Macroeconomic Dynamics, **24**, 2020, 1547–1573. Printed in the United States of America.
doi:10.1017/S1365100518000871

USER COSTS, THE FINANCIAL FIRM, AND MONETARY AND REGULATORY POLICY

Maksim Isakin
Cleveland State University

Apostolos Serletis
University of Calgary

We investigate how key monetary policy instruments and financial regulation affect the banking firm. We take the user-cost approach to the construction of prices for financial services and use quarterly data on the U.S. commercial banking sector, over the period from 1992 to 2016, obtained from the Federal Deposit Insurance Corporation. We use the symmetric generalized Barnett variable profit function to derive demands for and supplies of monetary and nonmonetary goods and provide evidence consistent with neoclassical microeconomic theory. We find that the compensated price elasticities of banking technology are small in magnitude. Yet a hypothetical policy experiment shows that even small changes in the holding costs of financial goods can result in significant changes in user costs and the quantities demanded and supplied.

Keywords: Commercial Banks, Generalized Barnett Variable Profit Function, Flexible Functional Forms.

1. INTRODUCTION

The current mainstream approach to monetary policy and business cycle analysis is based on the new Keynesian model and is expressed in terms of the interest rate on overnight loans between banks, such as the federal funds rate in the United States. This approach ignores the financial intermediary sector. As Adrian and Shin (2011, p. 602) put it, "in conventional models of monetary economics commonly used in central banks, the banking sector has not played a prominent role. The primary friction in such models is the price stickiness of goods and services. Financial intermediaries do not play a role, except as a passive player that the central bank uses as a channel to implement monetary policy." However, banks and other financial intermediaries have been at the center of the global financial crisis, and there is almost universal agreement that the crisis originated in the financial intermediary sector.

Address correspondence to: Apostolos Serletis, Department of Economics, University of Calgary, Calgary, Alberta, T2N 1N4, Canada. e-mail: Serletis@ucalgary.ca. Phone: (403) 220-4092. Fax: (403) 282-5262.

 1469-8056/19

It is not uniform view in the literature as to what financial intermediaries do. As Diewert et al. (2016) put it, "one of the most controversial areas in the field of economic measurement is the measurement of the real and nominal output of the banking sector. There is little consensus on all aspects of this topic: even the measurement of banking sector nominal outputs and inputs is controversial and there is little agreement on how to measure the corresponding real outputs and inputs." There is also a broader aspect to what banks do than just being financial intermediaries. Banks are allowed to create money and thus play an important role in the monetary policy transmission process. In this regard, the current approach to monetary policy also ignores the role of money, as the short-term nominal interest rate is the sole monetary variable and there is no reference to any monetary aggregates.

In this paper, we model the monetary transmission mechanism provided by the financial intermediaries and estimate the monetary-production technology. Our objective is to identify the channels through which financial intermediaries affect the real economy, and investigate the implications for regulatory and monetary policies. We follow Hancock (1985) and Barnett and Hahm (1994) and build a microeconomic model of the financial firm which produces intermediation services between lenders and borrowers. We assume that the firm maximizes a multiproduct variable profit function subject to a feasibility of financial transformation. We take the user-cost approach to the construction of prices for monetary goods. The user costs determine intertemporal revenues and costs associated with holding different assets and liabilities. This approach permits monetary goods (such as cash and deposits of various types), other financial goods (such as loans), and physical goods (such as labor and materials) to be classified as inputs or outputs. Those items with a positive user cost are classified as inputs and those with a negative user cost are classified as outputs.

The user-cost approach explicitly takes into account each bank's cost of funds. A heterogeneity in banks' cost of funds makes the user costs of the monetary goods bank- and time-specific. Since user costs vary across banks, the same investment opportunity could be profitable for a bank with a low cost of funds and unprofitable for a bank with a high one. Thus, the optimal demands for and supplies of monetary goods vary across institutions. For a similar reason, the profitability of monetary goods changes over time causing the optimal demand and supply to change. These aspects of monetary technology are often ignored in the empirical literature. However, the recognition of the heterogeneity in banks' cost of funds is important for unbiased estimation.

We approximate the variable profit function with the symmetric generalized Barnett flexible functional form introduced in Barnett and Hahm (1994). This unit profit function satisfies the theoretical regularity conditions of a profit function globally. In particular, it is linearly homogeneous in prices and maintains convexity over the positive orthant. This is an important quantity as in practice most flexible functional forms fail to meet theoretical regularity over a whole sample—see, for example, Serletis and Shahmoradi (2005, 2007). In addition, Barnett and Hahm (1994) show that the generalized Barnett form retains its flexibility under a weak separability constraint. Diewert (2015) notes that the form is quasiflexible, that is, it has certain restrictions on the second-order approximation, and not parsimonious. Yet, we show that the latter has limited consequences with current computer power. The estimation of a system with six nonlinear equations is computationally feasible on a personal computer even with a large sample size. In this paper, we use a panel data on U.S. commercial banks from 1992 to 2016 with a total of 780,825 observations.

We find that the monetary technology is relatively inelastic: the supplies of outputs and demands for inputs of financial goods are mostly inelastic and the degree of substitutability among financial goods is low. These results are robust across three different bank samples. Our estimates confirm findings in Hancock (1985) and Barnett and Hahm (1994). However, we also find that small changes in holding costs may have significant effects on user costs and, therefore, on demands for and supplies of financial goods, even when the user cost elasticities are small in magnitude. These effects are especially strong when the holding costs are close to the discount rate and the user costs become close to zero. Our findings have important implications for monetary policy, as they suggest that the central bank can implement policy changes via changes in interest-rate policy instruments even when the economy operates close to the zero lower bound.

A substantial body of literature that studies the behavior of the financial firm builds on the neoclassical theory of the firm. The user-cost approach proceeds from the assumption that the financial firm maximizes its capitalized profit choosing the quantities of monetary and nonmonetary goods subject to technological constraints. There are various empirical studies that employ this framework and estimate production, cost, and profit functions of the financial firm. Hancock (1985) investigates the flexibility of banking technology using the translog variable profit function and a longitudinal sample of 18 New York and New Jersey banks. Barnett and Hahm (1994) tests the hypothesis of weak separability and the existence of consistent monetary aggregates in monetary-production technology using longitudinal observations on 41 Chicago banks. A large part of the literature estimates banks' cost, profit or output distance functions to investigate productivity and economies of scale in banking. For example, Hughes and Mester (1998) estimate banks' cost function and examine scale economies and banks' risk aversion. Berger and Mester (2003) analyze cost and profit productivity over the 1984–1997 period. Wheelock and Wilson (2018) present nonparametric estimates of banks' cost, revenue and profit functions and analyze the evolution of scale economies before and after the financial crisis of 2007–2008.

Sealey and Lindley (1977) develop an alternative theoretical framework of the banking firm. They posit that banks are financial liaisons between liability holders and those who receive bank funds. This intermediation approach implies that all bank assets including loans and leases, investments in securities, and reserves are outputs for the banking technology. At the same time, all bank liabilities such as deposits, other debt, and equity capital are inputs. Sealey and Lindley

(1977, p. 1253) claim that deposits, the most controversial part of liabilities, are an economic input since “these services require the financial firm to incur positive costs without yielding any direct revenue.” Classifying deposits as inputs has exposed the intermediation approach to criticism for neglecting substantial services that banks provide to their depositors. Over the years, the literature has attempted to disentangle the intermediation and deposit services of banks; see, for example, Berger and Humphrey (1992) for a discussion.

The user-cost approach developed by Barnett (1978) renounces the ex-ante classification of assets and liabilities into inputs and outputs and instead derives it from the contribution of financial goods to bank profit. In particular, an asset is considered to be an output if the return on investment into this asset exceeds the opportunity cost of funds. Similarly, a liability is classified as an output if the financial cost of this liability is less than the opportunity cost of funds. According to this classification scheme, deposits are likely to be labeled as outputs, especially, if only interest expenses are taken into consideration. Interestingly, because of differences in the cost of funds, the same asset or liability can be classified as input in one bank and as output in another. For this same reason, the nature of a financial good can change over time.

This paper is organized as follows. In Section 2, we build the model of a banking firm. In Section 3, we discuss the variable profit function approach. In Section 4, we discuss the symmetric generalized Barnett variable profit function as well as the procedure for imposing convexity in order to achieve theoretical regularity. In Section 5, we deal with data issues. Our primary focus is on empirical application, specifically to quarterly data on the U.S. commercial banking sector, over the period from 1992 to 2016, obtained from the quarterly Uniform Bank Performance Reports provided by the Federal Deposit Insurance Corporation. Section 6 discusses related econometric issues. In Section 7, we estimate the model and present the empirical results. In Section 8, we conduct a monetary and regulatory policy analysis. The final section concludes the paper.

2. THE USER-COST APPROACH

A formal representation of the user-cost approach can be found in Barnett (1976, 1987), Barnett and Hahm (1994), Barnett and Zhou (1994), Barnett et al. (1995), and Hancock (1985, 1991). Here, we begin with Barnett (1978) definition of user costs and Hancock (1985) application of the user-cost approach to the banking firm. In doing so, we construct user costs (per unit, with the unit taken to be one dollar per period) for the services from all assets and liabilities in a financial firm’s balance sheet. Following Hancock (1985), we model banks that maximize the capitalized value of variable profit over a certain period of time. Since we estimate the model using quarterly data, we assume that banks can fully adjust the quantities of the goods to optimal values every period. The absence of adjustment costs effectively reduces our setup to a one-period profit maximization problem.

