Bankers and bank investors: Reconsidering the economies of scale in banking*

Ronald W. Anderson[†] and Karin Jõeveer[‡]

September 13, 2012

Abstract

We study economies of scale in banking by viewing banks as combinations of financial and human capital that create rents which accrue to investors and bankers. Applying this approach to annual data of US bank holding companies since 1990, we find much stronger evidence of economies of scale in returns to bankers as compared to returns to investors. The scale economies appear to be particularly strong in the top size decile of banks measured by total assets. We find that rents accruing to bankers are particularly strong in banks with a relatively large share of non-interest income and that for the largest banks a reduction of net interest margin is associated with an *increase* in bankers' rents. We find incorporating observable proxies for funding efficiency and presence in wholesale banking activities greatly reduces the pure size effect.

1 Introduction

Banking sectors have developed in very different ways around the world however, despite this diversity of origins the structure of modern banking sectors is remarkably similar among major countries. Typically there is a small number of very large banks and a large number of medium and small institutions. In addition to having an extensive presence in retail banking the largest banks tend to dominate investment banking, market making, and in the provision of a number of other wholesale services to other financial institutions. And the deep involvement of the largest banks in the market for government debt naturally creates strong links between big banks on the one hand and central banks and national treasuries on the other.

*We have benefitted from comments by Andrew Ellul, Charles Goodhart, Gary Gorton, Kevin James, Fred Sommers and participants at the LSE's FMG Lunchtime Workshop and Regulation Seminar,the FMG Bank of England Conference "Complements to Basel,", the Multinational Finance Conference, and the European Meetings of the Econometric Society. All responsibility for errors and views expressed is our own.

†London School of Economics, r.w.anderson@lse.ac.uk

‡Queens University Belfast, k.joeveer@qub.ac.uk

Despite its prevalence, the desirability of big banking has been called into question by the massive public sector support for the banking sector starting since 2008, much of it going to the largest banks. In this period, the first wave of public interventions in the banking sector involved numerous take-overs of insolvent banks by large healthy ones, which tended to increase banking concentration. Subsequently, the thrust of regulation of banking has tended to reduce the power of the largest banks by separating deposit taking from proprietary trading (Volker Rule) or investment banking more generally (the "ring fencing" proposed by the UK's Independent Commission on Banking). Furthermore, heightened regulatory capital charges mandated by the revised Basel Accord and the prospect for a supplementary charge for institutions deemed to be "systemically significant" has created an incentive for many of the largest banks to shed assets and to retreat from market segments where they find themselves in a competitive disadvantage.

Still, many commentators view these measures as too timid and would welcome an active use of anti-trust remedies to break-up the biggest banks (Reich, 2009). A new, but growing, vein of academic literature has developed that tends to support further actions aimed at the structural reform of banking. Philippon (2012) assembles a variety of time series on US economic activity dating back to 1870 and argues that financial intermediaries' share of GDP has grown rapidly since 1950. He argues that "improvements in information technology seem to have been cancelled out by increases in other financial activities whose social value is difficult to assess." Such activities include proprietary trading and involvement in over-the-counter (OTC) derivatives markets. On the theoretical front Bolton *et al* (2012) construct a model where finance can be directed through either public markets or OTC markets, where the opaqueness of the latter allows intermediaries to extract rents. They argue that through this mechanism the financial sector can grow excessively large.

OTC derivatives, securities lending, and other wholesale market activities tend to be dominated by relatively few players. This may suggest that there may be extensive economies of scale in these types of banking activities. And the fact that smaller banks have little or no presence in these activities further suggest that there may be some economies of scope that operate across a range of such activities. The trading of credit default swaps is an example of what we have in mind. While CDS are used by a broad range of institutions, market making is concentrated in a handful of global institutions who among themselves account for a large fraction of total trading activity. Nevertheless, most of the literature on banking efficiency has failed to find any evidence of economies of scale in banking that could rationalize large fraction of banking services that are provided by the biggest banks. Berger *et al* (1993) summarize the literature and conclude that "...the average cost curve has a relatively flat U-shape, with medium sized firms being slightly more scale efficient than either very large or very small firms." They suggest that the minimum efficient scale was something less than \$300 million total assets. By way of comparison, a total assets of \$300 million was approximately the 6th percentile by size of bank holding companies in our data set of banks in 1990.

More recent contributions have produced some evidence of stronger scale economies than did the earlier literature. Examples are the studies of Hughes and Mester (2011) and Wheelock and Wilson (2012) which are based on more recent data and new methodologies. However, these methodologies are limited by either using very special parametric functional forms or by applying to a limited time periods. So the generality of the results might be questioned. For example, Hughes and Mester study one year only, 2007, and find strong scale economies in only in the most restrictive of the several functional specifications they consider. Wheelock and Wilson use the static cost function framework employed by earlier banking efficiency studies but find significant scale economies only after making an important modification in the empirical specification, namely, by normalizing cost by dividing by the estimated bankers' wage rate. We discuss the significance of this later in light of our own results. Some banking analysts have become doubtful that the traditional static efficiency approach is capable of capturing the advantages of large organizations that seems to be implied by the observed equilibrium distribution of bank sizes (De Young, 2010). In particular, small and large banks tend to offer very different ranges of services with the latter dominating wholesale banking services, as has already been mentioned.

The failure of earlier studies to find evidence of substantial scale economies which can account for the observed size of the largest banking institutions may due to two key assumptions built into their methodology. First, previous estimates have concentrated on cost efficiency in the provision of retail banking services rather than wholesale banking services. Typically, output of the bank is measured by such quantities as total loans by class (commercial, mortgage, consumer...). This take no account of monitoring and screening effectiveness that affects the profitability of these activities (i.e., there is no adjustment for quality of outputs). Second, the methodology assumes that inputs including labor are purchased on competitive markets. For example, the bankers wage rate is typically calculated as the ratio of total employee compensation divided by total employees. This may miss rents that are extracted by providers of inputs who possess bargaining power with the bank. One of the principal aims of our study is to relax these two assumptions.

If there are substantial economies of scale in certain areas of wholesale banking we might expect that this would endow the large banks with a significant degree of market power. However, even if economies of scale give rise to substantial rents, it does not follow that this will translate into highly profitable banks. If wholesale banking activities rely on special expertise or on strong client relationships, it may be that a large fraction of the benefits accrue to key bankers rather than to the banks that employ them. Indeed, it is often reported that experienced, successful bankers will move as a team from one bank to another one that seeks to build market presence and is willing to pay guaranteed bonuses or other inducements to attract the needed talent. Thus modern wholesale banking is an example of a knowledge based industry where substantial bargaining power is in the hands of managers. This has been described by Rajan and Zingales in their essay "The Governance of the New Enterprise". In their view the new enterprise is distinguished by a reduced importance of vertical integration and a shifting of power away from the headquarters. In their words,

> *But perhaps the most significant change has been to human capital. Recent changes in the nature of organizations, the extent and requirements of markets, and the availability of financing have made specialized human capital*

much more important, and also much more mobile. But human capital is inalienable, and power over it has to be obtained through mechanisms other than ownership. As the importance of human capital has grown, power has moved away from the top and is much more widely dispersed through the firm.

In this paper we postulate that banking activities create value through the combination of financial capital provided by investors, principally shareholders, and human capital provided by bankers. Put simply, the technology of banking is to combine money, brains, and effort to produce money. Depending upon the markets where the bank operates and the organization of the bank, the value created by the bank will accrue in varying degrees to shareholders and bankers. How the bank trades off between these two will reflect the relative bargaining power of bank shareholders (the "principal") and bankers (the "agent"). The "efficiency frontier" describes the maximum benefit to bankers for a given benefit to investors. Using a variety of returns to bankers' human capital and investors financial capital and using a variety of statistical techniques we explore the evidence of economies of scale and ask whether scale economies are attributable to banks' involvement in particular wholesale banking activities.

