

How uncertain are the welfare costs of inflation?

Hasan Bakhshi

Ben Martin

and

Tony Yates

hasan.bakhshi@bankofengland.co.uk;

ben.martin@bankofengland.co.uk;

tony.yates@bankofengland.co.uk

The views expressed in this paper are those of the authors alone, and not necessarily those of the Bank of England, nor of the Monetary Policy Committee. We are grateful to Peter Andrews, Spencer Dale, Andrew Hauser, Simon Hayes, Norbert Janssen, Jens Larsen, Ed Nelson, Ryland Thomas, and Alex Wolman of the Federal Reserve Bank of Richmond for comments and helpful conversations.

Copies of working papers may be obtained from Publications Group, Bank of England, Threadneedle Street, London, EC2R 8AH; telephone 020 7601 4030, fax 020 7601 3298, e-mail mapublications@bankofengland.co.uk

Working papers are also available from www.bankofengland.co.uk/wp/index.html

The Bank of England's working paper series is externally refereed.

Contents

Abstract	5
Summary	7
1 Introduction	9
2 The shopping-time model of money demand	12
3 A UK calibration of the shopping-time model	20
4 The impact of low inflation on the demand for money and welfare	27
5 Uncertainties in the estimates of the welfare costs of inflation	31
6 Conclusions	37
Appendices	38
References	42

Abstract

This paper quantifies for the United Kingdom the general equilibrium costs of individuals holding cash to economise on 'shopping time'. These are a subset of a wider range of costs caused by inflation. The paper tests whether or not money balances tend to a finite number as nominal interest rates tend to zero, and investigates how sensitive to this test are the welfare implications of rates of inflation above the Friedman rule (a zero nominal interest rate). The paper then explores how uncertainties about the shape of the money demand curve translate into uncertainties about these welfare costs of inflation. A key uncertainty is the existence of a satiation point for money balances. We show using Monte Carlo tests that without observations at nominal interest rates very close to zero, the power of satiation tests can be very low. This finding may also be important for evaluating whether/how monetary policy could stabilise the economy in the event of a shock large enough to require that nominal interest rates are driven close to zero.

Summary

Inflation makes it costly to hold non-interest bearing cash and reserves. For a long time economists have known that this cost could be eliminated if monetary policy acted to bring about a steady state of zero nominal interest rates, where there is no penalty to holding cash. (In such a regime, inflation would be equal to the negative of the real interest rate, the rate that equalises the return to holding cash and a risk-free real asset.) Since then a number of researchers have sought to quantify how much time (and utility) is thrown away as nominal interest rates rise above zero. But why all this effort, when the ‘optimal’ inflation rate has been worked out? The interest stems from the apparent consensus in modern monetary regimes that policy should aim at a positive rate of inflation. Those regimes are predicated on the notion of setting the costs of staying away from the ‘Friedman rule’ on the one hand against the costs of lowering inflation further on the other. These costs of lowering inflation are highly uncertain and difficult to model coherently, but could be important. Leaving aside measurement problems, researchers have examined whether low inflation could cause problems if nominal wages or prices are downwardly rigid. And they have also sought to quantify the costs of monetary policy becoming impotent as nominal rates hit the zero bound in regimes of very low inflation. Models tractable enough to calculate the costs of inflation are typically simplified to the point where the economic behaviours that could generate these ‘benefits’ of positive inflation are not included. So the interest in calculating the welfare cost of positive inflation is a pragmatic one. Absent an all-singing, all-dancing model that includes a zero bound and downward nominal frictions, take a monetary general equilibrium model, calculate the costs of positive inflation, and balance these in an informal way against the ‘benefits’.

This paper adds to the literature on quantifying the costs side of the inflation ‘balance sheet’. It makes two distinct contributions. First, it offers a UK calibration of some of the general equilibrium costs of inflation that complements the efforts of Bakhshi *et al* (1999) and Chadha *et al* (1998), which take partial-equilibrium approaches. The estimate is a ‘general equilibrium’ one in the sense that, following Wolman (1997), it takes as its benchmark a model of money demand due to McCallum and Goodfriend (1987), who argued that individuals hold cash in order to economise on shopping time.

This approach to calculating the welfare costs of inflation is distinct from an older literature that stretches from Bailey (1956), Sidrauski (1967), through Lucas (1994) and most recently to Chadha *et*

al (1998). The intuition behind those papers was that just as you can use the area under the demand curve for apples to calculate the consumer surplus that accrues from apple consumption when the market clears at a certain price, so you can do the same with money. If money is a consumption good then you can compare the area under the demand curve for money when nominal rates are x per cent with when they are zero, and thereby derive a measure of the benefit of reducing steady-state nominal interest rates. In the shopping-time model, money is modelled not as directly utility-providing but as enabling consumers to translate work into consumption more efficiently. So the fact that our paper offers a ‘general equilibrium’ estimate of the welfare cost of inflation comes from the chosen shopping-time model that tells us that inflation has consequences for consumption and leisure decisions, whereas the ‘money gives you utility all by itself’ approach says that it does not.

Our ‘general equilibrium welfare cost of inflation’ comes, following Wolman (1997), from applying a shopping-time money demand function to UK data and combining it with a model of consumption and leisure choice.

The second contribution of our paper is that we look at the uncertainties surrounding those welfare costs of positive inflation that we capture. One dimension of this is to look at how uncertainties surrounding our shopping-time parameter estimates translate into uncertainties in the welfare calculation. The intuition we seek to confirm is that the more ill-determined are the parameters of the estimated money demand equation, the greater will be the uncertainty about the welfare costs of inflation. Another dimension of uncertainty is whether or not real money balances tend to a finite number at zero nominal interest rates. Wolman (1997) develops a test of satiation interpretable within the shopping-time model that we implement for UK data. We also explore how powerful these tests are when there are few—if any—observations at very low nominal interest rates. Intuitively, if there are no observations near zero nominal rates, then we cannot observe whether real balances tend to a finite number or not. Simulations using our model support this intuition.

Introduction

In a recent paper, Lucas writes:

'In a monetary economy, it is in everyone's private interest to try to get someone else to hold non-interest-bearing cash and reserves. But someone has to hold it all, so all of these efforts must simply cancel out. All of us spend several hours per year in this effort, and we employ thousands of talented and highly-trained people to help us. These person-hours are simply thrown away, wasted on a task that should not have to be performed at all.' (Lucas (2000, page 247).)

For a long time, at least since Friedman (1969), economists have known that the task would not have to be performed at all if monetary policy acted to bring about a steady state of zero nominal interest rates, where there is no penalty to holding cash. (In such a regime, inflation would be equal to the negative of the real interest rate, the rate that equalises the return to holding cash and a risk-free real asset.) Since then a number of researchers have sought to quantify how much time (and utility) is 'thrown away' as nominal interest rates rise above zero.⁽¹⁾ But why all this effort, when the 'optimal' inflation rate has been worked out? The interest stems from the apparent consensus in modern monetary regimes that policy should aim at a positive rate of inflation.⁽²⁾ Those regimes are predicated on the notion of setting the costs of staying away from the 'Friedman rule' on the one hand against the costs of lowering inflation further on the other.⁽³⁾ These costs of lowering inflation are highly uncertain and are difficult to model coherently, but could be important. Leaving aside measurement problems,⁽⁴⁾ researchers have examined whether low inflation could cause problems if nominal wages or prices are downwardly rigid.⁽⁵⁾ And they have also sought to quantify the costs of monetary policy becoming impotent as nominal rates hit the zero bound in regimes of very low inflation.^(6,7) Models tractable enough to calculate the costs of inflation are typically simplified to the

(1) Examples are Bailey (1956), Friedman (1969), Lucas (1994, 2000), Wolman (1997), Chadha *et al* (1998), Feldstein (1999), Bakhshi *et al* (1999) and Dotsey and Ireland (1996), though there are many more.

(2) See, for example, Haldane (1995).

(3) Note that the 'Friedman rule' is taken to refer to the notion that the optimal rate of inflation is the negative of the real interest rate, rather than Friedman's celebrated rule for achieving a given rate of inflation by setting money growth equal to 'k per cent'.

(4) See Cunningham (1996) and Boskin (1996) for details of biases in price indices.

(5) See Akerlof *et al* (1996), Smith (2000) and Yates (1998) for examples of this literature.

(6) See King and Wolman (1996), Fuhrer and Madigan (1994), Reifschneider and Williams (1999), and King (1999) for examples.

(7) Many of the papers in Feldstein (1999), including Bakhshi *et al* (*op cit*), also allow for the temporary output loss that occurs when prices are (symmetrically) sticky, drawing on the literature on the costs of disinflation; see Ball, Mankiw and Romer (1988), for example.

point where the economic behaviours that could generate these ‘benefits’ of positive inflation are not included.⁽⁸⁾ So the interest in calculating the welfare cost of positive inflation is a pragmatic one. Absent an all-singing, all-dancing model that includes a zero bound and downward nominal frictions, take a monetary general equilibrium model, calculate the costs of positive inflation, and balance these in an informal way against the ‘benefits’.

This paper adds to the literature on quantifying the costs side of the inflation ‘balance sheet’. It makes two distinct contributions. First, it offers a UK calibration of some of the general equilibrium costs of inflation that complements the efforts of Bakhshi *et al* (1999) and Chadha *et al* (1998), which take partial-equilibrium approaches. The estimate is a ‘general equilibrium’ one in the sense that, following Wolman (1997), it takes as its benchmark a model of money demand due to McCallum and Goodfriend (1987), who argued that individuals hold cash in order to economise on shopping time. In this model a rise in nominal interest rates has two effects. First, it lowers the return to working at the margin, because individuals are unable to translate that work into the same amount of consumption: they either have to spend more time to shop for the same amount of consumer goods, or they have to incur more interest costs. Second, there is an income effect. The reduction in the real wage reduces lifetime wealth, which raises the marginal utility of wealth and hence of working, so workers will work more as a result.⁽⁹⁾ The ‘general equilibrium’ aspects of the costs of inflation come from accounting for these effects on consumption and leisure decisions by households.