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

There are N banks in the economy and each bank operates with a set of assets A and a set of liabilities L. We let P_t denote a general price index in period t, y_{it}^k the real balance, and h_{it}^k the holding cost/revenue (per unit) of the kth financial good of bank i in period t. In addition, we denote w_{it} wage per efficiency unit, l_{it} number of efficiency units of labor, and $\bar{y}_{it}$ the quantity of a fixed in the short-run good of bank i in period t.

We denote $b_t \equiv \prod_{s=1}^{t}(1+R_{is})^{-1}$, where R_{is} is the ith bank discount rate in period s. Then the capitalized value of variable profit over T periods is

$$V_{iT} = \sum_{t=2}^{T} b_t \sum_{k\in A} \left[\left(1+h_{i,t-1}^k\right) y_{i,t-1}^k P_{t-1} - y_{it}^k P_t \right] + \sum_{t=2}^{T} b_t \sum_{k\in L} \left[y_{it}^k P_t - \left(1+h_{i,t-1}^k\right) y_{i,t-1}^k P_{t-1} \right] - \sum_{t=2}^{T} b_t w_{it} l_{it}. \quad (1)$$

The expression in the first line of (1) represents the capitalized net revenue to the firm from holding assets. For an asset k (such as a loan), the net revenue during period t is equal to the initial nominal asset, $y_{i,t-1}^k P_{t-1}$, plus holding revenue incurred at the rate $h_{i,t-1}^k$, $h_{i,t-1}^k y_{i,t-1}^k P_{t-1}$, minus the total nominal asset at the end of the period, $y_{it}^k P_t$. The expression in the second line of (1) represents the capitalized net cost of holding liabilities L and capitalized labor expenses. For a liability k (such as a deposit), the net cost during period t is equal to the nominal liability at the end of the period, $y_{it}^k P_t$, minus the initial nominal liability, $y_{i,t-1}^k P_{t-1}$, and holding costs incurred at the rate $h_{i,t-1}^k$, $h_{i,t-1}^k y_{i,t-1}^L P_{t-1}$. The negative of the coefficients of real balances, y_{it}^k, in equation (1) are beginning of the period nominal user costs—see also Barnett (1978). Thus, the beginning of period nominal user cost for assets is

$$u_{it}^k = \left(\frac{R_{it} - h_{it}^k}{1+R_{it}} \right) P_t, \quad k \in A, \quad (2)$$

and that for liabilities is

$$u_{it}^k = \left(\frac{h_{it}^k - R_{it}}{1+R_{it}} \right) P_t, \quad k \in L. \quad (3)$$

If a liability has reserve requirement, the holding cost of this liability is given by

$$h_{it}^k = h_{it}^{*k} + g^k \left(R_{it} - q^k\right), \quad (4)$$

where h_{it}^* is the explicit interest paid on the liability, g^k is the reserve requirement ratio (a flat rate), and q^k is the interest rate on required and excess reserves (we assume equal rates).

Equations (2) and (3) imply that the user costs may be positive or negative. The sign of the user cost permits monetary goods (such as cash and deposits of various types) and other financial goods (such as loans) to be classified as inputs or outputs. Those items with a positive user cost are classified as inputs (because

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

variable profit is reduced when the quantity is increased) and those with a negative user cost are classified as outputs (because variable profit is increased when the quantity is increased). With this classification of goods, we can also perform a change of variables, transferring the sign from the user costs to the quantities (so that the variable profit function in the next section has only nonnegative prices as arguments), as follows:

$$v_{it}^k = \left|u_{it}^k\right|,$$

and

$$x_{it}^k = -\text{sign}\left(u_{it}^k\right) \times y_{it}^k, \quad k \in A \cup L.$$

Further, we add wage and labor as $v_{it}^l = w_{it}$ and $x_{it}^l = l_{it}$. Then we let n be the size of $A \cup L$ plus one and denote $\boldsymbol{v}_{it} = (v_{it}^1, \ldots, v_{it}^n)$ the vector of absolute values of all user costs and wage and $\boldsymbol{x}_{it} = (x_{it}^1, \ldots, x_{it}^n)$ the vector of all quantities, with $x_{it}^k < 0$ for inputs (including labor) and $x_{it}^k \geq 0$ for outputs. Finally, we let $\bar{x} = \bar{y}$ denote the quantity of a fixed in the short-run input good. This setup is consistent with the neoclassical microeconomic theory of the firm—see, for example, Mas-Colell et al. (1995).

In constructing the user costs of financial goods (in Section 5), we deviate from the Hancock (1985) and Barnett and Hahm (1994) approach and follow Diewert et al. (2016). Hancock (1985) uses longitudinal observations on the balance sheet of 18 New York–New Jersey banks (all members of Federal Reserve District 2), over the period from 1973 to 1978. Barnett and Hahm (1994) also use longitudinal observations on 41 Chicago (Federal Reserve District 7) banks, over the period from 1979 to 1983. Both Hancock (1985) and Barnett and Hahm (1994) calculate holding and user costs using data on interest rates, deposit insurance premium rates, reserve requirement rates, and service charges. Here we follow Diewert et al. (2016) and construct holding and user costs using data on realized bank interest income and expenses. In particular, the holding cost (revenue) is the ratio of interest expenses (income) to the value of the corresponding asset or liability. To transform the holding costs into user costs we choose a time-varying bank specific discount rate, in particular, the weighted average cost of raising capital via deposits, debt, and equity. With this discount rate, bank deposits are mostly classified as outputs because typically deposits are the cheapest source of funds (see Basu et al. 2011 for discussion). We discuss data and measurement matters in detail in Section 5.

3. THE VARIABLE PROFIT FUNCTION

As in Hancock (1985), we use the profit function to obtain the functional forms for the estimating equations. In this section and Sections 4 and 5, we omit time and bank indexes for simpler exposition. A bank's profit maximization problem is

$$\pi\left(\boldsymbol{v}, \bar{x}\right) = \max_{\boldsymbol{x} \in S(\bar{x})} \sum_{k=1}^{n} v^k x^k, \tag{5}$$

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

where $\pi(\boldsymbol{v}, \bar{x})$ is the variable profit function and $S(\bar{x})$ is the production possibility set. The variable profit function is: (i) nondecreasing in output prices and nonincreasing in input prices; (ii) homogeneous of degree one in $\boldsymbol{v}$; (iii) continuous in $\boldsymbol{v}$; and (iv) convex in $\boldsymbol{v}$. It is worth noting that *a priori* it is not known whether a financial good is an input or an output.

In principle, assuming an explicit functional form for the variable profit function and having data on prices and observed profit, one could estimate equation (5) directly. However, we can substantially improve the accuracy of the estimation if we simultaneously estimate the system of supply and demand functions induced by the variable profit function. We obtain the system of supplies of outputs and demands for inputs using Hotelling's lemma, differentiating (5) with respect to prices

$$x^i(\boldsymbol{v}, \bar{x}) = \frac{\partial \pi(\boldsymbol{v}, \bar{x})}{\partial v^i}, \quad i = 1, \ldots, n. \tag{6}$$

In using Hotelling's lemma, we implicitly assume that the bank is a price taker. While the degree of competition varies across different service lines of commercial banking and over time, the level of concentration in the U.S. commercial banking industry over the sample period is relatively low. Bolt and Humphrey (2015) estimate that in 2010 the average HHI for banks with total assets in excess of \$1 billion is 1364. For banks with total assets between \$100 million and \$1 billion, the HHI is 1132. According to the U.S. Justice Department's 2010 horizontal merger guidelines, markets with an HHI below 1500 can be considered to be unconcentrated. Bolt and Humphrey (2012) also propose a relative competition efficiency measure and find that, out of five service lines, business loans and security activities are most competitive while investment banking and other fee-based activities are least competitive. In this paper, we abstract from imperfectly competitive investment banking services such as securitization, underwriting, and securities brokerage. In Section 5, we discuss the classification of financial goods for the purpose of this study.

Estimation of (6) allows the calculation of own- and cross-price elasticities of supply and demand for the financial goods. These elasticities can then be used to investigate the effects of interest rate and user-cost changes on the production of financial goods (including both inputs and outputs). In particular, the elasticity of transformation can be calculated from the Hessian matrix, $\boldsymbol{H}$, as follows:

$$\sigma_{ij} = \pi \frac{\partial^2 \pi}{\partial v^i \partial v^j} \left[\frac{\partial \pi}{\partial v^i} \frac{\partial \pi}{\partial v^j} \right]^{-1} = \frac{\pi \boldsymbol{H}_{ij}}{x^i x^j}, \quad i, j = 1, \ldots, n, \tag{7}$$

and the compensated price elasticities of supply and demand as

$$\eta_{ij} = \sigma_{ij} \frac{v^i x^i}{\pi}, \quad i, j = 1, \ldots, n. \tag{8}$$

4. THE GENERALIZED SYMMETRIC BARNETT FORM

In this section, we discuss the flexible functional form that we use to approximate the unknown underlying variable profit function (5): the symmetric generalized Barnett form, introduced in Barnett and Hahm (1994). As shown by Diewert and Wales (1987), the symmetric generalized Barnett functional form is locally quasi-flexible. They define that to mean that it can locally attain all first derivatives, all levels, and all second derivatives, except for those in one column of the Hessian (and its corresponding identical row, since symmetric). The number of derivatives potentially missed by the symmetric generalized Barnett functional form is linear in the number of goods, while the number attained is quadratic in the number of goods, so many more than the ones possibly missed. Hence, while the generalized Barnett functional form cannot exactly attain all possible elasticities at a point, it comes very close. Its big advantage is that it has much better global regularity properties than many other well-known flexible functional forms.