The remainder of the paper is organized as follows. In Section 2 we set out a simple model that clarifies our approach and describes how efficiency differences can be analyzed. Section 3 present preliminary evidence on increasing returns to investors, bankers and banks overall. Section 4 explores what observable bank characteristics can account for observed differences in returns across scale. Section 5 refines this analysis by drawing upon detailed wholesale banking proxies derived from regulatory filings. Section 6 discusses the interpretation of our results and relation to the literature. Finally conclusions are given in Section 7

2 The model

As suggested in the introduction, we dispense with the assumption of a competitive market for bankers' services and instead model a bank using the principal/agent paradigm. We suppose that the total value (rent) produced by a bank is shared between shareholders (the principal) and bankers (the agent). The bargaining power of bankers derives from the fact that their actions can be only imperfectly monitored by shareholders. Following the literature on optimal contracting in face of moral hazard we suppose that the second best efficient allocation is described by the maximum payoff to the principal for any given level of payoff to the agent and that this relationship is decreasing and concave (see DeMarzo and Fishman, 2007). Letting b be the payoff to bankers and r be the payoff to shareholders, then the bank efficiency frontier is described by a relationship

$$b = f(r; x)$$

where x is a vector of control variables reflecting the bank scale, prices of other inputs, etc. We suppose $f'(r) < 0$ and $f''(r) < 0$ for all r.

In Figure 1 we depict two efficiency frontiers in the case where x is a scale variable taking on two values 'large' and 'small'. We measure the efficiency of the bank as the euclidian distance $e = (r^2 + b^2)^{1/2}$. Suppose a 'small bank' has $r = \$2.5$ and $b = \$2$ as depicted by point A in the figure. This bank is inefficient as seen by the fact that this point lies to the southwest of the efficiency frontier for small banks. Its efficiency is $e = (2.5^2 + 2^2)^{1/2} = \$3.2016$. Its relative efficiency is measured by comparing this to the point on small bank efficiency frontier intersected by the ray from the origin passing through point A. This is the point (3.1235, 2.4988) whose efficiency is \$4. Thus the relative efficiency of bank A is 3.2016/4 or 80.04%.

To compare efficiency of two different classes of banks we can calculate for each class the average efficiency (i.e., distance to the origin) of point along its efficiency frontier and then calculate the ratio of these two average efficiencies. For example, in Figure 1 the small bank efficiency frontier has been constructed as the quarter circle with radius \$4 and the large bank efficiency frontier is the quarter circle with radius \$5. Thus the average efficiency of small banks is \$4 and of large banks is \$5. So large banks have an efficiency 125% of small banks.

From Figure 1 it is also clear that omitting returns to bankers from an analysis of returns to scale can lead to erroneous conclusions. For example suppose that most observations for "small banks" are clustered close to the efficiency frontier in the neighborhood a payoff of 3.8 for shareholders. At the same time observations for "big banks" are clustered near the large bank frontier with payments of 3.1235 to shareholders. Then a regression of shareholders returns on size would find a negative relationship; whereas, the analysis combining shareholders and bankers returns reveals increasing returns to scale.

3 Do returns to bank investors and bankers vary systematically with scale?

In this section we consider whether bank returns vary systematically with the scale of the bank. Our data set covers bank holding companies that are regulated in the US, covering 1990-2010. The data include all balance sheet and income statement variables reported in Compustat Bank Annual Fundamental File.

As discussed in the introduction we distinguish returns to bank investors and to bankers. We start by studying these measures separately. We then combine the two obtain measures of total bank rent and consider evidence of economies of scale to this aggregate measure.

3.1 Bank investors

We take a bank's investors to be its shareholders because collectively they have active control rights in a going-concern bank and they are ultimately responsible for the bank's compensation policy. We represent shareholders return by return on equity ($niseq$), calculated as annual net income after tax (ni) divided by book equity (seq), i.e., $niseq = ni/seq$. As a check on the robustness of our results, in the Appendix we have also considered return on assets, and we find the same qualitative conclusions as in this subsection and in the regression analyses of *niseq* from the next section.

Table 1 presents descriptive statistics of the return on equity for US listed bank holding companies from 1994 through 2010 by size decile where size is represented by total assets (*at*). Comparing these indicators across size deciles provides some evidence that returns to investors tend to increase with bank size. This pattern is quite systematic for the median and also for the 25th and 75th percentiles. There is a wide dispersion of returns within an given size class. However, it does appear that the whole distribution of returns is translated to the right as size increases. Judging from the quartile estimates, we might guess that the relation between return on equity and bank size is approximately linear.

Figure 2 depicts this information graphically. We have plotted the cumulative sample distribution of returns to bank investors for the fifth to tenth size deciles of banks. We see that the distribution for the largest banks lies strictly above those of the smaller banks, i.e., the tenth size decile dominates the others in the sense of first-order stochastic dominance. While the ordering is not strict for the smaller size deciles because some of the curves cross at the left tail of the return distributions, the size ordering does hold in the middle and in the right tail of the return distributions. That is, high investor returns are more frequent for banks in the ninth size decile as compared to the eighth size decile etc.

Some of the large dispersion seen in Table 1 is the result of pooling our 17 years of data. We have also broken down median returns by size and by year (not reported). When we do so we find that returns to investors are typically non-decreasing by size and then strongly increasing between the 9th and the 10th size deciles. The exception to this pattern occurred in the crisis years 2008 and 2009. The fact that the largest banks had a sharper downturn in profitability during the crisis suggests that the largest were pursuing different (perhaps more risky) strategies or were in different market segments than smaller banks.

To summarize, based on a preliminary review of return on equity by bank size, we see there is some evidence of economies of scale in banking. However, there is considerable dispersion of return on equity within any size class, which suggests that there may a variety of other factors beside scale that are important determinants of investor returns in banking. In addition, there is some evidence that the very largest banks may be involved in different business segments and may be exposed to different risks than small and medium sized banks.

3.2 Bankers

While measuring returns to bank investors is relatively straight forward, calculating returns to bankers' human capital is anything but straight forward. Many studies of banker compensation focus on pay of CEO's and other top management in the firm. For several reasons, this is not the approach we take here. The most important reason is that detail compensation information is only reported for top managers. It therefore will not capture compensation to traders and others without top management status but who may have substantial performance related pay. Second, compensation reporting standards for these top managers have varied over time. Third, compensation packages of bank employees will include salary, cash bonus, stock awards (current and deferred), stock options (current and deferred), pensions contributions, plus a variety of perquisites. Anecdotal evidence suggests that compensation practices vary greatly across banks.

In light of these considerations we have taken the following approach to estimating returns to bankers. We estimate aggregate rents earned by bankers in a given bank as the total amount employee compensation in excess of what we estimate to be the "competitive wage bill" of the bank. Specifically we calculate staff costs per employee ($xlremp$) as the ratio of total staff costs (xlr) to total number of staff (emp), i.e., $xlremp = lxr/emp$. We calculate for each year the average annual compensation per employee for banks with at least 50 employees and total assets less than \$1 billion. This is our estimate of upper bound on the competitive wage of a banker in year t, w_t. Using this our estimate of total rents to bankers is calculated as the total compensation to employees minus total estimated compensation as the minimum of the bank's wage and the maximimum competitive wage, i.e., $xlrrent_{it} = max(xlr_{it} - w_t * emp_{it}, 0)$. By setting the size cut-off at \$1 billion we are including banks up to about the 65'th size percentile within our sample. As a robustness check we have used an alternative bankers' rent calculation which uses banks with at least 50 employees and total assets less than \$2 billion, i.e., up to about the 75'th size percentile. The qualitative results with this alternative rent measure are essentially the same as in this subsection and as in the regression results reported in Section 4. These results are presented in the Appendix.

Table 2 reports summary statistics of total rent per banker broken down by bank size deciles. A variety of interesting facts emerge from these statistics. First, the first quartile systematically equals 0 for each size decile. That is, many banks, independently of size, pay no significant rents to their employees. Second, comparing the medians across size deciles we see that only relatively large banks tend to pay out rents to their employees. Finally, and this is most important for our purposes, the biggest banks payout far more rents to employees than all other banks. For example, the median banker rent in the 10th decile is an order of magnitude greater than that of the 9th decile (\$5946 per employee versus \$424 per employee). Comparing firms at the third quartile, the difference is \$18,664 versus \$10,780. Stated otherwise, judging from bankers' rent payments there is evidence of significant economies of scale that are strongest in the region of the largest bank size category.