This approach to calculating the welfare costs of inflation is distinct from an older literature that stretches from Bailey (1956), Sidrauski (1967), through Lucas (1994) and most recently to Chadha *et al* (1998). The intuition behind those papers was that just as you can use the area under the demand curve for apples to calculate the consumer surplus that accrues from apple consumption when the market clears at a certain price, so you can do the same with money. If the market for apples clears at a price of ten pence, then we know that the benefit from consuming the last apple is ten pence. If the demand for apples is downward-sloping, then the benefit from consuming the next to last apple is greater than ten pence. By adding up all the marginal benefits along the demand curve then

(8) An exception is King and Wolman (1996).

(9) For a given level of productivity and hence real wages, people cannot generate the same consumption-leisure level as before. Either they economise on real balances and spend more time shopping (having less time for either leisure or work), or they reduce lifetime income by suffering the increased erosion of their monetary assets.

subtracting the price we can calculate the total consumer surplus. So with money. If money is a consumption good then you can compare the area under the demand curve for money when nominal rates are x per cent with when they are zero, and thereby derive a measure of the benefit of reducing steady-state nominal interest rates. In this older literature money provides utility in and of itself.⁽¹⁰⁾ This is not to suggest that proponents of models of this class believe that individuals get pleasure from money directly, but simply that the services money yields to facilitate household activities can be modelled as if money were indeed a consumption good like any other. In the shopping-time model, money is modelled not as directly utility-providing but as enabling consumers to translate work into consumption more efficiently. So the fact that our paper offers a ‘general equilibrium’ estimate of the welfare cost of inflation comes from the chosen shopping-time model that tells us that inflation has consequences for consumption and leisure decisions, whereas the ‘money gives you utility all by itself’ approach says that it does not.

Our ‘general equilibrium welfare cost of inflation’ comes, following Wolman (1997), from applying a shopping-time money demand function to UK data and combining it with a model of consumption and leisure choice.

The second contribution of our paper is that we look at the uncertainties surrounding those welfare costs of positive inflation that we capture. One dimension of this is to look at how uncertainties surrounding our shopping-time parameter estimates (embodied in our parameter standard errors) translate into uncertainties in the welfare calculation. The intuition we seek to confirm is that the more ill-determined are the parameters of the estimated money demand equation, the greater, other uncertainties aside, will be the uncertainty about the welfare costs of inflation. Another dimension of uncertainty is whether or not real money balances tend to a finite number at zero nominal interest rates. Friedman (1969) argued, for example, that at some point, the costs of hiring security guards to look after hoards of cash would mean that reductions in interest rates would not increase real balances. If, on the other hand, real balances tend to an infinite quantity then a larger portion of the costs of positive inflation accrue in the first few percentage points away from the Friedman rule than if they do not. The argument has been at the centre of the ‘money as a consumption good’ approach to calculating the welfare costs. For example, Lucas (2000) and Chadha *et al* (*op cit*) both argue that

(10) This literature typically assumes that agents’ money-holding decisions are additively separable to agents’ leisure and consumption decisions.

lines that tend to infinity at zero nominal rates ('log-log' curves) fit real money data better than those that do not ('semi-log' curves). Wolman (1997) develops a test of satiation interpretable within the shopping-time model that we implement for UK data. We also explore how powerful these tests are when there are few—if any—observations at very low nominal interest rates. Intuitively, if there are no observations near zero nominal rates, then we cannot observe whether the real balances line tends to a finite number or not. We generate data under models of satiation and non-satiation, perform econometric tests of satiation in these artificial worlds, and look to see what happens to our inference as we strip away observations close to zero nominal rates.

The rest of the paper proceeds as follows. We begin in Section 2 by setting out the shopping-time model, though do so briefly as readers can turn to McCallum and Goodfriend (1987) for a more detailed consideration. We then set out in Section 3 how we calibrate the model for UK data: beginning first with the estimation of the money demand equation, and then turning to the calibration of other key parameters in the 'real' part of the model. In Section 4 we undertake our benchmark welfare analysis. Section 5 discusses some uncertainties in our welfare analysis. Section 6 concludes.

2 The shopping-time model of money demand

A key question is why we have chosen the shopping-time model of money demand amongst other specifications. We are forced at the outset to rely on those theoretical approximations to microfounded money demand functions that have been successfully integrated into general equilibrium models of the macro-economy. The lessons learned from the literature begun by Kiyotaki and Wright (1989) and the problems that have defined its progress since then are ignored by this paper.

The shopping-time model of money demand assumes that time and real money holdings are (imperfect) substitutes for undertaking transactions that enable consumption. The model should not literally be taken to imply that 'shopping' is more efficient with money, but is a statement about money providing services to the holder which make all activities involved in undertaking consumption—of which shopping is one—easier (or less costly). We might, for example, equivalently label the model 'the liquidity management model of money demand', arguing that the

model is a device for approximating the fact that real money holdings can help to cut down the time required to manage transfers between real and cash assets needed to finance consumption.

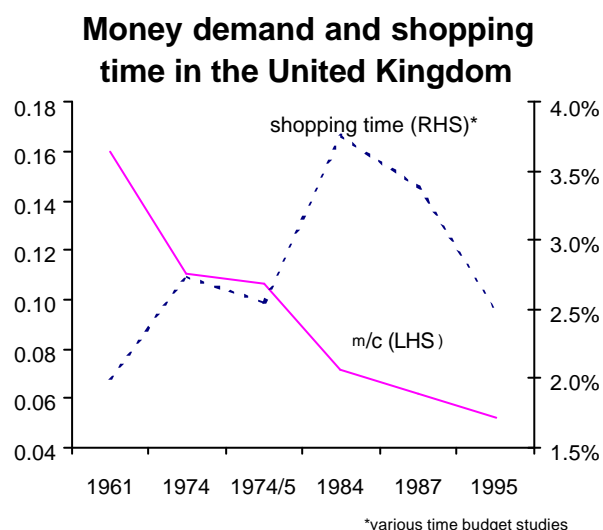
The model was formalised by McCallum and Goodfriend (1987) who employed a specification suggested first by Saving (1971). It is distinct from, but related to, models where households require ‘cash in advance’ to finance consumption (for example, Lucas and Stokey (1987)). In the shopping-time model, agents can substitute (real) cash for time. In cash-in-advance models they can’t. Cooley and Hansen (1989,1991) show how in a cash-in-advance model agents respond to higher steady-state inflation (and therefore nominal interest rates) by substituting consumption for leisure. In the shopping-time model if inflation rises, agents substitute away from holding cash towards devoting time to ‘shopping’. Karni (1973) made just this point. More recently, Aiyagari *et al* (1998) offer models with a similar flavour, which have consumers purchasing ‘credit services’ that help consumers substitute from goods that require ‘cash in advance’ towards goods that can be purchased with credit. The shopping-time model is also related to the Sidrauski (1967) or ‘money in the utility function’ model. In this model, as we outlined in the introduction, the inflation effect on money balances affects welfare, but not consumption. Formally, the term in money in the utility function is separable from those in consumption and leisure. With some appropriate restrictions, the shopping-time specification is equivalent to a ‘money in the utility function’ model (see, for example, Croushore (1993)), but in general it implies a ‘money non-separably (and indirectly) in the utility function’ functional form. This gives rise to the more complex general equilibrium effects of inflation (on consumption) that mean, as we have discussed, that estimates of the area under empirical money demand curves are insufficient to capture the full effects of inflation, a point made by many, including, for example, Dotsey and Ireland (1996).

Our choice to adopt the shopping-time specification is a pragmatic one. Although it does not approach the rigour of more properly microfounded approaches (like Kiyotaki and Wright (1989)), it allows us to derive an empirically testable equation. Moreover, it does so in a model that allows for a more subtle interpretation of the value of money than cash-in-advance models, and one that allows us to explore general equilibrium effects of inflation, which ‘money (separably) in the utility function’ models do not. A final advantage is one drawn out by Wolman (1997), who showed how the shopping-time model can be used to derive a theoretically interpretable empirical test for ‘satiation’ in real balances, which parallels the ‘log-log or semi-log’ debate in the papers by Lucas (1994) and Chadha *et al* (1998) (amongst others). One objective we have is to measure the

importance of the existence of a finite satiation point and to gauge the importance of being unsure about it. Wolman’s interpretation of shopping time allows us to do just that.

Some support for a shopping-time specification can be found by looking at a plot of inverse velocity and survey measures of shopping time (see Chart 1). This does not depict the structural relationship between money demand and shopping time in the shopping-time technology (see equation (4) on page 16), nor does it account for the many factors that have exogenously changed shopping time. We also acknowledge that the time measured in the surveys may constitute ‘leisure’. But the (weak) negative correlation between inverse velocity of notes and coin and measures of shopping time derived from time budget studies is at least suggestive that the assumed functional form for the shopping-time technology is not way off track.

Chart 1



The message to draw from this chart is that, though we have cautioned against employing a too narrow or too literal interpretation of the shopping-time model, the data suggest that our interpretation can at least *include* a literal one!

2.1 The shopping-time technology and money demand

The shopping-time technology, or the function that tells us how much shopping time is saved per unit of real balances, is a crucial part of our paper and therefore needs some discussion upfront. Since the remainder of the model is standard, we focus on drawing out the intuition of the

shopping-time model and its importance for the shape of the money demand relation. This section follows King and Wolman (1996) and Wolman (1997).