The generalized symmetric Barnett variable profit function is

$$\pi(\boldsymbol{v},\bar{x}) = \left(\sum_{i=1}^{n} a_{ii} v^i - 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij} \left(v^i v^j\right)^{1/2} + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \sum_{\substack{k=j+1 \\ k\neq i}}^{n} a_{ijk} \left(v^i\right)^2 \left(v^j v^k\right)^{-1/2} \right) \bar{x}, \tag{9}$$

where $a_{ij} \geq 0$ and $a_{ijk} \geq 0$. With this specification, the conditional variable profit function (9) is linearly homogeneous and globally convex in prices, $\boldsymbol{v}$. The convexity of the function follows from convexity of the summands in (9) and nonnegativity of the coefficients a_{ij} and a_{ijk}. Using duality and the envelope theorem, we obtain the demand and supply system induced by (9):

$$\begin{aligned} x^i/\bar{x} = a_{ii} &- \sum_{j:j>i} a_{ij} \left(v^j/v^i\right)^{1/2} - \sum_{j:j<i} a_{ji} \left(v^j/v^i\right)^{1/2} \\ &+ 2 \sum_{j:j\neq i} \sum_{k:k>j,k\neq i} a_{ijk} v^i \left(v^j v^k\right)^{-1/2} \\ &- \frac{1}{2} \sum_{j:j\neq i} \sum_{k:k>i,k\neq j} a_{jik} \left(v^j/v^i\right)^2 \left(v^i/v^k\right)^{1/2} \\ &- \frac{1}{2} \sum_{j:j\neq i} \sum_{k:k<i,i\neq j} a_{jki} \left(v^j/v^i\right)^2 \left(v^i/v^k\right)^{1/2}. \end{aligned} \tag{10}$$

With n goods, the generalized Barnett profit function (9) contains $n\left(n^2 - 2n + 3\right)/2$ free parameters. For $n = 6$ (as in our case), the number of free parameters is 81.

The convexity of the generalized symmetric Barnett variable profit function is ensured by imposing nonnegativity constraints on the parameters a_{ij} and a_{ijk}. In

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

particular, we reparameterize the model such that these coefficients are squared new parameters. See Barnett (1976, 1983) for a detailed discussion of squaring techniques and the asymptotic properties of such nonnegative estimators.

5. DATA AND MEASUREMENT MATTERS

We use quarterly data on the U.S. commercial banking sector, over the period from 1992 to 2016, obtained from the quarterly Uniform Bank Performance Reports (refined data from Call Reports) provided by the Federal Deposit Insurance Corporation. The sample covers federal and state chartered commercial banks, savings banks, savings associations (and as of July 21, 2011 thrifts), and insured U.S. branches of foreign chartered institutions. Although the range of activity of the savings associations has substantially expanded over the period under consideration, banks and savings associations are still distinct institutions subject to different regulation. In our analysis, we exclude from the sample savings associations and thrifts which account on average for 10.6% of total assets.

The resulting panel which contains 780,825 observations is unbalanced: only one third of all banks cover the entire 25 year period. During the period from 1992 to 2016 the number of banks declined from 11,535 to 5174 because of the intense consolidation in the U.S. banking industry. At the same time, the assets became more concentrated in the largest financial institutions, as can be seen in Table 1, which shows the distribution of assets over four groups of banks in 1992, 2002, and 2012. Our use of the unbalanced panel reduces the survival bias in parameter estimates. Olley and Pakes (1996) show that using an artificially balanced sample can lead to significant bias in parameter estimates due to the survival effect. They also demonstrate that an explicit selection correction has insignificant effect if the estimation is based on the unbalanced sample.

We focus on commercial banking and abstract from the possible interaction between commercial and investment activity. Yet, we take into account most of the noninterest expenses via modeling the demand for labor (which on average amounts to more than 70% of noninterest expenses) and including bank premises and fixed assets as a quasi-fixed factor. For example, in the case of JP Morgan Chase, according to its 2012 financial statements, the noninterest expenses amount to \$53 billion while the interest expenses are \$6 billion. To address heterogeneity, we estimate the model using subsamples formed based on the total amount of assets held. For instance, we separate small banks with average assets less than \$100 million and large banks with average assets exceeding \$1 billion (in 2009 U.S. dollars). Some of these large banks are bank-holding companies (and financial holding companies). For example, Morgan Stanley, a bank-holding company since 2008, is a major generator of electricity, and JP Morgan Chase controls the U.S. copper warehouse market.

We consider five variable financial goods, two assets, and three liabilities. They are debt securities and trading accounts (x^1), loans and leases (x^2), deposits (x^3),

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

TABLE 1. Size distribution of U.S. banks

Assets	1992 Number of banks	1992 Assets held	1992 Share of assets held (%)	2002 Number of banks	2002 Assets held	2002 Share of assets held (%)	2012 Number of banks	2012 Assets held	2012 Share of assets held (%)
Less than $100 million	8282	350	9.6	4125	213	2.9	1886	112	0.8
$100 million–$1 billion	3024	749	20.5	3583	963	13.1	3804	1152	8.5
$1 billion–10 billion	368	1117	30.5	359	1022	13.9	484	1275	9.4
More than $10 billion	51	1445	39.5	87	5175	70.2	88	11,042	81.3
Total	11,725	3662	100	8154	7374	100	6262	13580	100

Note: Bank assets are in billions of dollars.

TABLE 2. Assets and liabilities of U.S. banks

Year	Debt securities	Loans & leases	Fixed assets	Deposits	Other debt	Equity
1992	896.17	2350.14	101.55	2887.79	565.11	281.24
1993	1005.39	2494.14	95.55	2946.10	680.30	316.89
1994	1070.46	2732.42	98.06	3064.60	859.42	334.22
1995	1084.02	3018.06	104.07	3232.89	972.23	373.98
1996	1089.69	3287.96	120.50	3402.97	1045.93	400.43
1997	1208.68	3401.37	138.69	3612.81	1219.45	442.61
1998	1305.46	3685.39	161.61	3879.98	1353.08	489.22
1999	1342.44	3930.80	181.33	4030.09	1491.92	507.28
2000	1414.68	4265.11	189.30	4381.51	1597.59	558.78
2001	1540.90	4345.46	209.95	4614.87	1649.16	626.92
2002	1810.89	4605.62	218.11	4941.77	1809.33	683.18
2003	1995.99	4916.43	258.07	5301.89	1950.13	730.46
2004	2133.22	5430.44	383.96	5831.25	2044.53	892.65
2005	2142.43	5908.73	410.97	6312.85	2120.03	949.66
2006	2342.17	6579.50	471.59	6955.72	2397.20	1064.49
2007	2505.32	7324.41	549.58	7532.34	2788.04	1178.11
2008	2728.04	7350.35	536.23	8324.19	3153.02	1186.69
2009	2953.44	6862.70	544.73	8584.28	2240.37	1343.40
2010	3128.71	7096.62	542.43	8775.87	2239.20	1401.41
2011	3308.29	7354.94	512.53	9526.52	2023.54	1458.99
2012	3521.69	7747.82	510.09	10,307.37	1929.42	1541.98
2013	3393.49	8019.97	505.09	10,691.34	1829.34	1570.92
2014	3644.91	8406.48	489.11	11,271.19	2003.25	1658.05
2015	3657.74	8957.46	486.42	11,706.57	1947.34	1728.12
2016	3830.65	9362.53	492.19	12,330.07	1998.84	1791.26

Note: The values are in billions of U.S. dollars.

other debt (that is, debt other than deposits) (x^4), and equity (x^5). We also use a nonfinancial input, labor (x^6), and a quasi-fixed good, bank premises, and fixed assets ($\bar{x}$). The level of aggregation for these financial goods is mainly due to data limitation. Aggregation bias may be present, for example, because we combine demand deposits and term deposits with different maturities. In Table 2, we list asset and liability values (in billions of dollars) for each of the 25 years in the sample for each of the goods. The loans and leases on average account for over 60% of assets and debt securities and trading accounts account for about 26% of assets. Together with bank premises and fixed assets, these two assets on average account for about 88% of all assets. Deposits on average account for about 84%, other debt for about 5%, and equity capital for close to 11% of total liabilities.

Next we calculate the user costs of the financial goods based on realized bank interest income and expenses from holding these goods. Given available data, these *ex post* average realized user costs provide a reasonable approximation for

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

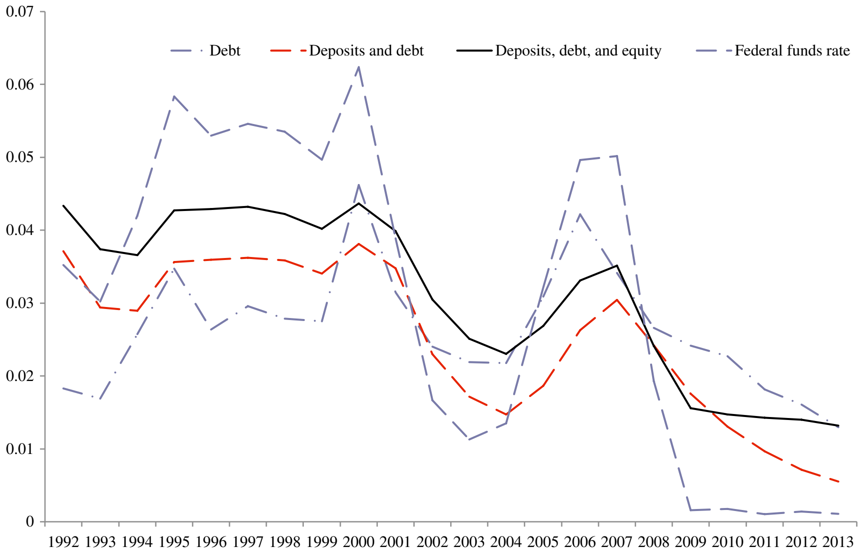
the *ex ante* marginal user costs defined by (2) and (3). We let $r^i, i = 1, \ldots, 5$, denote the interest income on the corresponding asset or the interest expense on the corresponding liability. We use the net income r^5 as a proxy to calculate the return to equity capital. Then, using the values of interest and noninterest income and expenses we calculate the holding costs and revenues for the assets and liabilities, as $h^i = r^i/y^i$ for $i = 1, \ldots, 5$.