Table 2 pools 17 years of data, and again it is useful to see whether these observations hold year by year. When we calculate median bankers' rent per employee by bank size and by year we find that the pattern of rent payment across different size categories has changed considerably over time. Since 1999 when reporting of head count and total compensation has been more regular than in prior periods, we see a distinct pattern of the largest banks paying the most rents to employees. For example, in 2006 just before the crisis, the median rent payout to employees of banks in the largest size decile was \$14,290 as compared to \$2632, \$2125 and \$1927 in the ninth, eighth, and seventh deciles respectively. Furthermore, this pattern continued during the crisis and after, from 2007 through 2010. Thus the poor performance of the largest banks in the crisis was felt by investors rather than bankers. Confining our attention to the data since 1999, we do find evidence of significant scale economies in the rents paid to bankers. In a two-way ANOVA of *mxlrrentemp* with year and size effects, the F-statistic of the hypothesis of equal means across size deciles is 2.11 which is significant at the 0.025 level.

In order to compare these estimates of bankers' rents with payout to bank investors we normalize by bankers' rents by the book value of common equity (mxlrrentseq=mxlrrent/seq). Table 3 reports summary statistics on our estimates of returns to bankers normalized by the book value of equity. Again we find evidence of economies of scale in the payoffs to bankers. Rents to bankers are more common in the larger banks and the returns to bankers in the largest size decile are significantly larger than in smaller deciles.

Figure 3 depicts the cumulative sample distributions of bankers' returns for banks segmented by size deciles. Similarly to Figure 2 the distributions tend to be ordered by size with the higher returns being associated with larger banks. However, the dominance of the largest 10% of the banks is even more dramatic when judged by returns to bankers. Thus even though there is a fair amount of dispersion in bankers' rents within this largest size class of banks, there is a clear tendency for the biggest rents to be paid out to bankers in the largest banks.

Clearly an important step in the construction of our measure of bankers' rents is the determination of the competitive bankers' wage. In our benchmark measure we treat this as common across all banks. However, it might be argued that different sized banks draw upon different labor markets. Accordingly, as a check on the robustness of our result we have calculated an alternative rent measure where we first calculate the average bankers' wage within each size decile. Then we calculate the bankers' rents in a particular bank as the amount of total compensation in excess of the competitive wage bill using the competitive wage rate for the relevant decile. This measure divided by total book equity we call $mxlrrentseq3$. We have plotted the sample distributions of these measure by size decile in Figure A in the appendix. This displays the same qualitative pattern as for our benchmark measure of bankers' rents. The distribution of rents in the top size decile dominates those in lower deciles.

Combining results from Tables 1 and 3 we arrive a first important insight from the approach that incorporates bankers' return as well as investor return into the analysis: the evidence of increasing returns to scale are significantly greater when bankers' return is included than when it is not. For example comparing medians in the 10th decile with those in the 9th decile, we see that total returns (investor plus banker) are higher by 3.56 per cent of total equity whereas investor returns are higher by only 2.1 per cent. As we will see this is a robust finding that continues to hold even after we incorporate a number of observable variables to control for sources of variation in returns.

For the purposes of better understanding determinants of scale economies it seems useful to use a measure of total bank returns as well as returns to investors and bankers separately. The simplest approach is to add the two returns as we have done in the numerical comparison just made. When calculate this measure and then calculate sample statistics broken down by size decile, we find the same pattern seen in Tables 1 and 3 above. There is evidence of economies of scale out to the largest size decile.

We have also calculated an alternative measure of total rents that exhibits the same concave shape as seen in Figure 1. Economically this captures a possible decreasing marginal returns to bank stake holders. That is, for banks within a given efficiency class, increasing returns returns to bankers to higher and higher levels will require greater sacrifices of investor returns as banks move into lines of business that are favorable to bankers' bargaining power. The measure that we use is

$$trentseq = (b^2 + r^2)^{\frac{1}{2}}$$

where now $b = 1 + niseq$ and $r = 1 + mxlrrentseq$. Again this measure exhibits economies of scale along the lines expected from Tables 1 and 3.

4 What can account for economies to scale in banking?

Until now we have (a) presented evidence of economies of scale both in returns to bank investors and to bankers and (b) shown that returns to bankers are increasing particularly in the largest banks. Furthermore, even though scale economies are present, we have found considerable variation of returns to bank investors and bankers that is not explained by size or year effects alone. In this section we explore other variables using regression analysis to see if they can account for this unexplained variation and possibly also account for the scale effects we have identified.

We consider models of the form,

$$return_{it} = \alpha_t + \beta X_{it} + \epsilon_{it}$$

where $return$ is a measure of bank returns, X is a vector of explanatory variables, i is the index of the bank, t is the fiscal year, and ϵ is an i.i.d. error term. As measures of the dependent variable $return$ we use return on equity, $niseq$, banker rents relative to equity, $mxlrrentseq$, and the total bank return measure based on our constant elasticity of transformation function, $trentseq$. The intercept α_t captures time variation affecting all banks.

The basic explanatory variable to capture returns to scale is total assets at. To allow for a nonlinear relationship affecting the largest banks, we use a dummy variable, $at10$, to indicate whether or not the observation pertains to a bank in the 10th size decile.

Other explanatory variables are chosen as indicators of funding efficiency, presence in wholesale banking services, and leverage. To capture the effect of funding costs on return we include net interest margin, nim. As the largest banks may have access to alternative sources of funding through wholesale markets including derivatives, we should recognize that this may impact funding costs. We allow for this by $nimat10$, a variable interacting net interest margin nim and the dummy for the largest size decile $at10$.

Our primary indicator of bank presence in market making and the provision of wholesale banking services is $niish$, the share of non-interest income in total bank revenues (non-interest income plus net interest income). To allow for the possible nonlinear effects for very large banks we use $niishat10 = niish * at10$. The share of non-interest income is a relatively crude proxy for wholesale banking activities, and ideally we would like to have detailed indicators of the presence in a variety of different wholesale markets. In the next section, we will use information in detailed regulatory filings available for some of the banks in our sample in order to explore this issue further.

We control for leverage with the capital ratio variable $ilev2$ calculated as the ratio of book equity to total assets. In addition, we control for general economic conditions with year dummies. Table 4 presents the summary statistics of our main explanatory variables. There is considerable variation in these variables. The low pair-wise correlations suggests that they are capturing very different aspects of bank characteristics.

Our main regression results are reported in Table 5 based on annual data between 1999 and 2010. The equations for $niseq$ and $trentseq$ are estimated by OLS. To take into account left-censoring of our dependent variable we report Tobit regressions for $mxlrrentseq$. Given that it is implausible to expect that the large number of very small US banks banks would have significant presence in wholesale markets we confine our regressions to the size deciles 6 through 10. Also outliers are removed (if $trentseq > 2$, $niish < -2$ or $niish > 2$)). T-tests are reported below coefficient estimates.

The first three columns of Table 5 report results based on the scale variables, at and $at10$ as well as yearly effects. The results are in line with our discussion in Section 3 where we found some evidence of scale economies in shareholders' return, but very strong evidence of scale economies in bankers' return. In the shareholders' return regression the scale variable enters positively but is statistically insignificant. However, the dummy for the 10th size decile, $at10$ is positive and highly significant. Both at and $at10$ are positive and highly significant in the regression for bankers' return, $mxlrrentseq$. When the two return measures are combined in $trentseq$ we also find strong evidence of positive scale effects that are increasing in the 10th size decile.

We now consider whether these strong scale effects persist when we introduce our additional explanatory variables into our regressions. In columns 4-6 of Table 5 we include our measures of funding efficiency, nim, and presence in wholesale banking, $niish$ as well as these variables interacted with the 10th size decile, $nimat10$ and $niishat10$. We see that these additional variables largely account for the pure scale effects. Total assets at is insignificant in all three regressions. And the 10th size decile dummy is now insignificant in the bankers' rent regression, $mxlrrentseq$.

