9</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>5.9</td>
<td>&lt; 0.01</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>12%</td>
<td>1161</td>
<td>0.03</td>
<td>5.8</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>5.6</td>
<td>&lt; 0.01</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>11%</td>
<td>1122</td>
<td>0.07</td>
<td>6.8</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>5.2</td>
<td>&lt; 0.01</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>1086</td>
<td>0.06</td>
<td>6.5</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>4.8</td>
<td>&lt; 0.01</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>9%</td>
<td>1023</td>
<td>0.06</td>
<td>6.1</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>4.5</td>
<td>&lt; 0.01</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>8%</td>
<td>952</td>
<td>0.05</td>
<td>5.0</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>3.7</td>
<td>&lt; 0.01</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>7%</td>
<td>872</td>
<td>0.05</td>
<td>4.9</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>3.1</td>
<td>&lt; 0.01</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>6%</td>
<td>795</td>
<td>0.04</td>
<td>4.2</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>2.8</td>
<td>&lt; 0.01</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>710</td>
<td>0.04</td>
<td>4.1</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>2.8</td>
<td>&lt; 0.01</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>4%</td>
<td>602</td>
<td>0.04</td>
<td>3.4</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>2.9</td>
<td>&lt; 0.01</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>3%</td>
<td>481</td>
<td>0.04</td>
<td>3.4</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>2.9</td>
<td>&lt; 0.01</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>2%</td>
<td>341</td>
<td>0.04</td>
<td>2.4</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>2.7</td>
<td>&lt; 0.01</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>1%</td>
<td>213</td>
<td>0.04</td>
<td>1.9</td>
<td>&lt; 0.01</td>
<td>&lt; 0.01</td>
<td>1.7</td>
<td>&lt; 0.01</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>126</td>
<td>0.07</td>
<td>2.0</td>
<td>0.03</td>
<td>&lt; 0.01</td>
<td>0.9</td>
<td>0.60</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>-1%</td>
<td>88</td>
<td>0.09</td>
<td>2.0</td>
<td>0.03</td>
<td>&lt; 0.01</td>
<td>0.8</td>
<td>0.74</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>-2%</td>
<td>50</td>
<td>0.18</td>
<td>2.0</td>
<td>0.04</td>
<td>&lt; 0.01</td>
<td>0.8</td>
<td>0.71</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>-3%</td>
<td>39</td>
<td>0.18</td>
<td>1.3</td>
<td>0.12</td>
<td>&lt; 0.01</td>
<td>0.7</td>
<td>0.75</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>-4%</td>
<td>25</td>
<td>0.04</td>
<td>0.4</td>
<td>0.39</td>
<td>&lt; 0.01</td>
<td>15.2</td>
<td>0.56</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 shows that as soon as the Nicaragua 1991 observation (π = 4523% > 3000%) is removed from the sample, β₁ declines from 1.01 to 0.68. The support for the proportionality theorem in the full panel hence depends on this single and extreme observation. The next hyper-inflation observations work in the same direction; as they are eliminated from the sample, β₁ drops to 0.58, 0.41 and 0.26. With further high-inflation observations removed, β₁ continues to decline monotonically below 0.10. After the threshold is lowered under 100 per cent, β₁ appears to stabilise and continues to move in positive territory. Moreover, though the reduction in magnitude is impressive, β₁ remains significantly positive until the sample size
tion in magnitude is impressive, $\beta_1$ remains significantly positive until the sample size gets smaller than 50 and the standard errors increase at inflation rates below $-2\%$, which is in deflationary territory.

Figure 3 plots the point estimates for $\beta_1$ as the cut-off value for $\pi$ is lowered to $-5\%$. The graphs represent the five specifications of equation (14) from table 2, where $\beta_{11}$ is from the full fixed effects model, $\beta_{12}$ the fixed country effects model, $\beta_{13}$ the country fixed year effects model, $\beta_{14}$ the model without fixed effects but with a linear time trend, and $\beta_{15}$ the OLS specification from equation (14').

![Figure 3](image)

The five point estimates illustrates the dramatic decline as a few hyper-inflation observations are removed from the sample, which is followed by gradual decline as the cut-off is lowered from 500% to some 20%. Beyond this, i.e. in the territory of moderate to low inflation as well as mild deflation, the point estimate for $\beta_1$ is getting higher, so that we end up with a U-shaped curve. Moreover, although the model specification does to some extent modify the magnitude of the point estimates, the story is basically the same. The major difference is that the models with fixed country effects produce somewhat lower point estimates in the in the 500 to 20 per cent cut-off range. Obviously, there are time-invariant country specific characteristics that to some degree predict the level of inflation, which leaves less leverage to $\beta_1$. At the margins, however, the model specification does not make a lot of difference.