According to equations (2) and (3), user costs are linear transformation of holding costs parameterized by the discount rate. By nature, the discount rate is specific for each bank and could vary over time reflecting a bank's riskiness. In general, there are several recognized methods of choosing the discount rate, such as, for example, the weighted average cost of capital (WACC) for mixed capital and the capital asset pricing model (CAPM) for equity capital. However, the literature shows no consensus on how to determine the discount rate for a bank. For example, a contentious issue in calculating a bank's WACC is accounting for the cost of deposits which on average account for more than 80% of the total liabilities of a commercial bank. At the same time, the interest rates paid by banks on deposit balances are usually quite low and do not reflect noninterest expenses associated with attracting and servicing deposits. In this regard, Diewert et al. (2016) discuss three options for the choice of the discount rate: (i) the average cost of raising financial capital via debt other than deposits; (ii) the weighted average cost of raising capital via deposits and debt; and (iii) the weighted average cost of raising capital via deposits, debt, and equity.

Although each of these methods can provide a reasonable proxy for the discount rate, the choice depends on the composition of bank liabilities. The first method might provide a biased estimate of the discount rate if the share of debt other than deposits in all liabilities is small. In our sample, the average share of debt other than deposits is about 5% of total liabilities and equity capital. At the same time, the second method can produce a significantly downward biased estimate because (1) the interest rate on deposits underestimates the cost of this source of funds (see Basu and Wang 2013 for a discussion) and (2) the method excludes equity capital which is, typically, the most expensive source of bank capital. Figure 1 shows the dynamics of the (across banks) average reference discount rates calculated using each of the three methods. In what follows, we choose the third method and calculate the weighted average cost of deposits, debt, and equity for each bank in every period.

Although there is a prior belief about which financial goods are inputs and which are outputs, the actual nature of these goods is determined by the sign of the corresponding user costs. Moreover, in our framework the sign of a particular user cost for any particular bank could change over time because both holding costs and discount rates vary over time. In particular, the sign of the user costs depends on the sign of the numerator in (2) and (3). In Table 3, we report the (across banks) average nominal user costs corresponding to the assets, y^1 and y^2, and the liabilities, y^3, y^4, and y^5, for each year from 1992 to 2016. For each year in the sample, we calculate the percentage of observations in the pooled sample

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press



Notes: The figure plots the dynamics of three possible average reference discount rates: (1) the cost of debt (other than deposits) capital, (2) the cost of deposits and other debt, and (3) the cost of deposits, other debt, and equity capital. The effective federal funds rate is presented as a reference.

FIGURE 1. Average reference discount rates.

of all 780,825 observations when each financial good is an input (i.e., the user cost is positive) and report the results in Table 3. On average, the most stable output is loans and leases, with a negative user costs in more than 99% of the observations. Another asset, debt securities and trading accounts, is an output in 92% of the observations. Deposits have a negative user cost in about 88% of the observations. Debt other than deposits is an output in about 67% of the observations. Finally, equity is an output only in 12% of the observations. Because of low or even negative return on equity capital during (and after) the period of the financial crisis, 2007–2011, equity became a relatively inexpensive source of funds and the observations with a positive user cost during this period dropped to the minimum of 81%. Finally, to obtain real quantities, we use the GDP deflator (from the St. Louis Fed FRED database) to proxy the general price index.

6. ECONOMETRIC ISSUES

We estimate the system of input demands and output supplies for the generalized symmetric Barnett model using the maximum likelihood method. We do not estimate the variable profit function itself, because it contains no additional information. In order to do so, we add a stochastic component, $\boldsymbol{\epsilon}_{it}$, as follows:

$$z_{it} = \boldsymbol{w}\left(\boldsymbol{v}_{it}, \boldsymbol{\theta}\right) + \boldsymbol{\epsilon}_{it}, \tag{11}$$

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

TABLE 3. User costs averaged across banks

Year	Debt securities	Loans & leases	Deposits	Other debt	Equity
1992	−0.364 / 0.000	−1.028 / 0.000	−0.217 / 0.000	−1.884 / 0.003	1.757 / 1.000
1993	−0.367 / 0.028	−0.946 / 0.000	−0.161 / 0.072	−0.450 / 0.108	1.641 / 0.936
1994	−0.343 / 0.021	−0.947 / 0.000	−0.158 / 0.064	−0.354 / 0.182	1.554 / 0.941
1995	−0.315 / 0.032	−0.959 / 0.000	−0.154 / 0.068	−0.359 / 0.250	1.451 / 0.935
1996	−0.316 / 0.033	−0.956 / 0.000	−0.155 / 0.072	−0.445 / 0.176	1.475 / 0.930
1997	−0.330 / 0.031	−0.946 / 0.000	−0.161 / 0.068	−0.409 / 0.205	1.488 / 0.933
1998	−0.310 / 0.051	−0.951 / 0.000	−0.156 / 0.085	−0.406 / 0.205	1.411 / 0.913
1999	−0.325 / 0.041	−0.940 / 0.000	−0.157 / 0.096	−0.337 / 0.253	1.424 / 0.903
2000	−0.342 / 0.036	−0.942 / 0.001	−0.162 / 0.098	−0.217 / 0.376	1.428 / 0.897
2001	−0.325 / 0.067	−0.955 / 0.000	−0.149 / 0.122	−0.315 / 0.276	1.319 / 0.877
2002	−0.340 / 0.055	−0.961 / 0.000	−0.186 / 0.079	−0.233 / 0.308	1.608 / 0.921
2003	−0.258 / 0.114	−0.971 / 0.000	−0.199 / 0.061	−0.153 / 0.348	1.678 / 0.934
2004	−0.278 / 0.086	−0.955 / 0.000	−0.212 / 0.050	−0.125 / 0.355	1.766 / 0.943
2005	−0.242 / 0.117	−0.978 / 0.000	−0.219 / 0.050	−0.085 / 0.433	1.822 / 0.943
2006	−0.218 / 0.156	−0.995 / 0.000	−0.205 / 0.068	−0.099 / 0.448	1.677 / 0.917
2007	−0.245 / 0.141	−1.011 / 0.000	−0.176 / 0.111	−0.186 / 0.373	1.407 / 0.865
2008	−0.378 / 0.047	−1.023 / 0.000	−0.165 / 0.150	−0.144 / 0.389	1.342 / 0.830
2009	−0.413 / 0.044	−1.063 / 0.000	−0.148 / 0.189	−0.035 / 0.478	1.259 / 0.805
2010	−0.343 / 0.077	−1.094 / 0.000	−0.174 / 0.137	−0.025 / 0.466	1.485 / 0.857
2011	−0.300 / 0.098	−1.119 / 0.000	−0.188 / 0.113	−0.054 / 0.427	1.567 / 0.886
2012	−0.238 / 0.149	−1.118 / 0.000	−0.205 / 0.080	−0.067 / 0.383	1.690 / 0.920
2013	−0.225 / 0.155	−1.096 / 0.000	−0.212 / 0.064	−0.095 / 0.333	1.798 / 0.938
2014	−0.253 / 0.110	−1.074 / 0.000	−0.223 / 0.048	−0.124 / 0.293	1.834 / 0.952
2015	−0.240 / 0.120	−1.058 / 0.000	−0.228 / 0.048	−0.148 / 0.261	1.850 / 0.952
2016	−0.242 / 0.107	−1.054 / 0.000	−0.234 / 0.037	−0.151 / 0.257	1.862 / 0.961

Notes: The table shows the user costs of financial goods averaged across banks in each year (the first number) and the fraction of observations when financial goods are inputs (the second number).

where $z_{it} = (z_{it}^1, ..., z_{it}^n)$ is the vector of input demands and output supplies. We assume that $\boldsymbol{\epsilon}_{it} \sim N(\mathbf{0}, \boldsymbol{\Omega})$, where $\mathbf{0}$ is a null vector and $\boldsymbol{\Omega}$ is the $n \times n$ unknown symmetric positive definite error covariance matrix. A vector-valued function $\boldsymbol{w}(\boldsymbol{v}_{it}, \boldsymbol{\theta}) = (w^1(\boldsymbol{v}_{it}, \boldsymbol{\theta}), ..., w^n(\boldsymbol{v}_{it}, \boldsymbol{\theta}))$ with $w^k(\boldsymbol{v}_{it}, \boldsymbol{\theta})$ is given by the right-hand side of (10). The full log-likelihood function for the system (11) over the pooled panel is

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{t=1}^{T}\left(nN_t \log(2\pi) - N_t \log(|\boldsymbol{\Omega}|)\right) - \frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{N_t}\left(z_{it} - \boldsymbol{w}(\boldsymbol{v}_{it}, \boldsymbol{\theta})\right)' \boldsymbol{\Omega}^{-1}\left(z_{it} - \boldsymbol{w}(\boldsymbol{v}_{it}, \boldsymbol{\theta})\right) \qquad (12)$$

where N_t denotes the number of banks in quarter t. The coefficients of the approximating form in $\boldsymbol{\theta}$ must be estimated together with the covariance matrix Ω. The log-likelihood function in (12) is computationally cumbersome, especially if the

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

dimensionality is high and the size of the sample is large. In our estimation, we assume no autocorrelation and use the concentrated log-likelihood function (see Greene (2012), p. 551 for more details)

$$\mathcal{L}_c\left(\boldsymbol{v}_t, \bar{x}_t, \boldsymbol{\theta}\right) = -\sum_{t=1}^{T} \frac{N_t}{2}\Big[n\left(1+\log(2\pi)\right) + \log|\boldsymbol{W}|\Big],$$

where

$$\boldsymbol{W} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_t}\sum_{i=1}^{N_t}\big(z_{it} - \boldsymbol{w}\left(\boldsymbol{v}_{it}, \boldsymbol{\theta}\right)\big)'\big(z_{it} - \boldsymbol{w}\left(\boldsymbol{v}_{it}, \boldsymbol{\theta}\right)\big).$$

To correct for possible heteroskedasticity in (11) we use the Huber–White estimator for the asymptotic covariance matrix.