In column 4 of Table 5 we see the coefficient on nim is positive and significant. That is, lower funding costs relative to returns on lending are passed on to investors in the form of higher return on equity. The coefficient on the interaction term $nimat10$ is negative and significant. That is, for the largest banks, low relative funding cost is a less important determinant of returns to equity. One interpretation of this is that large banks may be able increase profitability to investors by expanding activity (either on balance sheet or off balance sheet) even if this means higher funding costs on wholesale markets because they may have higher effective leverage or may benefit from scope economies (in the business called "cross-sales") to generate more fee business. The coefficient on $niish$ is positive and significant, implying that greater presence on wholesale markets is associated with higher returns on equity. However, this seems to apply to equally to banks of all sizes, because the inter-action term $niishat10$ is positive but insignificant. Thus the estimates in column 4 give us a first, preliminary answer to the question “what are the returns to scale in wholesale banking”. The answer is that we find positive but weak scale economies to wholesale banking when the performance metric is return on equity.

When we consider the role of the same explanatory variables in the determination of returns to bankers, Column 5 of Table 5, we find the coefficient of nim is positive and significant. That is, some of the advantage of greater funding efficiency accrues to bankers as well as well as to bank investors. However, in contrast with the results obtained for bank investors, the coefficient of $nimat10$ is negative and and highly significant. It is worth pausing to draw out the implications of this last result. If we combine the coefficients of nim and $nimat10$, we obtain a negative value that is statistically significant. That is, we find that a very large bank with high apparent funding efficiency pays out relatively little rent to its bankers. In contrast, a very large bank with low apparent funding efficiency pays out substantial rents to its bankers. Referring to the discussion in the introduction concerning bankers’ bargaining power, one explanation of this finding is that large banks may pursue growth by greater wholesale funding with narrower intermediation margins which may require bankers with specialist knowledge that would be difficult to replace and thus enabling such bankers to extract very attractive compensation packages.

Our wholesale banking activity proxies, $niish$ and $niishat10$ are both positive and highly significant in column 5 of Table 5. That is, greater presence in wholesale banking markets is associated with higher returns to bankers independently of scale and funding costs. This effect is even stronger for larger banks—the interaction term (niishat10) is positive and significant.

To summarize, we have found that rents accruing to bankers are particularly strong in banks with a relatively large share of non-interest income, i.e., from market making, investment banking, advisory, and trusteeship. Once we take this wholesale banking proxy into account, a pure scale effect disappears. The fact that we find a reduction of net interest margin is associated with an *increase* in bankers’ rents may suggest that increasing scale through greater reliance on wholesale funding may increase rents but these accrue disproportionately to bankers rather than bank investors.

In column 6 of Table 5 where we combine returns to bankers and bank investors in the total rent measure, $trentseq$, we find some evidence of a pure scale effect operating for the 10th size decile. The funding efficiency variable, nim enters positively and is highly significant; however, the funding effect is diminished somewhat in the top size decile. The wholesale market presence measure, $niish$ enters positively and is highly significant, and this effect is reinforced in the 10th size decile. We have also used the alterative specification of the sum of returns to bankers and bank investors and obtain very similar qualitative results.

Columns 7 to 9 of Table 5 introduces the capital ratio, $ilev2$, as well as controls for scale, funding efficiency and presence in wholesale markets. $ilev2$ is very close to the inverse of the leverage ratio used by US bank regulators and may be viewed as a control for the level of risk taken on by the bank. As emphasized by Hughes and Mester (2011) it may be that banks of different sizes also differ in their product mixes and that increases in return are obtained only by engaging in riskier activities. This is not borne out by the results in column 7 where $ilev2$ enters with a positive coefficient and is statistically significant. That is, controlling for other factors, greater leverage is associated with *lower* return on equity. In the bankers' return and total return regression (columns 8 and 9) $ilev2$ enters with the expected negative sign and is statistically significant. However, adding the control for leverage leaves the qualitative results as had been found without leverage in columns 4-6. Funding efficiency has a positive effect on returns to shareholders, bankers and the two combined, but this effect is diminished in the top size decile. Presence in wholesale banking enters with positive sign in all three regressions and this effect is increased in the largest size decile. Thus our measures of funding efficiency and investment banking activities are not simply capturing a leverage effect. When funding efficiency, wholesale banking and leverage are all included, the pure scale effects no longer appear to operate except in the top size decile for return on equity and combined shareholder/banker return.

In order to assess the robustness of our findings we have experimented with a wide variety of alternative specifications. One robustness check is to ask whether there is some omitted source of heterogeneity which can be captured as a pure firm effect. In fact, because of the very high level of structural change in U.S. banking in the last two decades as reflected in the large number of bank mergers and bank failures where large portfolios of liquidated assets were acquired by surviving banks, we do not view our data set as a ideal panel. Nevertheless, it is possible to formally treat it as a panel using the Compustat variable gvkey as the firm identifier. In Table 6 columns 1-3 we allow for firm random effects in addition to the controls we use so far. We find that introducing pure firm effects has very little impact on the results. Funding efficiency, nim enters with positive sign in all three regressions, but when interacted with the top size decile dummy the effect is negative and significant. The wholesale market proxy, $niish$, is positive and significant and the effect is reinforced in the top size decile. The capital ratio, $ilev2$ enters positively in the equation for return on equity and negatively in the bankers' rent and total return equations. These panel estimations do produce somewhat different estimates of the pure scale effects. The top size decile dummy $at10$ enters positively and is significant in the regressions for return on equity and bankers' rent. Finally, at enters negatively in the model for bankers' rents.

As was already noted, the non-interest income share, $niish$, is a fairly crude proxy for the bank's involvement in wholesale banking activity. For this reason, in the next section we will explore a variety of alternative indicators that are available for a subset of our banks that are required to make detailed regulatory filings. An alternative approach is to use an instrumental variables method as we have done in columns 4-6 of Table 6 where we report the second stage estimates from an IV regression where we instrument for both $niish$ and $niishat10$. The instruments are the observed open positions in derivatives, securities lending and in repurchase agreements as detailed in the next section. Compared to the results of columns 7-9 of Table 5 the main change is that interaction term $niishat10$ is now no longer significant. The other main qualitative findings are very similar to our previous results. Funding efficiency, nim enters positively in all three regressions, but the interaction term, $nimat10$ enters negatively. The presence in wholesale banking $niish$ is positive and highly significant.

The capital ratio, $ilev2$, enters positively in the shareholders' return regression and negatively in the bankers' return and total return regressions.

In addition to these robustness checks we have explored a number of other alternative specifications. The striking feature that comes out of this exercise is that the main qualitative conclusions that have been found in this section are very robust. Funding efficiency, nim has a positive effect on returns to shareholders, to bankers and to the two combined. However, this affect is diminished among banks in the top size decile. Presence in wholesale banking, $niish$, has a positive effect on returns particularly for bankers and then especially in the top size decile. Leverage is positively associated with bankers' returns but negatively associated with shareholder returns. Some of these additional tests are reported in the appendix. Table I reports the results for our benchmark regressions using return on assets rather than return on equity as our performance measure for bank investors. Table II modifies the specification from all three benchmark regressions to proxy size by the logarithm of total assets ($lnat$) instead of total assets. Table III uses two alternative measures of bankers' rents. Here $mxlrrentseq2$ (column 2) and $trentseq2$ (column 5) pertain to the total wage bill in excess of the competitive benchmark where the latter is based on banks with total assets of up to \$2 billion versus \$1 billion used in our other results. In column 3 of Table III we employ $mxlrrentseq3$, the measure of bankers' rents where the competitive wage is allowed to vary across banks in different size deciles as in Section 3.2 and in Figure A. All these alternative versions produce results remarkably similar to those we have discussed in detail in this section.