Since most economies nowadays manage to keep inflation below 20%, the territory at the right margin is the most relevant. Let us hence have a closer look at this range. From table 4, we infer that the fixed effects become insignificant as the 0% inflation cut-off is reached, so that we resort to the OLS specification, which does not require as many degrees of freedom as the fixed effects models, which may matter, since the sample size quickly goes down as we approach the right margin. Figure 4 shows the point of $\beta_{15}$ from equation (14') and its 95-percent confidence interval across the 20 to $-5$ range of cut-off values. The point estimate moves a bit more erratically then for most part of figure 3, which is due to the fact that the sample size is quickly decreasing (from 1335 at $\pi = 20\%$ to 25 at $\pi = -5$), so that the removed observations are relatively gaining in importance for the regression. Nevertheless, the point estimate does not continue to decrease monotonically, and it remains significantly positive.
mate does not continue to decrease monotonically, and it remains significantly positive throughout.

Figure 4

Taken together, these analyses leave no doubt that for our sample period, i.e. the years since 1991, the classical proportionality theorem does not hold, as $\beta_1$ is far from stable. At the same time, this analysis confirms that the quantity theory continues to have explanatory power. $P = f(M)$ with $f' \geq 0$ is confirmed for all cut-off values, even for samples that comprise only deflationary observations, where one might speculate that the relationship between monetary growth and changes in the price level is governed by untypical rules.

How does this finding relate to the existing body of empirical literature? Let us first acknowledge that the number of empirical studies on quantity theoretical statements is vast, so that it would take more space than available here to give anything like a comprehensive summary. Let us hence look at a few recent studies that either find support for the quantity theory of money or tend to reject it and see how our results relate to this literature.

In what then seemed to be the final verdict on the quantity theory of money, a large number of studies found that in the US, the once stable relationship between money and nominal GDP had collapsed in the 1970s.41 Benjamin Friedman accordingly concludes:

"whatever economic conditions might have warranted reliance on monetary growth targets in the 1970s and early 1980s had long disappeared by the 1990s, so that the Federal Reserve's 'downgrading' of these targets was indeed an appropriate response to changing circumstances." (1996: 2)

Meanwhile, the sceptics would usually concede that money matters for inflation. A typical conclusion is that by Bachmeier and Swanson (2005: 583), who find that in the US from 1979 to 2003, M2 had "marginally predictive content for inflation". The concession, however, has no practical consequence, as the preferred theory for the explanation or prediction of inflation remains the Phillips curve or the output gap, i.e. concepts that directly or indirectly relate to the NAIRU. Supply side oriented economics has certainly become the dominant paradigm for the analysis of inflation.

---

41 For an authoritative summary of this view, see e.g. Blinder (1998).
On the other hand, a large number of recent studies still find that the quantity theoretical explanation of inflation is a useful analytical device. CROWDER (1998), e.g., reports that in the US during the 1980s and 1990s, the trend growth rate of the monetary base was a significant determinant of CPI inflation. AKSOY and PIKORSKI (2006) find that when the US stock of money is corrected for foreign holdings, the quantity theoretical relationship between money and nominal output has persisted beyond the early 1980s. Taking account of the fact that the US dollar was the dominant international currency when the relationship between the dollar and nominal GDP in the US apparently broke down in the 1970s hence to some degree re-establishes the quantity theory’s predictive performance, implying that the mainstream may have abandoned it prematurely. If the US are a special case, the fact that empirical research that tends to attract most attention refers to US data may have led to a wrong generalisation.42

We now turn to a selection of studies covering other economies. BALTENSPERGER ET AL. (2001) analyse data from Switzerland and show that excess growth of M3 is highly and correlated with subsequent inflation. First findings from the euro area (TRECROCI and VEGA 2000; ALTIMARI 2001) confirm that after the introduction of the common currency, broad money aggregates like M3 reflect information about future inflation in that they significantly improve money-less indicator forecasts. From the other side of the globe, similar findings are reported. Although the Reserve Bank of New Zealand was one of the first central banks that not only abandoned monetary targeting, but would not even refer to monetary aggregates for the purpose of forecasting inflation, a number of analyses (COLLINS ET AL. 1999; RAZZAK 2002) show that broad money continues to be a predictor of future inflation in forecasts that fare not worse than models that refer to the Phillips curve or the output gap and hence are mostly regarded superior to approaches that build on the old fashioned quantity theory.