The shopping-time technology says that for a given level of consumption, shopping time can be reduced by holdings of real balances. We further impose homogeneity so that shopping time is a function of the ratio of money balances to consumption:

$$(h_t) = h \left(\frac{m_t}{c_t} \right) \text{ for all } t \quad (1)$$

Time spent shopping (or in leisure) is at the expense of time spent working, and therefore the opportunity cost of a unit of shopping time is the real wage w_t . Agents maximise a utility function which has terms in consumption and leisure time (specified below). To reduce the cost imposed by time spent shopping, agents can increase their money balances but this will be offset to a certain extent by the opportunity cost of holding (non-interest-bearing) money. This is the value of interest forgone R_t , discounted from when that interest would be received (ie next period), at rate

$$b = (1 + R_t)^{-1}.$$

Consumers choose shopping time and money balances so that the marginal cost of holding a unit of money balances is balanced by the marginal benefit of reduced shopping time (both measured in monetary units). Formally:

$$\frac{R_t}{1 + R_t} = -\frac{w_t}{c_t} h' \left(\frac{m_t}{c_t} \right) \quad (2)$$

This expression implicitly defines the demand for money. King and Wolman's choice of function form for $h_t'()$ nests different explicit forms for the money demand function. They choose $h_t'()$ to be of the form:

$$h' \left(\frac{m_t}{c_t} \right) = f - \left(\frac{m_t}{Ac_t} \right)^{\frac{1}{n}} \quad (3)$$

where, following Wolman's notation, the parameter n is negative. The nature of the money demand function at very low interest rates therefore depends on whether $f=0$ or whether it is strictly positive. At a zero nominal interest rate, money holdings are chosen so that the marginal benefit (in reduced shopping time) is driven to zero. If $f>0$ then there is a finite level of money balances consistent with the efficiency condition at zero nominal interest rates. But if $f=0$ the level of money balances consistent with the efficiency condition will be infinite at zero interest rates.

With this explicit functional form for $h'()$, we can derive an explicit money demand curve from the efficiency condition above:

$$\frac{m_t}{c_t} = A \left(\frac{c_t}{w_t} \frac{R_t}{1+R_t} + f \right)^n \quad (4)$$

We will fit this equation to UK data in Section 3.

We shall also make use of the fact that the satiation level of money balances (as a share of consumption) at zero nominal interest rates is Af^n , which is infinite if f is zero since n is assumed negative.

Once again, if $f>0$ then there exists some positive level of real balances at which the marginal benefit to consumers from holding money, accounting for the marginal reduction in shopping time, is zero. But although a satiation level of real balances is definable in the context of a shopping-time model, this does not establish the economic plausibility of satiation, not least since the shopping-time model has to be thought of as an approximation of the microfoundations for money demand. To illuminate the debate Wolman (1997) offers two competing quotes from Lucas and Friedman. Thus Lucas (1994, page 16) comments: 'Managing an inventory always requires *some* time, and a larger average stock must always reduce this time requirement, no matter how small it is.' On the other hand, Friedman (1969, page 17) puts it: '[As nominal interest rates fall] the retailer dispenses with an errand boy to economise on cash balances, which is a gain, but, at some point, he must hire guards to protect his cash hoard.'

We do not form a prior about the plausibility of satiation, but follow Wolman in allowing it to be decided by the data.

It is useful to note that this functional form for money demand can nest the log-log and semi-log forms of the money demand equation familiar from other empirical studies. Consider the case when the nominal interest rate is low enough for the approximation $R/(1+R)=R$ to be reasonable. Taking logs of the money demand equation when $f > 0$ we get:

$$\ln m_t = \ln A + \ln c_t + \nu \ln f \left(1 + \frac{nc_t}{fw_t} R_t \right) \quad (5)$$

which for small R_t we can use the approximation $\ln(1+x)=x$ to get:

$$\ln m_t = \ln A + \ln c_t + \nu \ln f + \frac{nc_t}{fw_t} R_t \quad (6)$$

(Recall the parameter n is assumed to be negative).

If the constant $f=0$ this approximation is no longer valid, but taking logs of (4) and rearranging gives:

$$\ln m_t = \ln A + (1+n) \ln c_t - n \ln w_t + n \ln R_t \quad (7)$$

Given the functional form for $h_r(\cdot)$ the derivative of the shopping-time function, it can be integrated with respect to m/c to obtain the expression for shopping time h :

$$h \left(\frac{m_t}{c_t} \right) = f \left(\frac{m_t}{c_t} \right) - \frac{u}{1+u} A^{-\frac{1}{u}} \left(\frac{m_t}{c_t} \right)^{\frac{1+u}{u}} + \Omega, \text{ for } \frac{m_t}{c_t} < A f^u, \quad (8)$$

$$h\left(\frac{m_t}{c_t}\right) = \frac{A}{1+u} f^{1+u} + \Omega = \bar{h}, \text{ for } \frac{m_t}{c_t} = A f^u$$

where \bar{h} represents the minimum possible shopping time.

2.2 The standard problems for consumers and firms, Wolman (1997)

The representative agent maximises utility subject to a budget constraint and a time constraint, which we can represent formally as follows:

$$\text{Max}_{c_{t+j}, l_{t+j}} E_t \sum_{j=0}^{\infty} \mathbf{b}^j u(c_{t+j}, l_{t+j}) \text{ where } u(c_{t+j}, l_{t+j}) = \ln c_{t+j} + \gamma \ln l_{t+j} \quad (9)$$

subject to:

$$P_{t+j} c_{t+j} + M_{t+j} + \frac{B_{t+j}}{1+R_{t+j}} = M_{t+j-1} + B_{t+j-1} + P_{t+j} w_{t+j} n_{t+j} + D_{t+j}$$

$$n_{t+j} + l_{t+j} + h_{t+j} = 1$$

c_{t+j} and l_{t+j} are consumption and leisure respectively in period $t+j$. \mathbf{b} is the agent's time discount factor. P_{t+j} is the price level, M_{t+j} denotes the agent's nominal money holdings, B_{t+j} is the agent's holdings of one-period zero-coupon nominal bonds, R_{t+j} is the nominal interest rate on these bonds, w_{t+j} is the real wage, n_{t+j} is the fraction of time the agent spends working, D_{t+j} denotes dividend payments from firms, l_{t+j} is the fraction of time spent as leisure, and h_{t+j} is the fraction of time spent shopping.

Given the shopping-time specification, the first-order conditions for the consumer's optimisation problem are:

$$c_{t+j} : u_c(c_{t+j}, l_{t+j}) = \mathbf{I}_{t+j} P_{t+j} \left(1 - w_{t+j} h_c \left(\frac{m_{t+j}}{c_{t+j}} \right) \frac{m_{t+j}}{c_{t+j}^2} \right) \quad (10)$$

$$l_{t+j} : u_l(c_{t+j}, l_{t+j}) = \mathbf{I}_{t+j} P_{t+j} w_{t+j} \quad (11)$$

$$B_{t+j} : 1 + R_{t+j} = \frac{I_{t+j}}{bE_{t+j}I_{t+j+1}} \quad (12)$$

$$M_{t+j} : I_{t+j} - bE_{t+j}I_{t+j+1} + I_{t+j}P_{t+j}w_{t+j}h\left(\frac{m_{t+j}}{c_{t+j}}\right) \frac{1}{P_{t+j}c_{t+j}} = 0, \quad (13)$$

$$\text{where } h\left(\frac{m_{t+j}}{c_{t+j}}\right) = f - A^{-\frac{1}{\alpha}} \left(\frac{m_{t+j}}{c_{t+j}}\right)^{\frac{1}{\alpha}}$$

The money demand equation (4) comes from manipulating the FOCs for money and bond holdings ((13) and (12) respectively).

The representative firm maximises discounted expected future profits, where the discount rate for period $t+j$ is the consumer's marginal rate of substitution between a unit of wealth in periods t and $t+j$ (see equation (12)). Formally, the firm's problem is:

$$\text{Max}_{k_{t+j+1}, n_{t+j}} E_t \sum_{j=0}^{\infty} b^j \frac{I_{t+j}}{I_t} \left(P_{t+j} a_{t+j} f(k_{t+j}, g^{t+j} n_{t+j}) - w_{t+j} P_{t+j} n_{t+j} - P_{t+j} i_{t+j} \right) \quad (14)$$

subject to

$$y_{t+j} = a_{t+j} f(k_{t+j}, g^{t+j} n_{t+j})$$

$$k_{t+j+1} = k_{t+j} (1 - d) + i_{t+j},$$

where a_{t+j} is the firm's (multiplicative) productivity parameter at time $t+j$, $f(.,.)$ is the firm's production function (assumed to exhibit constant returns to scale), k_{t+j} and n_{t+j} are the firm's capital stock and employment levels, g denotes labour-augmenting technical progress (assumed exogenous), w_{t+j} is the real wage that firms pay, i_t is investment and d is the depreciation rate of the firm's capital stock. Note that the model we are adopting omits the possibility that money may help firms as well as consumers economise on 'shopping time'. We do so partly for simplicity and partly as a practical matter, since only about 10% of cash in the United Kingdom is held by firms.

Other authors have modelled firms' money demand separately (see, for example, Fischer (1974)), though not, so far as we are aware, in the context of assessing the welfare costs of inflation⁽¹¹⁾

The first-order conditions for the firm's optimisation problem are therefore simply:

$$n_{t+j} : w_{t+j} = a_{t+j} g^{t+j} f_n(k_{t+j}, g^{t+j} n_{t+j}) \quad (15)$$

$$k_{t+j+1} : P_{t+j} = \mathbf{b} E_{t+j} \frac{I_{t+j+1}}{I_{t+j}} (P_{t+j+1} a_{t+j+1} f_k(k_{t+j+1}, g^{t+j+1} n_{t+j+1}) + P_{t+j+1} (1-d)) \quad (16)$$

Equation (15) is the standard condition equating the real wage and the marginal product of labour. Equation (16) is the firm's first-order condition with respect to next period's capital stock. It states that the increase in future profits associated with a higher capital stock in the next period just offsets the decrease in current profits resulting from an increase in investment.