To summarize, we have found evidence of positive economies of scale that are stronger when bankers' returns are taken into account than when they are not. However, differences in product mix, notably the presence in wholesale banking activities, seem to capture an important part of return differences between larger and smaller banks. We have emphasized that modern banking products are diverse and this is particularly true of what may be considered wholesale banking activities, i.e., services and products that cater particularly to institutional clients. These include securities trading and lending, OTC market making, trusteeship activities, securities underwriting and other investment banking activities. The importance of wholesale banking in accounting for increasing returns to scale is apparent only when returns to bankers are taken into account either as a performance measure by itself or in combination with returns to shareholders. Our results also suggest that the use of wholesale funding to expand the scale of the banks operations can increase bank returns even if it squeezes the bank's average net interest margin. However, these increased returns accrue largely to bankers rather than bank shareholders.

5 Alternative measures of wholesale bank activity

In order to provide a more detailed picture of wholesale banking activities we have combined our annual data set of bank holding companies with data contained in detailed regulatory filings by firms directly supervised by the Federal Reserve. These two data sets, the former derived from Compustat and the latter contained in the FRY9c filings, differ in several respects. The Compustat data cover some bank holding companies not included in the FRY9c data because they were not regulated by the Federal Reserve in the period covered. This is the case for example of financial firms who had their origins as Savings and Loans Institutions, with Washington Mutual being a prominent and very large example. The FRY9c data include some banking holding companies excluded from our Compustat based data. Merging the two data sources is further complicated by the fact that Compustat and the Federal Reserve do not share a common firm identifying variable. This last problem is partially solved by using a cross listing prepared by the research department of the New York Fed of Federal Reserve Identifiers (RSSD9001) and CRSP identifiers (PERMCO) for listed bank holding companies operating in 2008. Then linking these through the WRDS merged Compustat/CRSP data set we were able to obtain a matching of Fed identifiers and Compustat identifiers for many of the firms in our sample. However, this strategy does not work for banks that disappeared from our sample prior to 2008 through mergers, failures or otherwise. Ignoring this would have introduced a possibly significant sample selection bias in the analysis. To deal with this problem we have done a detailed search for firms with end dates in the Federal Reserve sample prior to 2008. This involved string matching of words contained in the Federal Reserve's Legal Name (RSSD9017) with the company names (conm) in the Compustat data set and then verifying that matched data set was correct, e.g., by comparing total assets in a given period. The result is a matched set of Compustat identifiers (gvkey) and Fed identifiers (RSSD9001) covering the period 2000 to 2010.

Table 7 lists summary statistics for selected wholesale banks measured as stock outstanding at year end, all normalized by total assets. These indicators are reported by almost all the banks in the top size decile. We give sample characteristics for these large banks between 2000 an 2010. There are many missing values among smaller banks, which might be indicative of zero positions. However, we have excluded these missing observations from the sample.

REPO is the amount of repurchase agreements outstanding at year end net of reverse repurchase agreements. This is reported on banks balance sheets and is a measure of reliance on wholesale funding. Tier 1 capital ratio (ilev2, above) is also based on the bank holding companies' consolidated balance sheets. The remaining variables are off-balance sheet, year-end stock values relative to total assets obtained from FRY9c filings. SEC LENDING is the gross amount of securities lending. FUTURES are gross nominal value of all exchange traded futures outstanding. FORWARDS are gross nominal amounts of OTC forward contracts outstanding. OPTIONS are gross nominal value of exchange traded options contracts written. OTC OPTIONS are gross amount of over the counter options contract written. All these derivatives positions include contracts on interest rates, equities, foreign exchange, commodities and other derivatives.

By comparing mean, minimum and maximum values we see there is a wide range of involvement in these various wholesale market activities. This reflects a significant degree of specialization among banks. For example, among the largest 10% of banks, the mean of securities lent by banks is 12.3% of total assets. However, many do no securities lending; while a handful of specialized banks do large amounts with the largest amount observed in our sample being 4.731 times the total balance sheet of the bank (reflecting a very high involvement in lending and rehypothecation). Similarly, quite a few of the relatively large banks eschew derivatives trading altogether; whereas some are very active derivatives houses whose total nominal exposures are many multiples of their balance sheets. In this regard we have chosen to work with gross notional exposures, since we think this is a better indicator of market making activity than would be net exposures which might be more indicative of economic exposure in a derivatives book. Also, it will be noticed that notional exposures tend to be larger for OTC derivatives than for exchange traded derivatives reflecting the centralized clearing of the latter. Another indication of the specialization within wholesale banking is provided by the correlations reported in Table 8 which are not high in many cases. Finally from this same table we see that the correlations between the capital ratio and the wholesale market indicators are negative but fairly low suggesting they capture quite different information about bank strategy. Nevertheless, there is some weak positive association between use of leverage and presence in wholesale markets.

We now consider whether these additional measures of wholesale market activity have an effect on returns to investors and bankers and whether these might be better proxies than the share of non-interest income variable (niish) that we have used until now. We do this first by rerunning regressions as in columns 7-9 Table 5 but with the new measure introduced both in levels and interacted with the top decile dummy, at10. Table 9 reports the results with REPO and REPOAT10. REPO is positive and significant in the investor return and total return regressions. That is, there is some evidence that increased reliance on wholesale market funding may be productive for investors and that this applies across the size spectrum. However, this does not seem to be particularly beneficial to bankers. Otherwise the coefficients and t-statistics are very much as those reported in columns 7-9 Table 5. That is, the introduction of controls for REPO finance does not alter the qualitative results found previously for funding efficiency, share of non-interests income, leverage and pure scale effects. This same property is true when we introduce the other detailed controls for wholesale market activity separately. Consequently, in Table 10 we report the regression coefficients and t-statistics of these variables omitting details regarding the other controls in the regressions.

We would consider securities lending, forward contracts and OTC options contracts as most representative of wholesale market activities as these involve significant counter-party risk which must be controlled through ISDA master agreements and collateral arrangements that are appropriate for institutional clients of a sufficient scale. Dealing in futures and exchange traded options is more accessible to retail oriented institutions, although, of course the big banks will also be present there. Focussing on securities lending, forwards and OTC options a striking pattern is seen in comparing coefficients across rows and columns. In the return to bankers regressions (column 2) these variables enter with positive and statistically significant coefficients, but the effect is greatly reduced or even totally eliminated by a negative and significant effect for these variable when interacted with the top size decile dummy, at10. In contrast, these variables are all insignificant in the regressions on investor returns (column 1).

That is, it looks like when smaller banks move into wholesale market activities like securities lending and OTC derivatives, it seems to boost earnings but these returns seem to accrue to bankers who are able bargain for enriched compensation contracts. This is reminiscent of the many stories one can find in the business press of smaller banks who pay over the odds to recruit a top-rated team of investment bankers.

For exchange traded derivatives this pattern does not hold very cleanly. Futures is positive and significant in the bankers' regression and this effect is reduced when interacted with at10. However, there is no significant effect for exchange traded options for bankers and no effect for either futures or exchange traded options in the bank investor regressions.

Finally, we have introduce all the wholesale market control variables both in levels and interacted with the top size dummy to see whether collectively these detailed controls would alter the qualitative results found previously. This is reported in Table 11. The main point to notice is that once we include all the wholesale market proxies, the pure scale effects, including the nonlinearity at the tenth decile, are insignificant in all three regressions. Otherwise, the qualitative conclusions from Section 4 are maintained in large part but with one significant exception. The coefficient of non-interest income (niish) now is insignificant in the bankers return equation. That is, the detailed wholesale market indicators we have introduced do seem to account for contributions of non-interest income in generating rents for bankers in banks below the top size decile. However, this does not seem to explain all the contribution for the biggest banks, as the non-interest income interacted with the top size dummy enters the bankers' regression positively and is highly significant. Otherwise, the results are similar to those found previously. There is a strong funding effect (nim) for investors that is weakened among the biggest banks. There is a positive funding effect in the bankers regression, but this does not apply to the largest banks since the level coefficient is counterbalanced by a negative coefficient on the interaction term. Finally, a higher capital ratio (ilev2) is associated with greater returns to investors but lower returns to bankers.