The estimation is performed in C++ using the concentrated log-likelihood function. The regularity conditions of the variable profit function are checked as follows:

- Monotonicity requires that the variable profit function is nondecreasing in output prices and nonincreasing in input prices. Since $v_i > 0$ (by definition) and $\pi > 0$, monotonicity is checked by direct computation of the estimated expenditure on each good relative to variable profit, since $\text{sign}(\partial\pi/\partial v_i) = \text{sign}(v_i x_i/\pi)$.
- Convexity requires the Hessian matrix of the variable profit function, $\boldsymbol{H}$, to be positive semidefinite. It is checked by performing a Cholesky factorization of that matrix and checking whether the Cholesky values are nonnegative [since a matrix is positive semidefinite if its Cholesky factors are nonnegative—see Lau (1978, Theorem 3.2)].

Finally, using the parameter estimates $\widehat{\boldsymbol{\theta}}$ we calculate elasticities of transformation and compensated price elasticities using equations (7) and (8), respectively. We also apply the first-order Taylor expansion (the Delta method) to obtain the standard errors for these elasticities. For example, the standard errors of the price elasticities are

$$\sigma\left(\eta_{ij}\right) = \sqrt{\left(\nabla_{\boldsymbol{\theta}}\eta_{ij}|_{\hat{\boldsymbol{\theta}}}\right)' Var(\hat{\boldsymbol{\theta}})\, \nabla_{\boldsymbol{\theta}}\eta_{ij}|_{\hat{\boldsymbol{\theta}}}}$$

where $\nabla_{\boldsymbol{\theta}}\eta_{ij}|_{\hat{\boldsymbol{\theta}}}$ is the gradient of the elasticity η_{ij} with respect to $\boldsymbol{\theta}$ evaluated at $\hat{\boldsymbol{\theta}}$ and $Var(\hat{\boldsymbol{\theta}})$ is the variance–covariance matrix of the estimate $\hat{\boldsymbol{\theta}}$.

7. EMPIRICAL EVIDENCE

7.1. Theoretical Regularity

In this section, we provide the estimates of the generalized Barnett system and discuss the elasticities of transformation and compensated price elasticities. Table 4 contains parameter estimates and the percentage of theoretical regularity (convexity and monotonicity) violations for the model. The model is estimated with the convexity conditions imposed globally.

TABLE 4. Generalized Barnett parameter estimates

	Assets		
Parameter	All banks	Less than $100 million	More than $1 billion
a_{11}	−6.731 (2.197)	−1.874 (0.615)	−0.930 (0.908)
a_{12}	−1.138 (0.344)	−0.611 (0.205)	−1.056 (0.870)
a_{13}	3.427 (1.255)	2.848 (0.942)	−0.259 (0.152)
a_{14}	−0.131 (0.042)	0.038 (0.013)	−0.999 (0.452)
a_{15}	0.985 (0.273)	−0.035 (0.012)	−0.170 (0.011)
a_{16}	3.448 (0.977)	−15.042 (5.036)	1.214 (0.814)
a_{22}	2.077 (2.018)	−2.297 (1.924)	0.838 (0.848)
a_{23}	−7.628 (6.945)	2.791 (0.988)	0.886 (0.950)
a_{24}	0.134 (0.044)	−1.441 (0.483)	−0.309 (0.264)
a_{25}	4.710 (1.716)	−4.232 (1.411)	0.694 (0.378)
a_{26}	0.551 (0.537)	24.697 (21.916)	0.757 (0.757)
a_{33}	11.322 (12.870)	1.428 (0.483)	−1.161 (1.075)
a_{34}	1.440 (0.461)	−3.019 (1.013)	0.116 (0.103)
a_{35}	−3.541 (3.123)	0.230 (0.077)	1.220 (1.194)
a_{36}	2.604 (2.168)	9.451 (3.289)	0.909 (0.921)
a_{44}	−2.431 (0.748)	−6.285 (2.105)	2.313 (0.150)
a_{45}	0.234 (0.072)	−0.144 (0.048)	0.088 (0.068)
a_{46}	2.238 (0.739)	6.510 (2.184)	0.006 (0.006)
a_{55}	2.648 (1.454)	−8.367 (2.789)	0.149 (0.168)
a_{56}	4.467 (1.362)	0.237 (0.079)	−7.302 (3.195)
a_{66}	−4.358 (4.358)	−0.478 (0.479)	0.593 (0.593)
a_{123}	0.411 (0.125)	−0.277 (0.086)	−0.015 (0.008)
a_{124}	−0.000 (0.000)	−0.005 (0.002)	0.126 (0.059)
a_{125}	−0.752 (0.250)	−0.769 (0.256)	0.385 (0.253)
a_{126}	−0.714 (0.210)	−6.122 (1.781)	1.675 (0.671)
a_{134}	0.000 (0.000)	−0.001 (0.000)	0.007 (0.004)
a_{135}	0.079 (0.024)	0.007 (0.002)	0.035 (0.020)
a_{136}	0.018 (0.006)	0.126 (0.039)	0.477 (0.164)
a_{145}	−0.000 (0.000)	−0.007 (0.003)	−0.041 (0.022)
a_{146}	0.000 (0.000)	0.002 (0.001)	0.053 (0.025)
a_{156}	0.372 (0.121)	0.411 (0.137)	0.041 (0.017)
a_{167}	−0.007 (0.002)	−0.007 (0.002)	0.006 (0.002)
a_{166}	0.000 (0.000)	0.000 (0.000)	−0.010 (0.005)
a_{213}	0.003 (0.001)	0.047 (0.016)	−0.030 (0.009)
a_{214}	−0.007 (0.002)	0.011 (0.004)	−0.014 (0.005)
a_{215}	−0.000 (0.000)	0.000 (0.000)	−0.039 (0.020)
a_{216}	0.013 (0.004)	−0.000 (0.000)	0.035 (0.020)
a_{234}	0.008 (0.003)	−0.016 (0.005)	0.018 (0.005)
a_{235}	0.000 (0.000)	−0.002 (0.001)	0.012 (0.006)
a_{236}	−0.000 (0.000)	−0.001 (0.000)	−0.010 (0.005)
a_{245}	−0.010 (0.003)	−0.055 (0.018)	1.364 (0.138)
a_{246}	−0.211 (0.070)	−1.098 (0.364)	−0.043 (0.016)
a_{256}	−0.000 (0.000)	−0.007 (0.002)	0.027 (0.011)
a_{267}	0.087 (0.029)	0.189 (0.063)	0.208 (0.075)

TABLE 4. (Continued)

Parameter	Assets: All banks	Less than $100 million	More than $1 billion
a_{266}	−0.056 (0.019)	0.112 (0.037)	−0.019 (0.007)
a_{312}	−0.001 (0.000)	−0.031 (0.010)	0.075 (0.032)
a_{314}	−14.056 (2.646)	−4.459 (1.486)	0.251 (0.236)
a_{315}	2.332 (4.741)	−1.731 (0.610)	5.009 (3.841)
a_{316}	−0.000 (0.000)	0.024 (0.008)	−0.077 (0.033)
a_{324}	0.000 (0.000)	−0.044 (0.015)	0.049 (0.019)
a_{325}	0.773 (0.041)	1.559 (0.520)	−0.758 (0.620)
a_{326}	0.048 (0.016)	0.004 (0.001)	−0.251 (0.045)
a_{345}	0.009 (0.003)	−0.185 (0.061)	0.350 (0.027)
a_{346}	0.009 (0.003)	0.009 (0.003)	0.123 (0.017)
a_{356}	−0.008 (0.003)	−0.039 (0.013)	−0.002 (0.000)
a_{367}	0.398 (0.114)	0.020 (0.007)	0.367 (0.012)
a_{366}	−0.156 (0.048)	−0.036 (0.012)	−0.453 (0.348)
a_{412}	0.367 (0.096)	2.789 (0.929)	0.818 (0.242)
a_{413}	0.031 (0.009)	0.031 (0.010)	0.049 (0.018)
a_{415}	0.112 (0.032)	0.074 (0.025)	0.127 (0.008)
a_{416}	0.278 (0.085)	0.046 (0.015)	−0.583 (0.154)
a_{423}	−0.010 (0.003)	−0.009 (0.003)	−0.006 (0.002)
a_{425}	−0.005 (0.002)	−0.006 (0.002)	−0.001 (0.000)
a_{426}	0.000 (0.000)	0.000 (0.000)	−0.001 (0.000)
a_{435}	0.006 (0.002)	0.001 (0.000)	0.002 (0.001)
a_{436}	1.142 (0.475)	1.960 (0.653)	0.032 (0.033)
a_{456}	−0.000 (0.000)	0.001 (0.000)	−0.005 (0.002)
a_{467}	0.376 (0.076)	−0.060 (0.020)	−0.004 (0.004)
a_{466}	0.000 (0.000)	0.000 (0.000)	0.020 (0.009)
a_{512}	0.190 (0.071)	0.087 (0.029)	−0.505 (0.571)
a_{513}	0.000 (0.000)	−0.000 (0.000)	0.006 (0.003)
a_{514}	−0.188 (0.063)	−0.997 (0.341)	−0.011 (0.004)
a_{516}	−0.091 (0.031)	0.248 (0.082)	0.235 (0.072)
a_{523}	−0.000 (0.000)	−0.017 (0.006)	0.001 (0.001)
a_{524}	0.035 (0.012)	−3.210 (1.070)	−0.015 (0.007)
a_{526}	1.813 (0.571)	0.398 (0.129)	1.006 (0.924)
a_{534}	0.000 (0.000)	−0.102 (0.034)	−0.223 (0.121)
a_{536}	−8.137 (2.979)	−9.430 (3.143)	0.242 (0.238)
a_{546}	0.000 (0.000)	0.086 (0.028)	−0.101 (0.055)
a_{567}	−0.193 (0.060)	−0.390 (0.131)	0.222 (0.193)
a_{566}	−0.000 (0.000)	0.004 (0.001)	0.215 (0.127)
Observations	780,825	389,823	48,457
Violations (%)			
Convexity	0	0	0
Monotonicity	21	0	14

Note: Standard errors in parentheses.