3 A UK calibration of the shopping-time model

There are two important elements to our account of the UK calibration. The first is to describe the empirical work that gives rise to our calibration of the parameters in the 'money demand' equation, or the empirical counterpart of equation (4). The second is to describe our calibration of the 'real' part of the economy: the parameters in the other first-order conditions.

3.1 Parameters in the money demand function

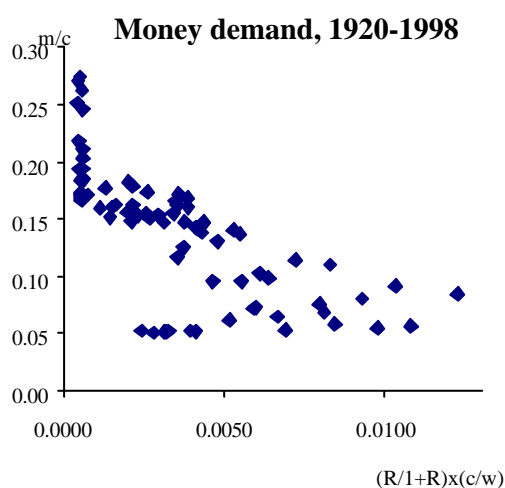
Our approach here follows Wolman (1997) in deriving estimates of the parameters A , \mathbf{f} and \mathbf{u} in the money demand function. In doing so there are conflicting demands on our methodology. The first demand is that we need an equation that is interpretable within our theoretical model. Without this we cannot carry out our general equilibrium welfare analysis. The second demand is that we would like an equation that fits the data and specifically an equation that performs in a way that allows us to

⁽¹¹⁾ Our observation that only 10% of cash in the United Kingdom is held by firms is supportive of, but not proof of, our implied assertion that modelling firms is not important for the welfare costs of inflation. This low proportion will be an equilibrium response to the current inflation rate and could be one that implies a high interest elasticity of demand (and a high welfare cost of positive inflation) for firms.

conduct inference. Recall that one of our goals is to test whether or not there is a satiation point, which amounts to testing for a zero value of f . If we were not interested in general equilibrium analysis, we could depart from the theoretical specification and fit the data with a vector autoregression. If we were not interested in conducting inference on our parameters then we could make do with imposing our theoretical equation on the data and proceeding. Readers should interpret what we do here as an attempt to reconcile these conflicting demands, rather than as a comment on the now enormous literature on empirical money demand functions (see the survey by Goldfeld and Sichel (1990) and Chadha *et al* (1998) for examples).

Chart 2 plots the two composite variables entering the money demand function in equation (4) using historical data for the monetary base, consumption, earnings and interest rates on seven-day deposit accounts in the United Kingdom from 1920 to 1998 (see Appendix 1 for details). The empirical problems that are likely to confront us in fitting a functional form to these data should be apparent from the chart.

Chart 2



We follow Wolman (1997) in deriving estimates of A , f and u using non-linear least squares (NLLS) and in using a bootstrapping procedure to generate confidence intervals around the central estimates which we can use to make inferences about satiation and to construct estimates of the uncertainty surrounding the welfare costs of inflation. The bootstrapping procedure has the advantage that we do not have to make distributional assumptions about the error term when conducting inference. (Note that the NLLS estimates are consistent and asymptotically normal,

though not efficient unless the error term is normally distributed (see Amemiya (1974)).⁽¹²⁾ So asymptotic results could be used to justify using standard inference techniques to test the null of non-satiation.) The procedure is described in Appendix 2.

Writing equation (4) in terms of the composite variables $y_t = \begin{pmatrix} m_t \\ c_t \end{pmatrix}$ and $x_t = \begin{pmatrix} R_t & c_t \\ i + R_t & wt \end{pmatrix}$, the equation we estimate is:

$$y_t = A(x_t + f)^n + e_t \quad (17)$$

The table below reports the central NLLS estimates.⁽¹³⁾ It also reports the two-sided 95% confidence bands around these central estimates, as generated by the bootstrapping procedure:

Table 1

<i>Coefficient:</i>	<i>A</i>	<i>f</i>	<i>n</i>
Central estimate	0.005282	0.001655	-0.5980
Upper conf bound	0.01385	0.004375	-0.39473
Lower conf bound	0.0008533	0.0005094	-1.02321

The estimate of f appears statistically significantly different from zero, echoing Wolman's result for the United States. This implies that real balances tend to a positive number at zero nominal interest rates.⁽¹⁴⁾ For comparison we estimated equation (17) imposing non-satiation ($f = 0$), giving the following estimates:

(12) Of course this assumes that both the dependent and independent variables are stationary. There is a strong prior that both $\frac{m_t}{c_t}$ and $\frac{c_t R_t}{w_t(1+R_t)}$ are stationary: all sensible theories of consumer behaviour predict that money demand and consumption cointegrate, on the one hand, and that consumption and wage income cointegrate on the other. In fact, using standard ADF tests it is not possible to reject the null of non-stationary at conventional significance levels. But these tests have notoriously low power. That is the case with our dataset too: using Kwiatkowski, Phillips, Schmidt and Shin (1992) tests, which test the null of stationarity, we cannot reject stationarity either (results available on request). We proceed with the assumption that both variables are indeed stationary.

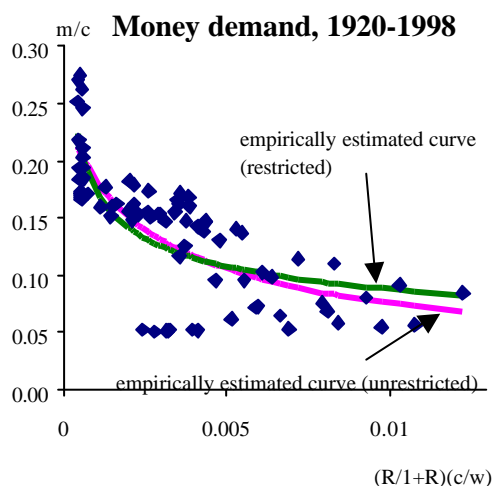
(13) We experimented with different starting values for the parameters in our NLLS routine and found that the estimates were robust.

(14) Chadha, Haldane and Janssen (1998) do not reject the log-log specification for money demand using a similar dataset to that used in this paper. But they do express discomfort at the implication that this specification implies non-satiation, or infinite money/GDP ratios at zero nominal interest rates.

$$y_t = 0.02243x_t^{-0.2961} \quad (18)$$

These two fitted curves are plotted in Chart 3 below. Note that on inspection the curves are extremely close to each other yet the economic implications of the two, in particular the welfare costs of inflation, are qualitatively different. Wolman (1997), in contrast, finds that for the United States the unrestricted and restricted curves are markedly different. Indeed, later in the paper we will ask whether it makes much difference to our calculation of the welfare cost of inflation whether or not real balances display satiation; we are going to conclude, in short, that in the United States it does, but in the United Kingdom it does not.

Chart 3



Our hypothesis test rejects the null of non-satiation in the UK demand for narrow money. Our parameter estimates imply that at zero nominal interest rates, the ratio of money to consumption equals 0.24. This in itself is insufficient to tell us what the satiation level is in current pounds per head. To do that we need to calculate consumption at zero nominal interest rates, and to do that we need to calibrate the remainder of our general equilibrium model.

3.2 Robustness of the empirical strategy

At this point we can usefully discuss two features of the empirical strategy. First, our response to the serial correlation in the errors. Second, our use of single-equation NLLS.

It is a well-known property in the empirical money demand literature that equation errors in a regression of the kind we estimated are significantly autocorrelated. (Once again, see Goldfeld and Sichel (1990), Chadha *et al* or Hoffman and Rasche (1996).) The serial correlation in the residual is arguably a sign of model mis-specification. Traditionally, researchers have responded by including a lagged dependent variable (or many) in the money demand function. Such lags could be motivated by the argument that consumers face costs in adjusting real balances, and allow us to preserve the usefulness of the equation for our welfare analysis (without such a motivation extra variables in the money demand equation would leave us unable to use it for the welfare calculation). But including a lagged dependent variable often fails to mop up the serial correlation completely and, more fundamentally, it is difficult to think of adjustment costs that are large enough to explain the high degree of serial correlation that is observed (see Goodfriend (1985) for an exposition of these points). Chadha, Haldane and Janssen's (1998) solution is to estimate an error correction model for money demand in the United Kingdom. Their approach stresses a distinction between long-run and short-run money demand where portfolio adjustment costs are used to justify the short-run (data-driven) dynamics. This distinction may turn out to be theoretically valid. For the moment, we have to rely on an equation that coheres with the shopping-time model and on the notion that that model can capture the fundamental welfare implications of different steady states, notwithstanding its inability, as currently specified, to capture all the dynamics. An alternative line of defence is offered by Goodfriend (1985). He argues that stochastic measurement error in the opportunity cost of holding money and in the scaling variable (consumption in the shopping-time model) can explain the observed serial correlation in the money demand function. So there is no need to rely on (implausibly high) portfolio adjustment costs. This argument would support using a theoretical model of money demand with no adjustment costs, such as the shopping-time model used in this analysis.⁽¹⁵⁾

Note too that we used single-equation NLLS. Although it is clear that our model would suggest that the regressors in the money demand equation (interest rates, consumption and earnings) are themselves endogenous, implying that unless we instrument in some way a single-equation method is inappropriate. We arrived at the single-equation NLLS strategy (like Wolman (1997)) after evaluating two other possibilities. One was to use GMM. We experimented with a GMM procedure using a variety of instruments (lagged values of the regressors) and, like Wolman (1997), found that

(15) A reply by Laidler (1985) in the same conference is sceptical of this line of reasoning.

the estimated coefficients were highly sensitive to which instruments were used. Without a theory to help guide our choice of instruments GMM is therefore problematic.⁽¹⁶⁾

Another alternative to our NLLS routine would have been to use a maximum likelihood procedure. Fuhrer, Moore and Schuh (1995) argue that ML estimates are generally unbiased, dynamically stable and economically plausible compared with GMM estimates. The problem in our case is that the residual from the money demand equation is far from normally distributed, so any of the standard distributional assumptions about the error are likely to be violated.