6 Interpreting the results

In this paper we have studied the returns to scale in banking by combining returns to shareholders and bankers to obtain a measure of total bank returns. We have found evidence of positive economies of scale that are stronger when bankers' returns are taken into account than when they are not. We find that these economies of scale to a significant degree can be accounted for by a combination of funding efficiency, leverage and involvement in wholesale banking activities. Specifically, when we include net interest margin, the capital to asset ratio, and the share of non-interest income in total net income, the pure scale effects are insignificant except in the top size decile. And when we include a large range of detailed proxies for wholesale banking activities, pure scale effects disappear altogether. Our results also suggest that the use of wholesale funding to expand the scale of the banks operations can increase bank returns even if it squeezes the bank's average net interest margin. However, these increased returns accrue largely to bankers rather than bank shareholders.

A basic premise of our approach is that the sharing of returns between investors and bankers is the result of a bargaining process where the relative bargaining power of bankers may be determined by the kinds of businesses that the bank undertakes. This is a fundamental departure from the assumption of competitive labor markets made in the traditional banking efficiency literature which found very little evidence of scale economies. Two recent papers maintain the assumption of a competitive market for bankers' services, but nevertheless, find some evidence of scale economies. Wheelock and Wilson (2012) find evidence of scale economics using a traditional static cost function approach by normalizing all input prices by the bankers' wage rate. Specifically, they adopt an empirical specification of the form $\frac{cost}{w3} = c(Y, w)$ where Y is a vector of bank outputs, $w = (K, C, w_1/w_3, w_2/w_3)$, K is physical capital, C is equity capital, and w_1, w_2 and w_3 are prices of purchased funds, deposits, and labor input respectively. Economies of scale are measured by the ratio $c(\lambda Y, w)/c(Y, w)$ for $\lambda > 1$. In line with standard practice in the literature they estimate w_3 as total banker compensation divided by employment. As discussed in Section 3.2 this tends to increase systematically with the size of the bank. Furthermore, as discussed in Section 4 funding costs also tend to rise with scale as many large banks rely more heavily on wholesale funds, with the effect that w_1/w_3 tends to be comparatively stable across size. Thus Wheelock and Wilson find that that ratio of total cost to the estimated wage rate tends to be decreasing with scale and interpret this as cost efficiency. In our case, we view the higher compensation of bankers as the result their ability to capture to some degree the increased returns realized by larger banks. Hughes and Mester (2011) also assume competitive input pricing and estimate scale economies using U.S. banking data for 2007. Using a standard cost function approach and including observed levels of capital as a control, they find *decreasing* returns to scale. It is only when they adopt an alternative specification which constrains the firm to choose an optimal (as opposed to observed) level of equity capital given an estimated shadow price of capital that they find increasing returns to scale.

While we have shown that observed economies of scale in total bank returns can be accounted for by observed proxies for funding efficiency, leverage and presence in wholesale banking markets, we do not claim that we have established a clear causal link between our explanatory variables and returns. There are at least four sets of more fundamental explanations that can be put forward in trying to establish such a causal link: (a) implicit public subsidies to banks viewed as "too-big-to-fail", (b) differences in banker skills that are not observable in our data set, (c) operational efficiencies that produce scale and scope economies and (d) market power in some segments of banking markets.

In light of the financial crisis starting in 2007-08, many analysts might be inclined believe that too-big-to-fail (TBTF) can account for most of our findings. Indeed, our data set does include such giants as Citigroup, which received government assistance in the recent crisis, and JPMorgan Chase and Bank of America which arguably benefitted indirectly from government programs in disposing of distressed assets of failed banks. Nevertheless, we do not believe that TBTF can account for the bulk of our results. First, it should be noted that our data set includes extremely large banks that were allowed to fail. Wachovia and Washington Mutual are examples of this. The top decile where we document large scale economies includes many banks that are small when compared to the very biggest banks. Banks at the 90'th percentile in our sample were smaller than the largest by a factor of 1/200. To verify whether our results can be attributed principally to TBTF, we did two robustness checks on our benchmark regression results (columns 7-9 of Table 5). Table IV columns 1-3 reproduces those regressions with the top 1% of banks excluded from the sample. The qualitative results are the same as in Table 5. In Table IV columns 4-6 we rerun the models including a dummy variable *Too-big-to-fail* for those banks ranked within the top 20 systemically important institutions (SIFI) in 2007 using the measure of marginal expected short-fall (Acharya *et al*, 2012). Again, the same pattern of signs and significance are found for our benchmark explanatory variables and in this case, the dummy *Too-big-to-fail* is insignificant.

It seems a harder challenge to test whether any of the remaining three candidate explanations can be excluded, and, if not, the degree to which observed scale economies can be attributed to each. To test the hypothesis that apparent increasing returns to scale that accrue to bankers merely reflects skill differences of the bankers working in different size banks, it would be interesting estimate supply of bankers' services using hedonic wage equations à la Mincer taking into account education, experience, and possibly past performance. However, micro data of this sort are not readily available. Some insight can be gained with aggregate data as shown in Philippon and Reshef (2011). They combine data on education and sector of employment derived from the U.S. Census of Population with wage information derived from the Annual Industry Accounts. They document an increase in skill in the U.S. financial services industry since 1980 that correlates closely with increases in wage rates. The increases in skills and wages are most marked in "other financial services" as compared to banking and insurance. They argue that some of these differences may reflect demand driven factors such as the rise of corporate restructuring as well as changes in the competitive environment brought about by deregulation. They argue that skill differences do not explain all the time variation in earnings and that in 2006 earnings in financial services are some 40% higher than would be justified by education alone. They suggest that the unexplained variation might reflect increased rents accruing to financial services.

Bell and Van Reenen (2010) working with U.K. data and Kaplan and Rauh (2010) working with U.S. data both document a sharp rise in compensation levels in the financial sector since 1990 and argue that this accounts for a significant share of observed increased income inequality. In particular, Bell and Van Reenen document a sharp increase in productivity in the financial sector. They point out that standard rent seeking arguments (Van Reenen (1996)) would predict as a consquence the increase in compensation in the financial sector relative to other sectors. The puzzle is why the increased compensation should particularly strong for the most highly paid bankers. Kaplan and Rauh show that the increase in compensation of the top bank executives cannot explain the increased share of the highest incomes attributable to the financial sector. Making an assumption about the likely shape of the income distribution within bulge-bracket investment banks combined with anecdotal evidence about the minimal compensation of a managing director within these institutions, they estimate the total compensation to bankers who would fall within the 99.9th percentile of the US income distribution. Extrapolating in this way they can account for some of the increased income inequality observed since 1990. They argue that increased demand for generalist managers, globalization, and changes in social norms cannot account for these results. According to them a more plausible explanation is that big increases in pay of bankers below the top executive level was driven by increased demand for specialist skills. Bell and Van Reenen also suggest that increased demand for specialist skills is the most plausible explanation for the observed increase in extreme income inequality and suggest that this could be attributable to the dynamics of superstars as described by Rosen (1981).

This line of reasoning has been developed recently by Garicano and Rossi-Hansberg (2006) in a way that suggests that improved information technology may have been crucial in bringing about the observed changes in earnings distribution in many firms. They view large organizations as "knowledge hierarchies." Agents are differentiated within organizations according to their ability to acquire skills in solving problems. Some agents are "producers" who deal with relatively routine tasks; while, others are "problem solvers" who deal with those that are more complex. Large organizations are able to gain efficiency by placing problem solvers relatively high in the hierarchy and sending the infrequently arriving, hard-to-solve problems up to the problem solver. Thus the specialist problem solver's skill is applied over a large scale and is not wasted on dealing with frequently arising but routine problems. However, the ability to achieve this efficiency gain will depend upon the cost of communication within the hierarchy. Improvements in communication technology allow specialists to leverage her knowledge. They show that reductions in the cost of accessing knowledge results in an increase in wage inequality within the organization.