We allow for several types of bank heterogeneity. First, banks can be heterogeneous in the level of bank profit captured by the intercept of the profit function. The duality transformation eliminates a potentially heterogenous intercept from the estimated demand and supply system of equations (11). Second, we control for bank heterogeneity by estimating the model for several bank size classes, acknowledging differences in the degree of competition and institutional structure among small and large banks. In particular, we estimate the model for three bank groups, based on asset size, as follows: all banks (780,825 observations), banks with assets less than \$100 million (389,823 observations), and banks with assets in excess of \$1 billion (48,457 observations). We cannot use bank dummy variables as bank fixed effects because of the large number of banks. Moreover, the nonlinearity of the demand and supply system precludes from using the within or the first difference transformations.

As can be seen in Table 4, the generalized Barnett model satisfies convexity of the variable profit function at every data point in each of the samples. We find, however, that the imposition of convexity does not always assure economic regularity, as there are monotonicity violations at some data points; only in the case of banks with assets less than \$100 million, we report zero monotonicity and zero convexity violations. Our evidence regarding economic regularity supports Barnett's (2002, p. 199) argument that "although unconstrained specifications of technology are more likely to produce violations of curvature than monotonicity, I believe that induced violations of monotonicity become common, when curvature alone is imposed. Hence, the now common practice of equating regularity with curvature is not justified." Although convexity of the variable profit function is not sufficient for regularity, we believe that in the context of our model the convexity condition is the most crucial. That is, although monotonicity is a desirable property, in our framework, since each financial good may be an input in one period and an output in another period, the monotonicity condition indicates whether the model can predict the sign of a financial good at a point in the sample.

In what follows, we use the generalized Barnett model to produce inferences about the elasticities of transformation, σ_{ij}, and the compensated price elasticities, η_{ij}. In this regard, according to equations (7) and (8), the Hessian matrix of the variable profit function, $\boldsymbol{H}$, is the basis for the calculation of the elasticities of transformation and the price elasticities of supply and demand. The satisfaction of the convexity condition with the generalized Barnett model suggests that our estimates of the elasticities of transformation and the own- and cross-compensated price elasticities (reported in what follows) are well behaved.

7.2. Elasticities of Transformation

The estimated (symmetrical) elasticities of transformation, σ_{ij}, calculated using equation (7) at the sample means, are shown in Table 5 for each of the three bank samples. We expect the on-diagonal elements for all six goods to be positive and this expectation is clearly achieved. The off-diagonal elements indicate the

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

TABLE 5. Elasticities of transformation

	Asset class	Debt securities	Loans and leases	Deposits	Other debt	Equity	Employees
Debt securities	all banks	2.545 (0.797)	−0.003 (0.000)	−0.047 (0.042)	−0.028 (0.027)	0.002 (0.003)	2.168 (1.421)
	< $100 million	12.766 (0.744)	−0.005 (0.000)	−0.209 (0.103)	−0.975 (1.084)	0.003 (0.004)	1.760 (2.687)
	> $1 billion	7.499 (0.001)	7.900 (0.002)	−0.382 (0.000)	−0.002 (0.000)	0.000 (0.000)	3.726 (0.001)
Loans	all banks	−0.003 (0.000)	0.000 (0.000)	−0.000 (0.000)	−0.001 (0.003)	0.000 (0.000)	−0.003 (0.002)
	< $100 million	−0.005 (0.000)	0.000 (0.000)	−0.003 (0.021)	−0.000 (0.000)	0.006 (0.009)	−0.001 (0.001)
	> $1 billion	7.900 (0.002)	64.839 (0.005)	−3.386 (0.000)	−0.003 (0.000)	0.000 (0.000)	4.396 (0.001)
Deposits	all banks	−0.047 (0.042)	−0.000 (0.000)	0.263 (0.213)	−0.008 (0.009)	0.385 (0.330)	0.184 (0.142)
	< $100 million	−0.209 (0.103)	−0.003 (0.021)	8.013 (4.575)	−2.199 (0.231)	0.293 (1.623)	0.905 (4.781)
	> $1 billion	−0.382 (0.000)	−3.386 (0.000)	39.157 (0.026)	1.631 (0.000)	−0.000 (0.000)	2.972 (0.001)
Other debt	all banks	−0.028 (0.027)	−0.001 (0.003)	−0.008 (0.009)	16.717 (21.740)	0.012 (0.107)	1.302 (2.710)
	< $100 million	−0.975 (1.084)	−0.000 (0.000)	−2.199 (0.231)	14.033 (3.792)	0.079 (0.198)	0.006 (0.002)
	> $1 billion	−0.002 (0.000)	−0.003 (0.000)	1.631 (0.000)	0.013 (0.000)	0.000 (0.000)	0.063 (0.000)
Equity	all banks	0.002 (0.003)	0.000 (0.000)	0.385 (0.330)	0.012 (0.107)	0.586 (0.533)	−0.023 (0.026)
	< $100 million	0.003 (0.004)	0.006 (0.009)	0.293 (1.623)	0.079 (0.198)	23.153 (15.647)	−4.895 (2.666)
	> $1 billion	0.000 (0.000)	0.000 (0.000)	−0.000 (0.000)	0.000 (0.000)	5.660 (0.001)	−4.302 (0.001)
Employees	all banks	2.168 (1.421)	−0.003 (0.002)	0.184 (0.142)	1.302 (2.710)	−0.023 (0.026)	14.847 (5.755)
	< $100 million	1.760 (2.687)	−0.001 (0.001)	0.905 (4.781)	0.006 (0.002)	−4.895 (2.666)	1.962 (3.850)
	> $1 billion	3.726 (0.001)	4.396 (0.001)	2.972 (0.001)	0.063 (0.000)	−4.302 (0.001)	6.552 (0.001)

Notes: The elasticities are based on the estimates of the generalized symmetric Barnett variable profit function. The sample includes quarterly data over the 1992–2017 period. The elasticities are reported at the sample mean. Standard errors are in parentheses.

degree of substitutability or complementarity between financial goods. In particular, between two inputs or two outputs, if $\sigma_{ij} > 0$, they are substitutes, and if $\sigma_{ij} \leq 0$, they are complements. In the case of one output and one input, if $\sigma_{ij} > 0$, they are complements, and if $\sigma_{ij} \leq 0$, they are substitutes. As is evident from Table 3, the sign of the user-cost changes for several financial goods implying that these goods are inputs (positive user cost) in some periods and outputs (negative user cost) in other periods. For the whole sample, equity is an input and other financial goods are outputs at the sample mean.

According to the model, most of the elasticities of transformation are statistically and/or economically insignificant. This result is consistent with the findings in Hancock (1985) and Barnett and Hahm (1994). For example, all elasticities of transformation between equity and other financial goods are statistically and economically insignificant. However, we find several cases of highly elastic financial goods. For example, in the subsample of large banks, loans and leases are strong complements with deposits ($\sigma_{ld} = -3.386$) and strong substitutes with debt securities ($\sigma_{sl} = 7.900$). It is also worth noting that several elasticities of transformation have different signs for different subsamples of banks. For instance, debt securities and loans and leases are essentially weak (but statistically significant) complements for all banks and small banks, but they are strong substitutes for large banks. Similarly, deposits and other debt are unrelated for all banks, strong complements for small banks and strong substitutes for large banks.

7.3. Compensated Price Elasticities

We report the own- and cross-compensated price elasticities in Table 6 for each of the three bank groups. The elasticities are evaluated at the arithmetic sample mean, and numbers in parentheses are standard errors. According to the results, the financial technology is relatively inflexible, consistent with Hancock's (1985) findings.

As equation (8) suggests, the sign of the own-price elasticities of each financial good is indicative of whether the financial good is an input or output at the sample mean. For example, loans and leases are outputs for all banks and small banks but an input for large banks. At the same time, debt securities and trading accounts are output in the whole sample and both subsamples. The own-price elasticity of debt securities and trading accounts is approximately unity for large banks but significantly smaller in other subsamples.