Given the uncertainties surrounding the empirical calibration of the money demand function, we checked for robustness in two ways: by changing the sample period, and by allowing for a possible time trend in the shopping-time technology.

The observations nearest the origin (with low values of inverse velocity and low interest rates) clearly worsen the fit of the empirically-estimated money demand curve. These observations all occurred in the 1990s when narrow money demand was historically low, given the prevailing low levels of nominal interest rates. One possibility is that there is a change in some fundamental parameter that means the money demand equation experienced a ‘structural break’ in the 1990s. We investigated the possibility of structural breaks by re-estimating the satiation level implied by equation (4) over the sample period excluding the 1990s. The estimated ratio of money balances to consumption, m/c , at zero nominal interest rates is 0.22 (compared with 0.24 for the full sample).

The closest antecedent for our paper, Wolman (1997), studies the United States but only covers the period up to 1992. Evidence in Ball (1998) suggests that unusually high narrow money velocity in the 1990s has been a phenomenon in the United States too. All this might suggest that there has been a structural break in narrow money demand in the 1990s. But Ball (1998) argues against this view

(16) The result that the estimates from the GMM procedure are sensitive to choice of instruments is not new. Mao (1990) and Nelson and Startz (1990) find that the performance of the GMM estimator and the J-test of over-identifying restrictions is sensitive to the choice of the number of lags used in the instrument set. Fuhrer, Moore and Schuh (1995) report Monte Carlo evidence that GMM estimates are often biased, dynamically unstable and economically implausible (wrongly signed). They also argue that the main problem lies in the sensitivity of the results to the choice of instrument. Holman (1998) discusses the pros and cons of GMM estimators for estimating a model of money demand closely related to the shopping-time model used in this paper. She also reports that the parameter estimates are highly sensitive to which instruments are used.

for the United States for M1. He argues that econometric studies of money demand in the United States that exclude the 1990s give biased estimates of parameters in the money demand function.

Our second robustness check was to look for evidence of a deterministic trend. A trend can be motivated by thinking of technical progress in the process by which money enables consumers to economise on real balances. Or, put another way, there could be a trend improvement in the amount of shopping time that a given level of real balances implies. This captures the idea that shopping has become easier for consumers, with developments such as the growth of superstores, containing a wider range of goods, and home shopping. We included a deterministic time trend in equation (4), but found it to be insignificant and it produced economically implausible parameter estimates.⁽¹⁷⁾

We are left to continue, albeit tentatively, beginning with the assumption that there appears to be evidence of a satiation point in real balances, just as Wolman (1997) found for the United States, but in contrast to the argument that Lucas (1994) made for a ‘log-log’ specification—which implies non-satiation. Recall that Chadha *et al* (*op cit*) found that what they call ‘long-run money demand’ appears to be ‘log-log’, implying non-satiation for the United Kingdom but that ‘short-run’ money demand appears to be ‘semi-log’. To reiterate our earlier discussion, our results are not readily comparable with those. Ours come from applying an equation from a shopping-time model to data: the Chadha *et al* results are explicitly fitted to the data, with the benefits and costs (already set out in Section 2) that adopting that strategy implies.

3.3 Parameters of the benchmark economy

The welfare analysis will quantify the benefits of moving from some benchmark inflation rate to the Friedman rule. So the second stage of the model calibration is to characterise the benchmark economy.

The inflation rate in the benchmark economy is taken as the average retail price index (RPI) inflation rate over the 1920-98 sample period. This comes out as 4.1%. The production function is

(17)With exogenous shopping-time technical progress, \mathbf{a} , equation (4) would become

$$\left(\frac{m_t}{c_t}\right) = A e^{a t} \left(\frac{R_t}{i + R_t} \frac{c_t}{w_t} + \mathbf{f}\right)^u.$$

Cobb-Douglas. We take the 0.44 value for the capital share from Holland and Scott (1998). The depreciation rate is assumed to be 9% per annum.

The preference parameters \mathbf{y} and \mathbf{b} are set so that the annual real interest rate is 4.5% and the steady-state hours worked is equal to 26.3% of the time endowment (calibrated using average hours worked data for manual workers over our sample period—comparable data for non-manual workers are unavailable for the United Kingdom (sources are discussed in Appendix 1). The 4.5% real interest rate is the average post-tax real rate of return on equities from the Credit Suisse First Boston Equity-Gilt Study (1999) over our sample period. The riskless real interest rate relevant for calculating the real opportunity cost of holding money is calibrated at 2.2%. This is the sample average nominal interest rate (3.8%) minus the sample average of an *ex ante* inflation expectation based on a simple ARMA model (1.6%)—compared with Wolman’s estimate of 1%.⁽¹⁸⁾ There is no endogenously arising equity risk premium in the model: there are two distinct real rates calibrated to mimic one. Our methodology here parallels Wolman’s. Note that the assumed riskless real rate of return is crucial for determining the welfare costs of any given rate of inflation, since it determines the Friedman rule (which is the negative of this riskless real rate). But the elasticity of welfare to changes in steady-state inflation and in particular the behaviour of real balances at zero nominal rates will be independent of this assumption.

Given the estimated parameters of the shopping-time technology, the constant of integration, Ω , is chosen so that the time spent shopping in steady state is equal to 2.8%. This is the average taken from several time budget studies in the United Kingdom (Gershuny and Robinson (1988), ONS Omnibus Survey (1995) and Jenkins and O’Leary (1997)).

4 The impact of low inflation on the demand for money and welfare

Our econometric evidence rejected the null hypothesis of non-satiation, and provided us with point estimates of the coefficients of the money demand function. These estimates give an estimate of the ratio of nominal money balances to nominal consumption at zero nominal interest rates, and our

(18) The ARMA model for RPI inflation is estimated up to 1920, and the expectations beyond that constructed as out-of-sample forecasts. That the average inflation expectation is some way below the sample average inflation rate is due to the experience of the 1970s, when the model gives us low inflation and the data give us very high inflation. We take the stance, therefore, that the 1970s was a ‘surprise’.

estimate of this ratio is 0.24.⁽¹⁹⁾ This in itself does not provide us with an estimate of the satiation level of nominal balances in current prices. But using our calibrated model, we can work out how real consumption varies from the long-run average as inflation is reduced to the Friedman rule. This in turn can be expressed in current prices, allowing us to compute the current-prices per capita satiation level of narrow money. We estimate this to be around £2,000 per person on average throughout the year, in 1998 prices. This compares with current per capita holdings of narrow money of around £450.⁽²⁰⁾

As well as providing an estimate of the nominal per capita satiation level, the general equilibrium model calibrated to UK data can provide estimates of the welfare costs of different inflation rates as they are implied by the shopping-time model. Recall that we are doing so without a model that captures behaviours that could lead to ‘benefits’ of inflation that could be set against these costs. These estimates will serve as a benchmark: we will then go on to explore uncertainties in the welfare calculation emanating from the empirical calibration of the money demand function.

The welfare analysis involves using the utility function in the consumer’s problem to evaluate the returns to different steady-state inflation rates. These returns are expressed as an answer to the following question: ‘what proportion of annual total income would an individual forego/accept as compensation in order to move to a lower/higher rate of inflation?’ What we are calling ‘total income’ is the sum of consumption and leisure, where leisure units are converted into ‘income’ by valuing them at the calibrated effective real wage under the initial (in our case sample average) inflation rate. This effective real wage is the marginal product of labour adjusted for the wedge that the sample inflation rate drives between the returns to work from the point of view of the firm and the returns once the shopping-time/real-balance costs of consumption are taken into account.⁽²¹⁾

(19) This estimate comes from evaluating the expression $m/c = A \left(\frac{c}{w} \frac{R}{1+R} + f \right)^n$ at $R=0$ and at our central coefficient estimates presented in Table 1.

(20) One potential criticism of our estimate of the nominal value of the per capita satiation level might be that, since most people hold on average less than £400, the estimate of the satiation level should be scaled down appropriately. But note that we do not define the nominal per capita satiation level relative to the average per capita holdings of money balances, it is defined relative to per capita nominal consumption at the Friedman rule. But to the extent that our econometric estimates are based on a level of per capita money balances in excess of those held by many people, our coefficient estimates may be distorted and hence the per capita satiation level estimate could be indirectly affected.

(21) Intuition: when inflation is above the Friedman rule, the worker cannot transform an hour of work into the same amount of consumption: either she holds less real balances, and has to spend more time shopping, and take less leisure, or she holds the same real balances and loses out because of the erosion in their value.