This argument establishes a link between our finding of economies of scale that can be accounted for by the degree of involvement in specialized wholesale banking activities on the one hand and the Kaplan & Rauh and Bell & Van Reenen evidence that the financial sector seems to account for increased income inequality on the other. Improvements in information technology which require significant costs of investment can give large banks a cost advantage by employing new methods of intermediation between funding (e.g., deposit taking, trustee, and depository) and placement (lending, trading and corporate finance). A broker-dealer's system to support clients' shorting securities is a good example. Shorting involves identifying an available security to borrow, providing and maintaining collateral, and finding a replacement security, and delivering it on-time at the time of closing the borrowing leg. This is a complex set of tasks which if done by separate entities can accumulate significant operational costs. Potentially, there can be large benefits to integrating these tasks in a single organization, but to realize these gains the information systems used for each of the stages need to be linked. Making these linkages as seamless as possible can generate non-routine problems whose resolution requires specialist bankers. The bank that assembles the systems and teams that can achieve these operational advantages may have an important competitive edge that persists as rival banks struggle to catch-up or choose not to do so because they had lost the first-mover advantage. This may result in value for the bank, but because the specialist team may be difficult to replace were they to leave, they may be able to reap a large fraction of this value for themselves.

There are other examples of how information systems have changed banks' business models and may have contributed to the realization of scale and scope economies. DeYoung *et al* (2011) document the role of credit history data sets and credit scoring models in increasing the distance between lending institutions and their borrowers, thus expanding a bank's market reach. In a similar vein, Basel II introduced the use of internal models in the calculation of regulatory capital. The impact studies carried in the run-up to Basel II implementation showed that the costly, sophisticated systems which were adopted by the largest banks translated into a significant lightening of the credit risk and market risk charges. It is these same systems that have allowed the largest banks to meet heightened capital requirements mandated by Basel III by optimizing their portfolios to achieve reductions of risk-weighted assets. Ellul and Yerramilli (2010) provide evidence that investments in risk management expertise have had a significant impact on the bank performance in the adverse market conditions.

Again, the advantage achieved by early movers with efficient internal systems may bestow upon them a persistent competitive advantage. Thus decomposing the source of superior performance into factor productivity and market power is a difficult task. This is a problem that is well known in other areas where technology and market structure interact. An example outside of finance is that of economic geography where there is well-documented, robust evidence of "agglomeration economies", i.e., a positive relationship between total factor productivity of firms and the density of their location. However, how much of the improved productivity can be ascribed to cost savings and how much to market power remains an open and much debated question (Combes *et al* (2011)).

The interpretation of our results that comes from this line of reasoning is that large banks achieve a degree of operational efficiency which gives them a competitive edge over smaller rivals, and this advantage allows them to retain some producers' surplus that creates value for their shareholders and for bankers themselves. This interpretation of the forces that have given rise to the large scale economies in banking that we have documented seems to suggest a very different vision of the financial efficiency than that put forward by Phillipon (2012). One of his key arguments is that in the United States the share of GDP devoted to the financial sector has grown significantly since 1950 and that this trend closely correlates with an increase in the share of fixed investment spending devoted to IT in the financial sector. This is in stark contrast with the retail and the wholesale sectors which also saw a trend toward increased share of IT investment but where there has been a steady decline in the sectors' values added shares which apparently reflect efficiency gains (Philippon, 2011). How can the two set of results be reconciled?

In our view, an explanation of Philippon's findings is likely to be found in three factors that he does not fully consider and which differentiate U.S. finance from the retail and wholesale sectors. These are the globalization of finance, trends in aging and the financing of retirement, and increased income inequality. The globalization of finance means that Wall Street has played a central role in intermediating between savers and real investment on a global scale. This is a trend that has been present since wide-spread liberalization of capital flows starting in the 1970's. As a result, the relevant metric for gauging the scale of the American financial sector's value added is not American GDP but something much larger. The second missing factor is that increased longevity without a commensurate increase in retirement ages has meant that there has been a net increase in the transfers between the active population and the retirees which needs to pass through some channel which may be private or public or both. Furthermore, in the last three decades there has been an increased reliance on defined contributions, privately provided pensions. Therefore, an increasing share of these larger transfers is passing through the private financial sector. Finally, increasing income inequality during this same period may have reinforced this trend, since wealthy people live longer and are also more likely to save for a bequest motive. That is, increased income inequality means resources have flowed to segments of the population who naturally devote a higher fraction of their income on financial services.

7 Conclusion

Studies of economies of scale in banking based on pure cost efficiency measures and assuming competitive input pricing fail to satisfactorily account for the preponderance of a small number of very large banks within most financial systems. In this paper we have argued that a more fruitful starting point in understanding this phenomenon is to view banks as combining financial and human capital to create rents which are then allocated to investors and bankers through a bargaining process that will reflect the mix of businesses operated by the bank. To implement this approach empirically we have measured bank performance not only by investor returns (measured as return on equity) but also by an estimate of bankers' rents. Applying this approach to annual data of US bank holding companies since 1990, we find much stronger evidence of economies of scale in returns to bankers as compared to returns to investors. The scale economies appear to be particularly strong in the top size decile of banks measured by total assets. We find that these economies of scale are to a significant degree attributable to a bank's involvement in wholesale banking activities. The importance of wholesale banking in accounting for increasing returns to scale is apparent only when returns to bankers are taken into account. Our results also suggest that the use of wholesale funding to expand the scale of the banks operations can increase bank returns even if it squeezes the bank's average net interest margin. We find these increased returns accrue largely to bankers rather than bank shareholders. We suggest that the the most plausible explanation for our results is that large banks achieve a degree of operational efficiency which gives them a competitive edge over smaller rivals, and this advantage allows them to retain some producers' surplus that creates value for their shareholders and for bankers themselves.

References

Acharya, V., L. Pedersen, T. Philippon and M. Richardson 2012, "Measuring Systemic Risk" NYU Stern School Working Paper.

Bell, B. and J. Van Reenen , 2010 "Bankers' Pay and Extreme Wage Inequality in the UK," *Working Paper, Centre for Economic Performance, London School of Economics.*

Berger, A. N., W.C. Hunter, and S. G. Timme , 1993 "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present and Future," *Journal of Banking and Finance.* 17, 221-249.

Bolton, P., T. Santos, and J.A. Scheinkman , 2012, "Cream skimming in financial markets" Columbia University Working Paper.

Combes, P.P., G.Duranton, L. Gobillon, D. Puga, and S.Roux , 2011 "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *Working Paper, Aix-Marseille School of Economics*, forthcoming *Econometrica.*

DeMarzo, P. and M. Fishman, 2007, "Optimal Long-term Financial Contracting," *Review of Financial Studies* 20, 2079-2128.

DeYoung, R. , 2010, "Scale Economies Are a Distraction," *The Region.* Federal Reserve Bank of Minneapolis.

DeYoung, R., W.S. Frame, D.Glennon, and P. Nigro 2011, "The Information Revolution and Small Business Lending: The Missing Evidence," *Journal of Financial Services Research* 39, 1933.

Ellul, A and V.Yerramilli, 2010, "Stronger Risk Controls, Lower Risk: Evidence from U.S. Bank Holding Companies," working paper, Indiana University.

Garicano, L. and E. Rossi-Hansberg , 2006 "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics.* 121, 1383-1435.

Hughes, J. P., and L. J. Mester, 2011, "Who Said large Banks Don't Experience Scale economies? Evidence from a Risk-Return-Driven Cost Function," *Federal Reserve Bank of Philadelphia working paper.*

Kaplan, S. and J. Rauh , 2010 "Main Street and Wall Street: What Contributes to the Rise of the Highest Incomes?" *Review of Financial Studies.* 23, 1004-1050.

Philippon, T. 2012 "Has the U.S. Finance Industry Become Less Efficient?" NYU Stern School Working paper.

Philippon, T. 2011 "Finance vs. Wal-Mart: Why are Financial Services so Expensive?", forthcoming in A. Blinder, A. Lo, and R. Solow *Economic Lessons from the Financial Crisis.*

Philippon, T. and A. Reshef, 2011, "Wages and Human Capital in the U.S. Financial Services Industry: 1909-2006," NYU Stern School Working Paper, forthcoming, *Quarterly Journal of Economics.*

Rajan, R. and L. Zingales, 2000, "The Governance of the New Enterprise" in X. Vives (ed.) *Corporate Governance*. Cambridge University Press.