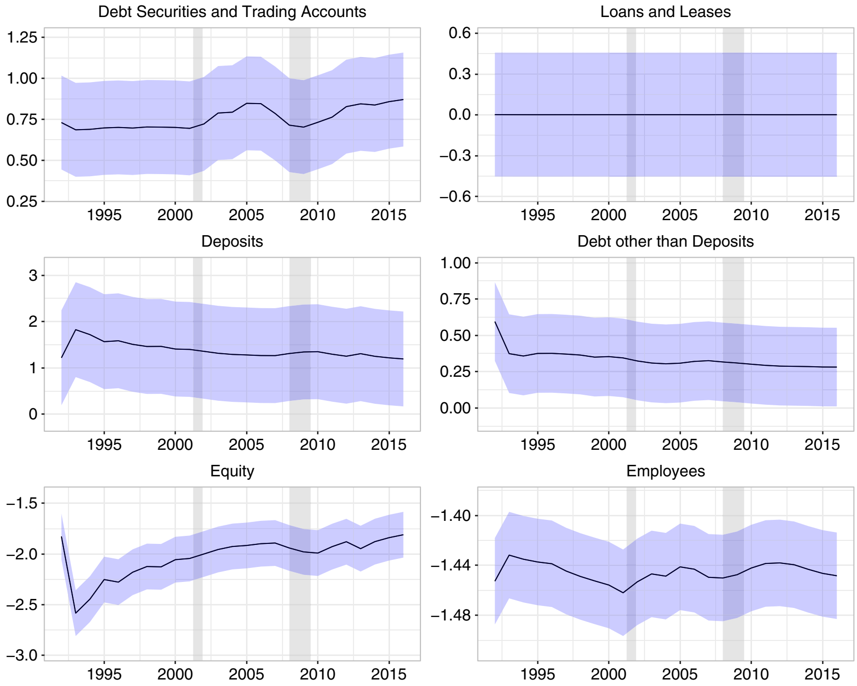
Most price elasticities are significantly less than unity. For loans and leases only in the subsample of large banks two elasticities are greater than unity. They are the own-price elasticity $\eta_{ll} = -1.500$ with a standard error less than 0.001 and the elasticity with respect to the user cost of debt securities $\eta_{ld} = 2.000$ with a standard error of less than 0.001. For the whole sample, all price elasticities of loans and leases are less than 0.1. A small magnitude of most price elasticities implies that the monetary transmission process is insensitive to interest rates. Therefore, the effectiveness of conventional monetary policy can be limited (the next section provides a more detailed analysis).

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

TABLE 6. Compensated price elasticities of supply and demand

% change in quantities	Asset class	% change in price					
		Debt securities	Loans and leases	Deposits	Other debt	Equity	Employees
Debt securities	all banks	0.736 (0.048)	−0.269 (0.008)	−0.249 (0.001)	−0.001 (0.000)	−0.007 (0.004)	−0.211 (0.037)
	< $100 million	1.151 (0.068)	−0.542 (0.025)	−0.018 (0.007)	−0.020 (0.022)	−0.000 (0.000)	-0.571 (0.059)
	> $1 billion	0.984 (0.000)	−0.492 (0.000)	0.000 (0.000)	−0.003 (0.000)	−0.000 (0.000)	−0.489 (0.000)
Loans	all banks	−0.001 (0.000)	0.001 (0.001)	−0.001 (0.000)	−0.000 (0.000)	−0.000 (0.000)	0.000 (0.000)
	< $100 million	−0.000 (0.000)	0.001 (0.001)	−0.000 (0.002)	−0.000 (0.000)	−0.001 (0.001)	0.000 (0.000)
	> $1 billion	2.000 (0.000)	−1.500 (0.000)	0.000 (0.000)	−0.003 (0.000)	0.000 (0.000)	−0.497 (0.000)
Deposits	all banks	−0.014 (0.009)	−0.017 (0.006)	1.383 (0.096)	−0.000 (0.000)	−1.335 (0.117)	−0.018 (0.005)
	< $100 million	−0.019 (0.009)	−0.293 (2.063)	0.677 (4.150)	−0.044 (0.008)	−0.028 (0.075)	−0.294 (2.030)
	> $1 billion	−0.015 (0.000)	0.028 (0.000)	−1.505 (0.000)	1.941 (0.000)	0.000 (0.000)	−0.448 (0.000)
Other debt	all banks	−0.008 (0.009)	−0.117 (0.361)	−0.042 (0.012)	0.336 (0.409)	−0.042 (0.307)	−0.127 (0.357)
	< $100 million	−0.088 (0.094)	−0.000 (0.001)	−0.186 (0.135)	0.284 (0.058)	−0.008 (0.011)	−0.002 (0.004)
	> $1 billion	−0.026 (0.000)	0.006 (0.000)	−0.007 (0.000)	0.034 (0.000)	−0.000 (0.000)	−0.007 (0.000)
Equity	all banks	0.001 (0.001)	0.001 (0.001)	2.028 (0.042)	0.000 (0.002)	−2.033 (0.037)	0.002 (0.001)
	< $100 million	0.000 (0.000)	0.588 (0.242)	0.025 (0.110)	0.002 (0.004)	−2.203 (1.966)	1.589 (2.093)
	> $1 billion	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	−0.500 (0.000)	0.500 (0.000)
Employees	all banks	0.627 (0.262)	−0.261 (0.234)	0.971 (0.111)	0.026 (0.052)	0.081 (0.014)	−1.444 (0.123)
	< $100 million	0.159 (0.241)	−0.064 (0.109)	0.076 (0.471)	0.000 (0.000)	0.466 (0.555)	−0.637 (0.217)
	> $1 billion	0.484 (0.000)	−0.121 (0.000)	−0.000 (0.000)	0.001 (0.000)	0.379 (0.000)	−0.742 (0.000)

Notes: The elasticities are based on the estimates of the generalized symmetric Barnett variable profit function. The sample includes quarterly data over the 1992–2017 period. The elasticities are reported at the sample mean. Standard errors are in parentheses.



Notes: The figure plots the dynamics of three possible average reference discount rates: (1) the cost of debt (other than deposits) capital, (2) the cost of deposits and other debt, and (3) the cost of deposits, other debt, and equity capital. The effective federal funds rate is presented as a reference.

FIGURE 2. Own-price elasticities of financial goods.

Another implication of our results concerns the monetary aggregation method. Traditional monetary aggregates are calculated as simple sum indices of the monetary components. Simple sum indices assume perfect substitutability between their components. Our findings suggest that monetary goods exhibit low substitutability and provide further evidence to Barnett's (2012) claim that the traditional (simple-sum) monetary aggregates are misleading.

In Figure 2, we show the own-price elasticities for all five financial goods and labor, calculated at the mean of each year's data. As can be seen, the own-price elasticity of loans and leases stays insignificant over the sample period. Other goods have statistically significant elasticities. Except for debt securities and trading accounts, the elasticities reveal no substantial fluctuations over the business cycle.

8. MONETARY AND REGULATORY POLICY ANALYSIS

Our estimates of the compensated price elasticities of banking technology are, in general, moderate or small in magnitude, a result consistent with Hancock (1985). However, since the prices of the financial goods are user costs, the interpretation

https://doi.org/10.1017/S1365100518000871 Published online by Cambridge University Press

of the elasticities requires taking into account the relation between holding costs and user costs. In some cases, small changes in holding costs may have significant effects on quantities even when the price elasticities are small. The reason for this is that the user costs are centered around the discount rates and the user costs are rates (percentages) by nature. In fact, if the holding cost of a financial good is close to the discount rate, then even small changes in the holding cost result in significant swings in the user cost.

8.1. Interest on Reserves

For years, the Federal Reserve did not pay interest on bank reserves. However, during the global financial crisis, legislation was passed and authorized the Fed to remunerate bank reserve holdings. Thus, since October 2008, the Federal Reserve operates a channel system of monetary control by paying interest on bank reserves. We analyze the effect of a 25-basis-point increase in the interest rate paid by the Fed on bank reserves on the demand and supply of financial goods, using compensated price elasticities.

From equations (3) and (4), it follows that the effect of a 1% increase in the interest rate on bank reserves on the user cost of deposits is $-g^d/(1 + R_{it})$, where g^d is the reserve requirement ratio. Assuming a reserve requirement ratio of 10% (the actual ratio depends on the amount of deposits and it is lower for small banks), a 25-basis-point increase in this rate results, on average, in a 5.2% increase in the volume of deposits and approximately a 0.9% decrease in debt securities and trading accounts. The effect on loans and leases is negative but less than 0.01%.

8.2. Reserve Requirements

Now we consider a decrease in the reserve ratio by 1%. According to equation (4), if the reserve ratio decreases by 1%, the user cost of deposits decreases by $\left(R_{it} - q^d\right)/(1 + R_{it})$, where q^d is the interest rate on required and excess reserves on deposits. Assuming a 0.5% interest rate on reserve balances (effective as of December 17, 2015), then a 1% decrease in the reserve ratio leads to approximately 3.5% increase in the amount of deposits, a 0.6% decrease in investments in debt securities, and an insignificant change in loans and leases.

8.3. Changes in the Federal Funds Rate

With the recent recovery of the U.S. economy, the Federal Reserve has moved its focus back to conventional monetary policy instruments. We consider the effects of an increase in the federal funds rate on the demand for and supply of financial goods in our model. We assume that the primary channel through which the federal funds rate affects the production of financial goods by banks is the interest rate on debt other than deposits (consistent with the discount window policy, interbank loans and money markets).

An increase in the federal funds rate has two opposing effects on the user cost of other debt. First, it increases the holding cost of other debt. Second, it raises the discount rate. However, since other debt is a relatively small component in the discount rate (the average weight of the cost of other debt in calculating the discount rate using the WACC method is about 5%), the magnitude of the second effect is economically insignificant. Therefore, the overall effect on the user cost of other debt is positive. We ignore the effect of the discount rate on the user costs of other financial goods.