Chart 4 below uses the sample average inflation rate as the base for comparison. It shows that agents would need slightly less than a 0.1% increase in ‘total income’ to leave their welfare unchanged were inflation to fall to the Friedman rule from the sample average (implying that nominal interest rates fall from their sample average to zero). Put another way, welfare would be just under 0.1% higher per period had the history of the steady-state inflation rate matched the Friedman rule. Wolman’s figure is higher at about 0.6%.⁽²²⁾

Chart 4

Welfare compared to baseline of 4.1 per cent inflation

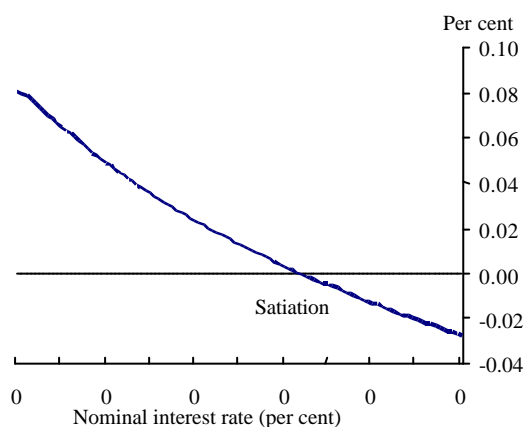
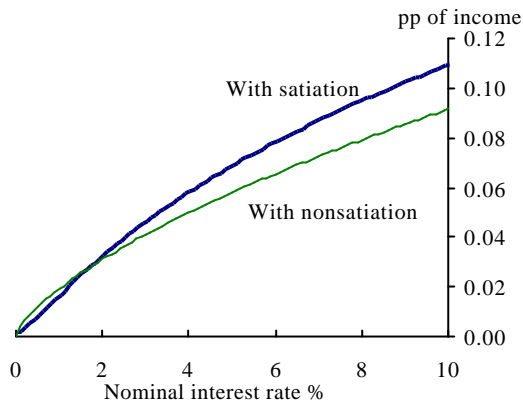


Chart 5 displays the same information, but records the gain from moving from a nominal interest rate of $x\%$ to the Friedman rule (nominal rates at zero). It also shows these welfare gains using the estimates of the money demand curve where we impose non-satiation.

(22) Wolman’s benchmark economy has a slightly different nominal interest rate, though this is unlikely to account for the difference.

Chart 5

Estimated welfare gain from moving from $x\%$ nominal interest rate to 0%



Whether we impose satiation or not, the estimated gains from reducing anticipated inflation are small: smaller than the costs that Chadha *et al* (1998), Wolman (1997) and Lucas (2000) found.

One of Wolman's key results is that the existence of satiation in the money demand curve determines where the bulk of the results of reducing inflation occur: he finds that under satiation the bulk of the gains come in reducing inflation to zero, whereas under non-satiation the bulk of the gains come in reducing inflation all the way to the Friedman rule. We examined the gains from reducing UK average inflation from 2.5% to 0% and to the Friedman rule under satiation and non-satiation. We found that under satiation 57% of the gains of moving to the Friedman rule were realised by moving to zero inflation. Under non-satiation, a little under 52% of the gain was realised at zero inflation. Therefore our results differ from Wolman's in two respects. Not only are our costs of a given inflation rate measurably smaller, but we find that the existence of a satiation point does not have as material an effect on the proportion of welfare gains that accrue by eliminating the last x percentage points of inflation away from the Friedman rule. Given a visual inspection of our Chart 5 and the equivalent in Wolman (1997), this is not surprising.

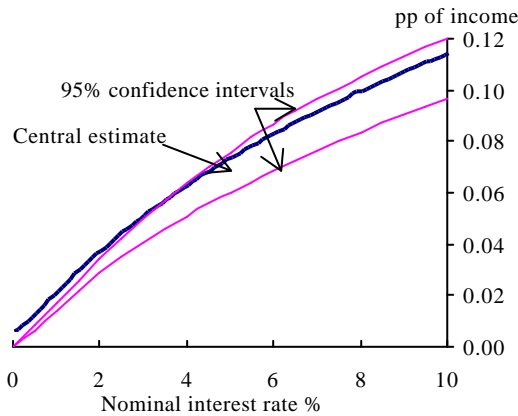
5 Uncertainties in the estimates of the welfare costs of inflation

5.1 How does uncertainty about parameter estimates in the money demand function translate into uncertainty about welfare gains?

In this part of the paper, we turn to look at uncertainties in the welfare costs of inflation that come from the empirical calibration of the money demand function. We leave to one side the uncertainties that would come from errors in the other parts of the model or in the calibration of those parts of the model, though those errors may be important. We also leave to one side the ‘fundamental’ uncertainties about the validity of the shopping-time specification. Instead we focus on two issues. First, supposing the shopping-time model to be qualitatively correct, how do the bootstrapped distributions for the estimated parameters in the money demand function translate into distributions for the welfare costs of inflation? Second, can we measure the degree to which our inference about satiation in money balances is impaired by having few observations at or near zero nominal interest rates?

Chart 6

Estimated welfare gain from moving from x% nominal interest rate to 0%



The first step was to compute 95% confidence bands around our ‘welfare costs of inflation curve’, based on the uncertainty underlying the estimates of the three parameters in the money demand function. The method used to construct the confidence intervals is described in more detail in Appendix 3. Briefly, we used the sample distributions of the money demand parameters generated in the bootstrapping procedure used to conduct inference in Section 2 (and see Appendix 1) as

approximations to the true joint probability density function of the parameters. The rest of the model was then re-calibrated for each of a 1,000 draws of the parameter triplets, and then the welfare costs of inflation were computed as before for each of these re-solved models.

The confidence intervals (around the welfare costs coming from our base case of a satiated demand curve) are shown in Chart 6 above. Note that the sampled uncertainty about money demand parameters does not translate into large uncertainties about the welfare gains of moving from a given inflation rate to the Friedman rule. This may seem surprising, but recall that there are other sources of uncertainty in this calculation that we are ignoring but which could make these bands much larger. Indeed we might be more disturbed if the welfare gains were much more sensitive to the shopping-time function parameters than to the inflation rate. The error bands are probably small enough to enable us to infer, given all the other maintained assumptions in the model, that the profile of welfare costs accruing away from the Friedman rule are significantly different from the United States. But this is an important ‘given’: it is likely that the other sources of uncertainty we have put to one side make it hard to draw this conclusion from our study of money demand related uncertainties alone.

5.2 The impact of the lack of observations at low nominal interest rates on the hypothesis test for satiation

Another dimension of uncertainty is whether researchers can make accurate inference about satiation in real balances. Economies have spent little or no time, Japan excepted, at zero nominal interest rates. Intuitively, if we are trying to detect whether real balances tend to a finite number or not at zero balances, one might imagine that our ability to do this would be greater, the closer to and more frequent were our nominal interest rate data to zero. In the extreme, an economy that was pushed to a steady state at zero nominal interest rates would either manifest infinite real balances or it wouldn't.

To explore how our inference might be impaired by having few observations near zero nominal rates, we set up the following kind of ‘Monte Carlo’ experiment: generate some artificial data on real balances and nominal interest rates based on a model; add to it some random noise; run a regression like (4) on that model data, and on many other datasets generated in the same way; conduct hypothesis tests on \mathbf{f} ; see how often you get the right answer; strip out observations with nominal

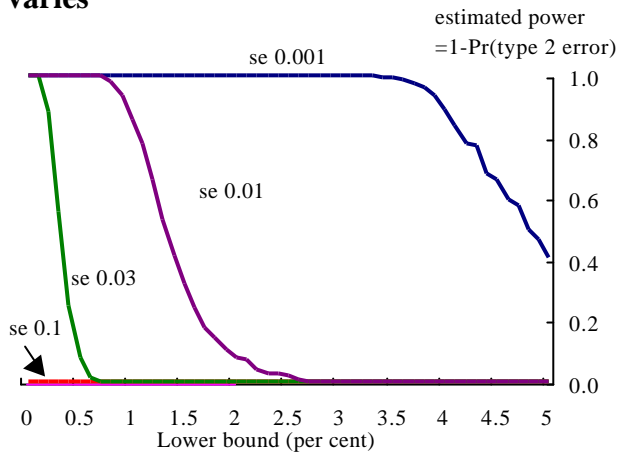
interest rates below a threshold value and see how often the regression gives you the right answer; and so on, increasing the threshold for nominal interest rates, recording how the frequency of correct inference changes as the sample shrinks. In each case we began with data generated for a uniform distribution of nominal interest rates between 0 and 14% (our sample maximum for the United Kingdom) in steps of one basis point.

Chart 7 contains our results for the case when the true model has satiation. To explain, take the line labelled 'se=0.03'. This line sets the variance of the noise added to the generated money demand observations to 0.03, approximately the variance of the residuals in our fitted money demand curve for the United Kingdom described in Section 3. The y-axis records the power of our satiation test; the proportion of the 1,000 regressions where we make the correct inference about the satiation point from the t-ratio on $f^{(23)}$: specifically, the amount of times we reject $H_0: f=0$ at the 5% significance level, or equivalently 1 minus the probability of a type 2 error. The x-axis indicates the lower threshold for nominal interest rates in the sample. For a threshold value of 1%, we have stripped out all observations with interest rates below this before running the regression and the hypothesis test. So the 'se=0.03' line shows how the power of the test falls from one initially, to around 0.3 at a threshold value of 0.5%, which is the lowest observation for nominal interest rates in our UK sample. This shows very clearly how econometricians will find it harder to conclude that there is satiation, even if the true model exhibits it, the further is the lowest nominal interest rate observation from zero. This result makes our finding (and Wolman's) all the more striking. Note too that the greater the noise in the data, the more we are dependent on low interest rate observations to make good inference (or, in terms of the chart, the earlier is the point along the x-axis that the power falls off).

(23) Note that our artificial datasets have 1,000 observations so we can rely on the asymptotic properties of NLLS estimators and use the t-ratios for inference, rather than using bootstrapped errors as we did for the actual UK historical data.

Chart 7

estimated power as sample lower bound varies



We can explore why the power falls off as we strip the sample of low interest rates, and why this fall-off is related to the variance of the error term. Re-write the regression (17):

$$y_t = A(x_t + \mathbf{f})^v + \mathbf{e}_t \tag{17}$$

as

$$y_t = Ax_t^n \left(1 + \frac{\mathbf{f}}{x_t} \right)^v + \mathbf{e}_t \tag{19}$$

Compare this with the specification when non-satiation is imposed:

$$y_t = Ax_t^n + \mathbf{e}_t \tag{20}$$

There are two situations in which testing specification (19) will lead us to accept the restricted model (20). One is when our sample estimate of \mathbf{f} is close enough to zero to be thought of as statistically insignificantly different from zero—as would be expected. The other is when the ratio \mathbf{f}/x_t is very close to zero. But as the nominal interest rate increases x_t also increases, and so the bracketed term in specification (19) will tend to 1, making this model look very like the specification under satiation, especially when an additive disturbance confuses the picture.