Other studies refer to panel data and cover large number of countries. These analyses tend to support the quantity theoretical explanation of inflation. DWYER and HAFER (1999), e.g., find a close relationship between excess money growth and subsequent inflation corresponding to the proportionality theorem in a panel of 79 countries covering the period from 1987 to 1997, i.e. a period when the mainstream had already dismissed the quantity theory. In a recent study, DE GRAUWE and POLAN (2005) refer to a panel of 160 countries covering yearly observations from 1969 to 1999. They find that the strong link between money growth and inflation that can be identified in the complete sample is "almost wholly due to the presence of high-inflation and hyper-inflation countries in the sample" (p. 256). Referring to low inflation countries only (defined as countries that experience less than 10% inflation on average), they conclude that the relationship is "weak, if not absent" (ibid.).

DE GRAUWE and POLAN’S study is closely related to ours, both methodologically and in its perhaps most important finding: The close relationship between excess money growth and inflation that shows up in panels that include very high and hyper-inflation inflation observation is getting considerably weaker when these are removed from the sample. Yet, while they are ready to dismiss the quantity theory for low-inflation countries, our analysis confirmed that the quantity theory, though certainly not proportionality between money and prices still holds for samples that are restricted to low inflation or deflationary observations.43

Summing up this necessarily short and selective survey, it is fair to conclude that both sceptics and proponents of the quantity theoretical explanation of inflation can refer to sound evidence to support their claims. After the world moved to a paper standard in the 1970s, the relationship between money and prices became considerably looser, notably in the US, but it

42 For this argument, see also BRUGGEMAN ET AL. (2003) and CALZA and SOUZA (2003).
43 This may be due to the different panels or to the fact that DE GRAUWE and POLAN set cut-off values for average inflation in a country across the sample period, while we apply our cut-off values to observations rather than countries. We shall certainly come back to this issue in another paper.
did not break down completely. Moreover, while the theory appears to lose ground in a low inflation environment, money growth generally remains a significant predictor of subsequent inflation, although the magnitude of this effect is far below proportionality.

5 The present status of the theory

As the basis of monetarism in the 1970s, the quantity theory had significant influence on academic theory as well as on the practice of monetary policy. Since then, we have seen a continuing reduction in its relevance. In the USA money targeting was abandoned in the early 1980s, when the correlation between the excess supply of money and inflation had apparently collapsed. Ever since, this has characterised the dominant point of view, at least in the English-speaking world: monetary aggregates are inappropriate as intermediate targets.

After monetary targeting was given up by the SNB in 1999, and markedly downgraded by the ECB in 2003, the last prominent central banks finally dismissed the quantity theory as a basis to implement money policy.

In addition to this, particularly in the English-speaking world, not even monetary analysis, i.e. taking money as an indicator rather than an intermediate target, is any longer considered worthwhile as guidance for monetary policy. This practice is based on two main points that question the practical usefulness of the quantity theory:

1. Problems with measurement and definition of monetary aggregates: the appropriate monetary aggregate cannot be measured with sufficient accuracy;
2. Goodhart's law: the correlation between the stock of money and nominal output breaks down as soon monetary policy relies on it to achieve its goals.

Regarding the first point, there are indeed numerous examples that demonstrate the difficulty in defining and measuring monetary aggregates. We just mention two recent cases, the far-reaching revision of Swiss monetary aggregates by the Swiss National Bank in 1995 and the re-booking of German savings balances from M3 to M2 after the EMU came into effect. Yet, in our view, problems with definitions and measurement are in principle manageable; and the frequent revisions of monetary statistics around the world show that the problem is recognised and endeavours are being made to cope with it. From this perspective, the quantity theory will continue to produce predictions with empirical content as long as payments are customarily settled with money. Now, are we witnessing that economic transactions are decoupled from what we would in economic terms define as money, i.e. a financial stock variable that delivers convenience rather than return and is characterised by inelastic supply? We do not know what tomorrow will bring, but presently, conventional liquid assets still clearly prevail, and "electronic money" is advancing slowly (OECD 2002). Surely, as liquid assets are getting less tied to the monetary base, the quantity theory will continue to lose significance for the conduct of monetary policy. However, in as much as the stock of money can be considered to be exogenous, it would in our view clearly be premature to consign the quantity theory into the realm of historical dogma.

The second point is the in our view more substantive argument against the quantity theory: Goodhart's law states that a statistical regularity falls apart, when a policy maker tries to exploit it. In the case of monetary targeting, this could be because the financial sector and/or the public undermine effective implementation of a money supply rule through financial

44 COLLINS ET AL. (1999: 7) explicitly mention New Zealand, the US, Canada and the UK.
45 See DALZIEL (2000).
46 See GOODHART (1981: 5): "Any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes." For an in-depth discussion of Goodhart's law, see EVANS (1985).
innovation that facilitates substitution of illiquid and liquid financial assets. This is reflected by the observation that monetary targeting, when it was introduced in the 1970s, usually referred to a narrow monetary aggregate (the Bundesbank e.g. first targeted the monetary base), which was then widened, and in the end the target was usually broad money (M3). Apart from that, the public is to some degree free to adapt its demand for money to the conditions set by monetary policy, so that, in technical terms, $V^Y$ will turn into an endogenous variable, which jeopardises a policy that builds on endogeneity of $V^Y$ in equation (12). In the end, central bank money could be avoided altogether, which would drive the money multiplier towards zero and thus deprive the monetary authority of its traditional means to manage the quantity of money held by the public.