Consider the effect of a 25-basis-point increase in the federal funds rate on the demand for and supply of financial goods. In line with our previous argument, a 25-basis-point increase in the federal funds rate results in a substantial decrease of 45% in the user cost of other debt. Given the compensated price elasticity of debt other than deposits, the expected percentage decrease in the quantity of other debt is approximately 15%. However, the effects on the other financial goods are negligible. This finding gains support from the recent evidence of large excess reserves accumulated by banks when the cost of capital has approached zero.

8.4. Changes in Investments Returns

It is also to be noted that in the aftermath of the global financial crisis, the Federal Reserve and many central banks around the world have departed from the traditional interest-rate targeting approach to monetary policy and have focused on their balance sheet instead, using nonconventional monetary policy, such as quantitative easing and credit easing. There has also been a move towards tougher standards in prudential regulation for banks, mostly in the form of higher regulatory capital requirements. Financial firm production can be influenced by nonconventional monetary policy, as well as by regulatory policy requirements, even when the policy rate is at the zero lower bound. The transmission mechanisms of such policies include traditional interest-rate channels that operate through the cost of borrowing and lending, other asset price channels, as well as the bank lending channel.

In general, the user-cost approach has no internal constraints related to the zero lower bound constraint on the policy rate and can be helpful in analyzing nonconventional monetary policy. Effectively, the unconventional monetary policy used by the Federal Reserve has also been reducing (long term) interest rates. For example, the Federal Reserve's quantitative easing program, which consisted of monthly purchases of government and mortgage bonds, raised the prices of these financial assets and lowered long term yields, while simultaneously increasing the monetary base. Here the changes in the interest rate affect banks mainly through the user costs of loans and leases and debt other than deposits. The analysis of the effects of these changes is similar to that for conventional monetary policy.

Since in our model banks optimally choose the quantities of financial goods, we cannot model the direct purchases of bank (problematic) assets. Instead, we can assume that asset purchases increase the return on bank investments ("debt securities and trading accounts" in our model) and, therefore, user revenues of these assets. A 25-basis-point increase in the return of bank investments leads to a growth of investments in debt securities and trading accounts by approximately 9.8%. The effects on other financial goods are less than 0.1% or statistically insignificant.

9. CONCLUSION

In this paper, we build on the path-breaking work by Hancock (1985) and Barnett and Hahm (1994) and develop an estimable model of the microeconomics of the financial firm, using recent state-of-the-art advances in microeconometrics. We assume that the deposit-taking financial firm produces intermediation services between lenders and borrowers and maximizes variable profit (total revenue less variable cost). We also follow the user-cost approach and define outputs as those assets or liabilities that contribute to a bank's revenue and inputs as those assets or liabilities that contribute to a bank's cost of production. With the calculation of user costs for financial goods, we use a flexible specification for the variable profit function in order to derive demands for and supplies of financial goods.

In constructing user costs, we deviate from Hancock (1985) and Barnett and Hahm (1994) and follow Diewert et al. (2016) using data on realized bank interest income and expenses in order to classify financial goods (debt securities and trading accounts, loans and leases, deposits, other debt, and equity) as inputs or outputs. We estimate the generalized symmetric Barnett variable profit function with the convexity conditions imposed globally in order to produce inference consistent with neoclassical microeconomic theory. We also take the literature to a new level by using quarterly panel data on all commercial banks in the United States, over the period from 1992 to 2016 (a total of 780,825 observations).

We find that the generalized symmetric Barnett model is able to provide inferences about the microeconomics of financial firm production consistent with neoclassical microeconomic theory, although the imposition of convexity globally does not always assure full theoretical regularity, as pointed out by Barnett (2002). The model produces well-behaved elasticities of transformation, and price elasticities of supply and demand. We show that the form can be estimated on a personal computer with a large sample even though the generalized Barnett form is not parsimonious as noted by Diewert (2015).

We find that most supplies of outputs and demands for inputs of financial goods are relatively inelastic, that the degree of substitutability/complementarity among financial goods is generally low and have low variability over time, and that production of financial firms is relatively insensitive to changes in user costs—consistent with Hancock (1985) and Barnett and Hahm (1994). These results are robust across different bank samples based on asset size. However, we find that small changes in holding costs may have significant effects on user costs and the demands for and supplies of financial goods, even when the user-cost elasticities are small in magnitude and nominal interest rates are close to zero. This has significant implications for monetary policy, as it suggests that the central bank can implement policy changes via changes in interest-rate policy instruments even when it operates near the zero lower bound constraint. While modelling a financial firm that operates at the zero lower bound is beyond the scope of this paper, one might embed the zero lower bound constraints into our model. This is an area for potentially productive future research.

Acknowledgments

We would like to thank William Barnett, Erwin Diewert, James Swofford, an anonymous Associate Editor, two anonymous referees, and participants at the Inaugural Conference of the Society for Economic Measurement at the University of Chicago, August 18–20, 2014, for comments that greatly improved the paper.

REFERENCES

Adrian, T. and H. S. Shin (2011) Financial intermediaries and monetary economics. In: B.M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics* (Vol. 3A), pp. 601-650. Amsterdam: North-Holland.

Barnett, W. A. (1976) Maximum likelihood and iterated Aitken estimation of nonlinear systems of equations. *Journal of the American Statistical Association* 71, 354–360.

Barnett, W. A. (1978) The user cost of money. *Economics Letters* 1, 145–149.

Barnett, W. A. (1983) New indices of money supply and the flexible Laurent demand system. *Journal of Business and Economic Statistics* 1, 7–23.

Barnett, W. A. (1987) The microeconomic theory of monetary aggregation. In W. A. Barnett and K. Singleton (eds.), *New Approaches to Monetary Economics*, pp. 115–168. Cambridge, UK: Cambridge University Press.

Barnett, W. A. (2002) Tastes and technology: Curvature is not sufficient for regularity. *Journal of Econometrics* 108, 199–202.

Barnett, W. A. (2012) *Getting it Wrong: How Faulty Monetary Statistics Undermine the Fed, the Financial System, and the Economy*. Cambridge, MA: The MIT Press.

Barnett, W. A. and J. H. Hahm (1994) Financial-firm production of monetary services: A generalized symmetric Barnett variable-profit-function approach. *Journal of Business and Economic Statistics* 12, 33–46.

Barnett, W. A., M. Kirova and M. Pasupathy (1995) Estimating policy-invariant deep parameters in the financial sector when risk and growth matter. *Journal of Money, Credit, and Banking* 27, 1402–1430.

Barnett, W. A. and G. Zhou (1994) Financial-firms' production and supply-side monetary aggregation under dynamic uncertainty. *Federal Reserve Bank of St. Louis Review* 76, 133–165.

Basu, S., R. Inklaar and C. Wang (2011) The value of risk: Measuring the service output of U.S. commercial banks. *Economic Inquiry* 49, 226–245.

Basu, S. and J. C. Wang (2013) Technological Progress, the "User Cost of Money," and the Real Output of Banks. Working paper: No. 13-21.

Berger, A. N. and D. B. Humphrey (1992) Measurement and efficiency issues in commercial banking. In: Z. Griliches (ed.), *Output Measurement in the Service Sectors*, pp. 245–300. Chicago: University of Chicago Press.

Berger, A. N. and L. J. Mester (2003) Explaining the dramatic changes in performance of US banks: Technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation* 12, 57–95.

Bolt, W. and D. Humphrey (2012) A Frontier Measure of U.S. Banking Competition. DNB Working Paper: No. 359.

Bolt, W. and D. Humphrey (2015) A frontier measure of U.S. banking competition. *European Journal of Operational Research* 246, 450–461.

Diewert, W. E. (2015) A Note on the Flexibility of the Barnett and Hahm Functional Form. Vancouver School of Economics, Working Paper: No. 2015-1.

Diewert, W. E., D. Fixler and K. Zieschang (2016) Problems with the measurement of banking services in a national accounting framework. In J. M. Hartwick (ed.), *National Accounting and Economic Growth*, Chapter 6, pp. 117–176. Cheltenham, UK: Edward Elgar Publishing.

Diewert, W. E. and T. J. Wales (1987) Flexible functional forms and global curvature conditions. *Econometrica* 55, 43–68.

Greene, W. H. (2012) *Econometric Analysis*, 7th ed. Upper Saddle River: Pearson Prentice Hall.

Hancock, D. (1985) The financial firm: Production with monetary and nonmonetary goods. *Journal of Political Economy* 93, 859–880.

Hancock, D. (1991) *The Theory of Production for the Financial Firm*. Boston, MA: Kluwer Academic.

Hughes, J. P. and L. J. Mester (1998) Bank capitalization and cost: Evidence of scale economies in risk management and signaling. *The Review of Economics and Statistics* 80, 314–325.

Lau, L. J. (1978) Testing and imposing monotonicity, convexity, and quasi-convexity constraints. In M. Fuss and D. McFadden (eds.), *Production Economics: A Dual Approach to Theory and Applications*, Vol. 1, pp. 409–453. Amsterdam: North-Holland.

Mas-Colell, A., M. D. Whinston and J. R. Green (1995) *Microeconomic Theory*. Oxford: Oxford University Press.

Olley, G. S. and A. Pakes (1996) The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64, 1263–1297.

Sealey Jr., C. W. and J. T. Lindley (1977) Inputs, outputs, and a theory of production and cost at depository financial institutions. *Journal of Finance* 32, 1251–1266.

Serletis, A. and S. Shahmoradi (2005) Semi-nonparametric estimates of the demand for money in the United States. *Macroeconomic Dynamics* 9, 542–559.

Serletis, A. and A. Shahmoradi (2007) Flexible functional forms, curvature conditions, and the demand for assets. *Macroeconomic Dynamics* 11, 455–486.

Wheelock, D. C. and P. W. Wilson (2018) The evolution of scale economies in U.S. banking. *Journal of Applied Econometrics* 33, 16–28.