Our estimate of f was 0.0016, and the sample average deposit rate was about 3.75%. Our (satiated) model implies that at this level of deposit rates, the average x_t is an order of magnitude larger than f ie around 0.01, and we may reasonably expand the bracketed expression using the binomial theorem:

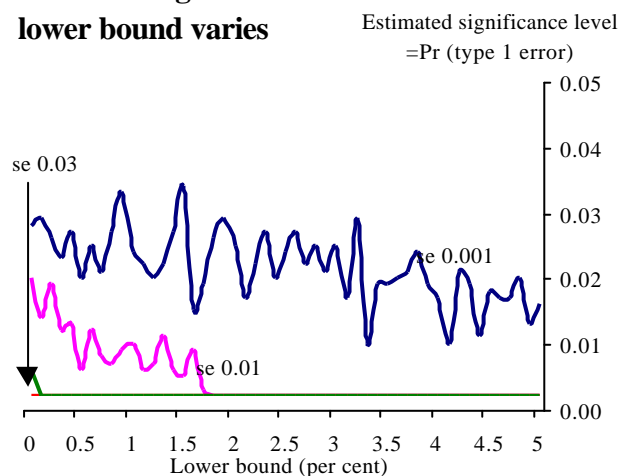
$$y_t \approx Ax_t^n \left(1 + \frac{nf}{x_t} \right) + e_t = Ax_t^n + \frac{Anf}{x_t} + e_t \quad (21)$$

For nominal interest rates near the sample average, and based on our parameter estimates, the second term would be a factor of 1,000 smaller than the first term due to the term fn . If the error term has a standard error sufficiently large, the term involving f will be swamped at moderate interest rates.

Turn now to our tests when the true model is non-satiated ($f=0$). In the same way as before, we explore our inference on the null hypothesis that there is non-satiation as we increase the lower bound on nominal rates. The results are shown in Chart 8.

Chart 8

Estimated significance level as lower bound varies



The chart shows that the significance level (the chance of rejecting $H_0: f=0$ when H_0 is true) falls as we increase the lower bound. Moreover, it falls off more quickly the greater is the variance in the

error term⁽²⁴⁾ So, just as with our first experiment, the higher is the floor to interest rates, the more our tests are biased towards finding non-satiation. In this example, we chose to test a hypothesis that was actually true, so here the sample stripping acts to help us reach the right answer. But note that this is fortuitous. The chance of correctly rejecting a range of false hypotheses about positive values of f (the power, in other words) is falling as the floor to nominal interest rates rises, just as in the first experiment in Chart 7.

To summarise, Chart 7 shows that when satiation is true, if you do not have low nominal interest rate observations, then there is a low probability of rejecting non-satiation. In contrast, Chart 8 shows that when non-satiation is true, even if you do not have low nominal interest rate observations, then there is still a low probability of rejecting non-satiation. So if there is non-satiation you should find it even with a small number of low interest rate observations. But if there is satiation it is difficult to find this in the data if there are a small number of low interest rate observations. That asymmetry in our Monte Carlo simulations should make us (and Wolman) even more confident in our results finding satiation.

If we were prepared to assume that the basics of the general equilibrium model we use are correct, and that the shopping-time model in particular is correct, we could draw some comfort from these experimental tests: namely, by noting that both ourselves and Wolman (1997) managed to find satiation in money balances even though, as we have shown, the tests for satiation may be biased towards finding the opposite when the sample has few near-zero interest rate data.

Regardless, these results show that our inference about satiation may be severely hampered by the fact that it is based on datasets with few observations close to zero nominal rates. On the face of it, that it may be difficult (or impossible) to decide between the two models would seem to matter more for the United States, for which Wolman found satiation changed the costs of staying $x\%$ away from the Friedman rule quite a lot, than for the United Kingdom. This is true with one important qualification. Whether real balances tend to infinity at zero nominal interest rates may be important for policy-makers engaged in monetary policy in a dynamic setting, the costs of steady-state inflation aside. Were a shock to arrive that was large enough to require a stabilising monetary impulse which

(24) The significance level generated when the standard error equals 0.1 is drawn on the chart, but is approximately zero whatever the lower bound.

was in turn large enough to drive nominal interest rates to the region of zero, whether/how monetary policy could work would depend in part on whether there is a satiation point in real balances. Recent contributions by Goodfriend (1999) and McCallum (1999) meditate on the legitimacy of, and the policy options in the face of, a zero nominal bound to interest rates in more detail.

6 Conclusions

The literature on calculating the welfare costs of inflation is now a large one. Our increment to it is as follows. First, we have presented a ‘general equilibrium’ estimate of a subset of the costs of inflation for the United Kingdom, and contrasted this with a similarly derived estimate by Wolman (1997) for the United States. (Our estimated costs were smaller). They were ‘general equilibrium’ estimates in the sense that, as suggested by the chosen (shopping-time) model, they accounted for the distortion to consumption, work and leisure decisions and constructed using an explicit utility function for the representative consumer: in this sense they departed from the tradition established by Bailey (1956) through Lucas (1994) and recently Chadha *et al* (1998) of treating money as a consumption good and measuring the welfare costs of inflation by integrating under a money demand curve (of varying coherence with theory). This general equilibrium estimate is interesting insofar as it can be used to set against the features of an economy (not modelled here) which could generate an optimal inflation rate away from the Friedman rule.

Our second increment to the literature was to show how the empirical money demand uncertainties translated into uncertainties about those general equilibrium welfare costs. We showed that the variances around the central estimates of the parameters in our money demand function do not translate into large uncertainties about the welfare cost of inflation. We showed that in economies with few or no observations of nominal interest rates close to zero it may be very difficult to make good inference about whether real money holdings tend to some finite number or not as nominal rates go to zero. In the United States this matters quite a lot for figuring out the welfare costs of staying at an inflation rate $x\%$ away from the Friedman rule. In the United Kingdom it does not. But this uncertainty about the existence of a satiation point in money demand is important for the task of assessing the (gross) benefits of higher inflation.

Appendices

Appendix 1: Data and sources

Consumption

Source: ONS, Sefton and Weale (1995). Consumption data are in current prices and seasonally adjusted. After 1955, the series is AIIIX from the ONS National Accounts. Prior to this, the series is from Sefton and Weale (1995), Table A2, consumers expenditure 1920-1990.

Nominal average earnings

Source: ONS, Liesner (1985), British Labour Statistics (1971). Between 1963 and 1998 the source is ONS. From 1990 onwards, the series is LNMQ, whole economy average earnings, seasonally adjusted. Between 1963 and 1990 the average earnings series is DNHS. Between 1940 and 1963, the series is from Liesner (1985), 'Average weekly earnings of manual workers'. Between 1920 and 1968 the series is from British Labour Statistics (1971), 'All manual workers, all industries and services'.

Deposit rate

Source: Bank of England, Capie and Webber (1985). From 1983 to 1998 this series is AJNR from Financial Statistics. Prior to this it is from the *Bank of England Statistical Abstract* seven-day deposit rate, and before this from Capie and Webber (1985), Interest on deposit accounts.

Monetary base

Source: Capie and Webber (1985), Bank of England. Prior to 1982 this the monetary base series from Capie and Webber. Post-1982 the growth rate of M0 is used to project this series forward.

Population

Source: ONS, Mitchell (1988).

Post versus pre-tax data?

Note that we use pre-tax data. A strict interpretation of the shopping-time model would require using post-tax wages as this is the appropriate measure of the opportunity cost of leisure and shopping. (An even stricter application would be to model the government tax and expenditure

choices, and their impact on consumers' marginal decisions explicitly.) But, in practice, we do not have average marginal income tax rates ('effective' tax rates) for labour income going back over our sample period. One option is to use the basic rate of income tax (that we do have for the whole sample period) as a proxy for the 'effective' rates. We choose not to do that here: the basic rate of income tax does not even cointegrate with the effective tax rates data we have for different household types for the 1948-82 sub-sample from Dilnot *et al* (1984). So using the basic rate would be a poor proxy for the effective rates in our steady-state analysis. In the absence of a suitable effective tax rate series we use the pre-tax wage. Wolman (1997) finds that using post-tax wages instead of pre-tax wages makes little difference to his results for the United States.

Appendix 2: Bootstrapping procedure to generate sampling distributions for parameters in the money demand function, taken from Wolman (1997)

The bootstrapping procedure is done in three stages. First, equation (4) is estimated using NLLS. The residuals exhibit significant serial correlation. This is a general property of empirical money demand functions—even when a lagged dependent variable is included—see Goodfriend (1985) and Laidler (1985). The second stage is to fit an AR model to the residuals, producing a new set of whitened residuals.⁽²⁵⁾ The final stage is to draw randomly (with replacement, that is, leaving the chance of drawing a particular residual the same, regardless of whether that value has already been drawn) from the population of whitened residuals. With each draw from the whitened residuals, we put back in the serial correlation from the AR model to give a new residual with the same serial correlation properties as the residual in stage 1. Combined with the parameter estimates from stage 1 and the original x_t data (the composite consumption, wage, nominal interest rate series) we can then use the new residuals to construct a series for \tilde{y}_t (real balances). For each newly constructed value of \tilde{y}_t we re-run NLLS to generate a different estimate for the coefficients. These estimates are collected to form the bootstrapped distributions.⁽²⁶⁾

(25) In fact an AR (1) model fitted the residuals reasonably well.