Goodhart's law makes a strong argument against application of the quantity theory with $M$ as an exogenous control parameter, in predicting that the observed relationship between $M$ and $\pi$ is weakened or collapses as soon as the policy maker refers to it. At the same time, it follows that as long as $M$ is not used as a control variable, it will continue to reflect information about $\pi$. In this sense, abandoning monetary targeting was a precondition for re-establishment of a stable relationship $P = f(M)$. Ironically, the progressive dismissal of the quantity theory since the 1980s may have contributed to the fact that it now appears alive as ever. And indeed, it seems that there is a new consensus emerging, where money is disregarded as a control variable but continues to be – or is resurrected – an a valuable indicator of inflationary pressure. Even in a central banking environment, which prides itself not to pay attention to money, it is acknowledged that neglect of monetary analysis could have gone too far.

To conclude, we hence shall not suggest an epilogue. The quantity theory may be the oldest theory that still features prominently in economics, but in the history of science, the observation "the older you are, the sooner you die" does not always hold. As we have argued, the vitality of the quantity theory lies in the fact that it has been able – explicitly or implicitly – to apply an identity, the equation of exchange, to ever changing historical circumstances and institutional change. It emerged as mediaeval rulers manipulated the coinage, and as the Black Death was followed by price increases. In early modern times it contributed to understand the price revolution. Later it was the theoretical basis for controversies about the links between bank money, paper money and bullion. When the last remainders of the gold standard vanished in the last century, it was referred to as a theory that advocated monetary targeting, but it was then quickly (in historical terms) abandoned. Nowadays, it continues mainly as a theory of inflation. It is presently challenged by the fact that financial innovation appears to

---

47 GREENSPAN (2003) summarises: "... in the past two decades, what constitutes money has been obscured by the introduction of technologies that have facilitated the proliferation of financial products and have altered the empirical relationship between economic activity and what we define as money, and in doing so has inhibited the keying of monetary policy to the control of measured money stock." See also GOMME (1998), who relates the failure of monetary targeting in Canada to this issue.

48 Notice that without demand for central bank money, not only would the foundation for monetary targeting be stripped from the central bank, but the bank rate policy would be obsolete too, so that this argument can also be brought forward against inflation targeting via the central bank rate.

49 See e.g. SVENSSON (2003).

50 Referring to New Zealand, COLLINS ET AL. (1999: 20) remark: "Given the myriad of ways that money and credit aggregates can be influenced by definitions, by financial innovation and restructuring, and so forth, can we make anything of money and credit aggregates as economic indicators? Perhaps contrary to what one might have assumed ..., one need not look too hard to find reasonable evidence that money and credit aggregates may provide useful in gauging inflation and the developments in the real sector." And in his August 2003 speech at Jackson Hole, Alan GREENSPAN, after justifying why the Fed had early abandoned monetary targeting, comes back to the quantity theory: "Nonetheless, in the tradition of Milton Friedman, it is difficult to disregard the long-run relationship between money and prices."
make it increasingly difficult to define and measure monetary aggregates that are appropriate representations for $M$ in the equation of exchange. Yet, as the history of money shows, the quantity theory has survived profound changes to the currency system, so that we are confident that it will also cope with the next one and continue to produce predictions with empirical content.

References


BEUTIN, Ludwig and KELLENBENZ, Hermann (1973), Wirtschaftsgeschichte, Köln.


BORN, Karl Erich (1972), Die Entwicklung der Banknote vom "Zettel" zum gesetzlichen Zahlungsmittel, Mainz.


BRUGGEMAN, Annick, DONATI, Paola and WARNE, Anders (2003), Is the Demand for Euro Area M3 Stable? EZB, Frankfurt am Main.


EINZIG, Paul (1949), Primitive Money in its Ethnological, Historical and Economic Aspects, London.
FISHER, Irving (1911), Recent Changes in Price Levels and Their Causes, American Economic Review, 1(2): 37–45


HUME, David (1752), Political Discourses, Edinburgh.


KREMERS, Jeroen and LANE, Timothy D. (1990), Economic and Monetary Integration and the Aggregate Demand for Money in the EMS, IMF Staff Papers, 37: 777–805.


RITTMANN, Herbert (1975), Deutsche Geldgeschichte 1484–1914, München.


TAEUBER, Walter (1933), Geld und Kredit im Mittelalter, Nachdruck 1968, Köln.