(26) In fact, the bootstrapping procedure generates a *joint* sampling distribution of A , f and u .

Appendix 3: Computing the 95% confidence bands

Recall the description of the bootstrapping procedure in Appendix 2. We artificially generated a set of 1,000 triplets of the parameters A , \mathbf{f} and \mathbf{n} . This provided a sample distribution showing the relative frequencies of \mathbf{f} , which we took to be an approximation to the true joint probability density function. Similarly we have the sampling distributions of A and \mathbf{n} . So for any given inflation rate we can draw a triple $A, \mathbf{f}, \mathbf{n}$ from these sampling distributions to parameterise a money demand function. We would need to calibrate \mathbf{y} and \mathbf{W} (defined in Section 3.3) specifically for this parameterisation of the money demand function (and the other, unchanged benchmark parameters) and then compute the welfare gain in moving from the inflation rate x to the Friedman rule. We repeated this for 1,000 parameterisations of the money demand function. This provides 1,000 estimates of the welfare gain associated with that inflation rate. Taking these observations to be the sampling distribution, the 2.5% and 97.5% critical values can be obtained, representing a confidence interval around the central estimate. Repeating this for the range of inflation rates generates a confidence band around the central estimate.

References

- Aiyagari, R, Braun, A and Eckstein, Z (1998)**, 'Transaction services, inflation, and welfare', *Journal of Political Economy*, Vol. 106, No. 6.
- Akerlof, G A, Dickens, W T and Perry, G L (1996)**, 'The macroeconomics of low inflation', *Brookings Papers on Economic Activity*, No. 1, pages 1-59.
- Amemiya, T (1974)**, 'The non-linear two stage least squares estimator', *Journal of Econometrics*, Vol. 2, pages 105-10.
- Bailey, M (1956)**, 'The welfare cost of inflationary finance', *Journal of Political Economy*, Vol. 64, pages 93-110.
- Bakhshi, H, Haldane, A and Hatch, N (1999)**, 'Some costs and benefits of price stability in the United Kingdom', in Feldstein, M (ed), *NBER*.
- Ball, L (1998)**, 'Another look at long-run money demand', *John Hopkins University working paper*.
- Ball, L, Mankiw, N G and Romer, D (1988)**, 'New Keynesian economics and the output-inflation trade-off', *Brookings Papers on Economic Activity*, No. 1, pages 1-65.
- Boskin, M (1996)**, 'To order more accurate measures of the cost of living', Final Report to the Senate Finance Committee.
- British Labour Statistics (1971)**, *Historical Abstract 1886-1968*, Department of Employment and Productivity, HMSO.
- Capie, F and Webber, A (1985)**, *A monetary history of the United Kingdom 1870-1982*, Routledge.
- Chadha, J, Haldane, A and Janssen, N (1998)**, 'Shoe-leather: the costs reconsidered', *Economic Journal*.
- Cooley, T and Hansen, G (1989)**, 'The inflation tax in a real business cycle model', *American Economic Review*, Vol. 79, pages 733-48.
- Cooley, T and Hansen, G (1991)**, 'The welfare costs of moderate inflations', *Journal of Money, Credit and Banking*, Vol. 23, pages 483-503.
- Credit Suisse First Boston Equity-Gilt Study (1999)**.

- Croushore, D (1993)**, 'Money in the utility function: a functional equivalence to a shopping-time model', *Journal of Macroeconomics*, Vol. 15(1), Winter, pages 175-82.
- Cunningham, A W F (1996)**, 'Measurement bias in price indices: an application to the UK's RPI', *Bank of England Working Paper no. 47*.
- Dilnot, A, Kay, J and Morris, N (1984)**, 'The UK tax system, structure and progressivity', *Scandinavian Journal of Economics*, Vol. 86(2), pages 150-65.
- Dotsey, M and Ireland, P (1996)**, 'The welfare cost of inflation in general equilibrium', *Journal of Monetary Economics*, Vol. 37, pages 29-47.
- Feldstein, M (ed) (1999)**, 'The costs and benefits of price stability', *University of Chicago Press, for NBER*.
- Fischer, S (1974)**, 'Money in the production function', *Economic Inquiry*, Vol. 12(4), December, pages 517-33.
- Friedman, M (1969)**, 'The optimal quantity of money and other essays', Chicago, Aldine.
- Fuhrer, J and Madigan, B (1994)**, 'Monetary policy when interest rates are bounded at zero', *Federal Reserve Bank of Boston, Working paper No. 94/1*.
- Fuhrer, J, Moore, G and Schuh, S (1995)**, 'Estimating the linear/quadratic inventory model maximum likelihood versus generalised method of moments', *Journal of Monetary Economics*, Vol. 35, pages 115-57.
- Gershuny, J and Robinson, J (1988)**, 'Historical changes in the household division of labour', *Demography*, Vol. 25, No. 4, November.
- Goldfeld, S and Sichel, D (1990)**, 'The demand for money', in Friedman, B and Hahn, F (eds), *Handbook of Monetary Economics*, Vol. 1, Elsevier Science Publishers.
- Goodfriend, M (1985)**, 'Reinterpreting money demand regressions', *Carnegie-Rochester Conference Series on Public Policy*, Vol. 22, Spring, pages 207-42.
- Goodfriend, M (1999)**, 'Overcoming the zero bound on interest rate policy', paper for the Federal Reserve Bank of Boston conference on 'Monetary policy in a low inflation environment', 18-20 October.
- Haldane, A G (ed) (1995)**, 'Inflation targeting', *Bank of England*.

Hoffman, D and Rasche, R (1996), 'Aggregate money demand functions', *Kluwer Academic Publishers*, Boston.

Holland, A and Scott, A (1998), 'The determinants of UK business cycles', *Economic Journal*, Vol. 108, July, pages 1,067-92.

Holman, J (1998), 'GMM estimation of a money-in-the-utility function model: the implications of functional forms', *Journal of Money, Credit and Banking*, Vol. 30, pages 679-98.

Jenkins, S and O'Leary, N (1997), 'Gender differentials in domestic work, market work, and total work time: UK time budgets survey evidence for 1974/5 and 1987', *Scottish Journal of Political Economy*, Vol. 44, No. 2, May.

Karni, E (1973), 'The transaction demand for cash: incorporation of the value of time into the inventory approach', *Journal of Political Economy*, Vol. 81, pages 1,216-25.

King, M A (1999), 'Challenges for monetary policy: new and old', speech prepared for the symposium on 'New challenges for monetary policy', sponsored by the Federal Reserve Bank of Kansas City at Jackson Hole, Wyoming, on 27 August 1999, *Bank of England Quarterly Bulletin*, Vol. 39, No. 4.

King, R and Wolman, A (1996), 'Inflation targeting in a St Louis model of the 21st century', *Federal Reserve Bank of St Louis Review*, Vol. 78, May/June, pages 83-107.

Kiyotaki, N and Wright, R (1989), 'Money as a medium of exchange', *Journal of Political Economy*, Vol. 97, pages 927-54.

Kwiatkowski, D, Phillips, P, Schmidt, P and Shin, Y (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?', *Journal of Econometrics*, Vol. 54, pages 159-78.

Laidler, D (1985), 'Expectations and adjustment in the monetary sector revisited: a comment', *Carnegie-Rochester Conference Series on Public Policy*, Vol. 22, pages 243-54.

Liesner, T (1985), 'Economic statistics 1900-1983', *The Economist*.

Lucas, R E (1994), 'On the welfare cost of inflation', *Working paper No. 394, Stanford University, Centre for Economic Policy Research*.

Lucas, R E (2000), 'Inflation and welfare', *Econometrica*, Vol. 68, No. 2, pages 247-74.

Lucas, R E and Stokey, N L (1987), 'Money and interest in a cash-in-advance economy', *Econometrica*, Vol. 55, pages 491-513.

Mao, C (1990), 'Hypothesis testing and finite sample properties of generalised method of moment estimators: a Monte Carlo study', manuscript, *Federal Reserve Bank of Richmond*.

McCallum, B (1999), 'Theoretical analysis regarding a zero lower bound on nominal interest rates', paper for the Federal Reserve Bank of Boston conference on 'Monetary policy in a low-inflation environment', 18-20 October.

McCallum, B and Goodfriend, M (eds) (1987), 'Demand for money: theoretical studies', in Eatwell, J, Milgate, M and Newman, P (eds), *The New Palgrave: the Dictionary of Economics*, London: Macmillan; New York: Stockton Press, pages 775-81.

Mitchell, B R (1988), *British Historical Statistics*, Cambridge University Press.

Nelson, C and Startz, R (1990), 'Some further results on the exact small-sample properties of the instrumental variable estimator', *Econometrica*, Vol. 58, pages 967-76.

ONS Omnibus Survey (1995), Office for National Statistics, unpublished.

Reifschneider, D and Williams, J (1999), 'Three lessons for monetary policy in a low-inflation era', Board of Governors of the Federal Reserve System, *mimeo*.

Saving, T R (1971), 'Transactions costs and the demand for money', *American Economic Review*, Vol. 61(3), June, pages 407-20.

Sefton, J and Weale, M (1995), 'Reconciliation of national income and expenditure, balanced estimates of national income for the United Kingdom 1920-1990', Cambridge.

Sidrauski, M (1967), 'Rational choice and patterns of growth in a monetary economy', *American Economic Association Papers and Proceedings*, Vol. 57, May, pages 534-44.

Smith, J (2000), 'Nominal wage rigidity in the United Kingdom', *Economic Journal*, Vol. 110, pages 179-95.

Wolman, A (1997), 'Zero inflation and the Friedman rule: a welfare comparison', *Federal Reserve Bank of Richmond Economic Quarterly*, Vol. 83/44.

Yates, A (1998), 'Downward nominal rigidity and monetary policy', *Bank of England Working Paper no. 82*.