Reich, R. 2009, "Too big to fail: Why the big banks should be broken up, but why the White House and Congress don't want to" http://robertreich.org/post/257310704

Rosen, S. , 1981 "The Economics of Superstars," *American Economic Review.* 71, 845-858.

Van Reenen, J. , 1996 "The Creation and Capture of Economic Rents: Wages and Innovation in a Panel of UK Companies *Quarterly Journal of Economics.* 111, 195-226.

Wheelock D. C., and P. W. Wilson , 2012 "Do large Banks Have Lower Costs? New Estimates of Returns to Scale for U.S. Banks," *Journal of Money, Credit and Banking.* 44, 171-199.

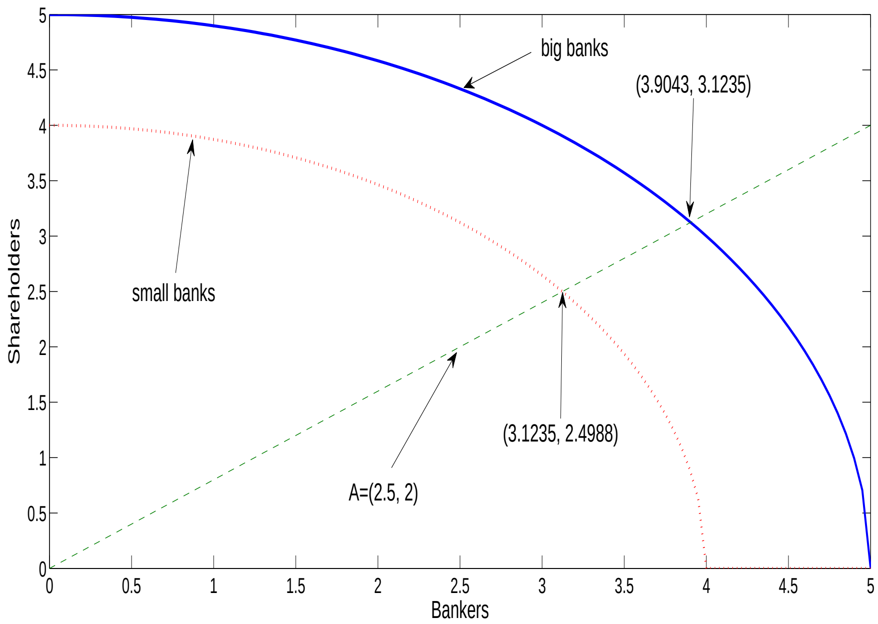

Figure 1: Measuring scale economies with two efficiency frontiers

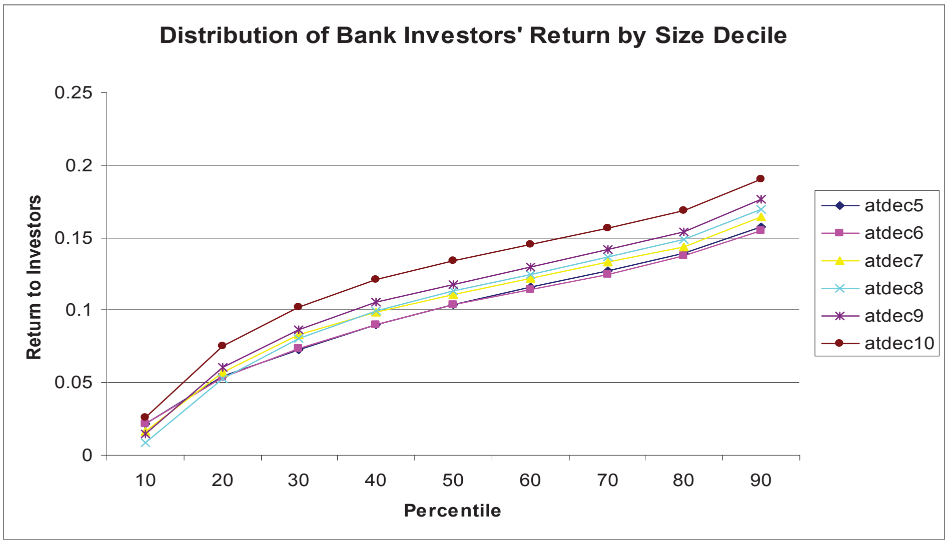

Figure 2: Cumulative Sample Distributions of Bank Investor Returns, Top Six Size Deciles

Table 1: Total Return to Investors, 1994-2010

	niseq				
Deciles of at	mean	25th tile	50th tile	75th tile	sd
1	-.021	.028	.056	.084	1.042
2	.058	.045	.076	.109	.268
3	.073	.054	.090	.119	.165
4	.081	.059	.098	.126	.116
5	.080	.062	.101	.131	.154
6	.042	.061	.101	.129	1.351
7	.067	.066	.107	.137	.367
8	.070	.067	.111	.143	.248
9	.060	.073	.115	.147	.393
10	.114	.095	.136	.164	.133

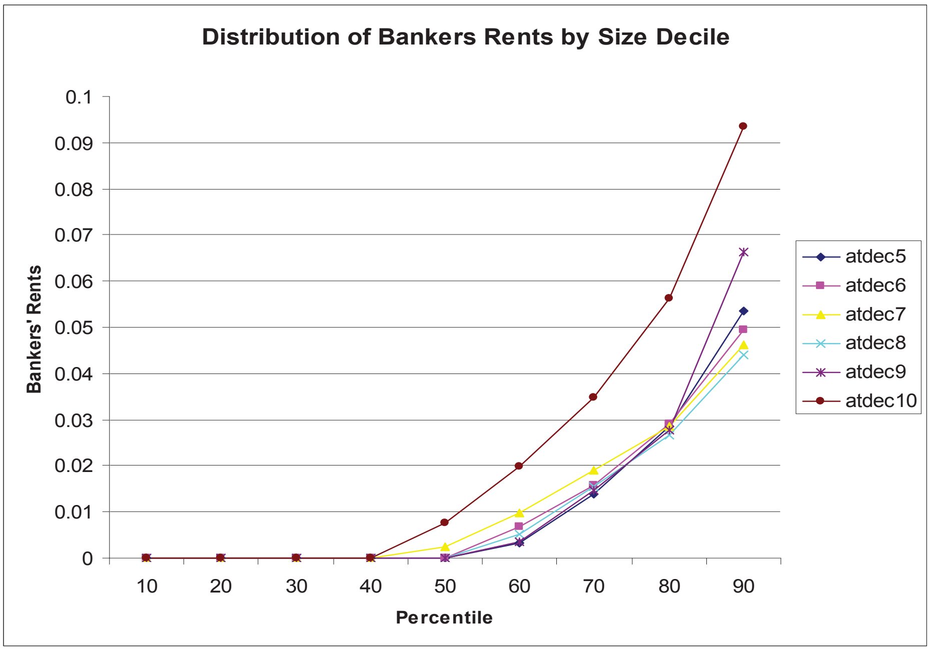

Figure 3: Cumulative Sample Distributions of Bankers' Returns, Top Six Size Deciles

Table 2: Rents per Banker ($1000), mxlrrentemp

Deciles of at	mean	25th tile	50th tile	75th tile	sd
1	6.602	0.000	0.000	6.697	14.247
2	6.102	0.000	0.000	7.139	13.045
3	4.385	0.000	0.000	4.113	9.495
4	9.923	0.000	0.000	5.915	99.463
5	8.652	0.000	0.000	6.953	89.001
6	9.048	0.000	0.000	9.738	79.654
7	7.287	0.000	1.199	9.558	12.763
8	6.809	0.000	0.722	9.377	13.108
9	9.122	0.000	0.424	10.780	18.171
10	13.224	0.000	5.946	18.664	21.160

Table 3: Returns to Bankers, mxlrrentseq

Deciles of at	mean	25th tile	50th tile	75th tile	sd
1	0.0142	0.0000	0.0000	0.0142	0.0729
2	.0.0186	0.0000	0.0000	0.0189	0.0552
3	.0.0142	0.0000	0.0000	0.0123	0.0356
4	.0.0165	0.0000	0.0000	0.0174	0.0403
5	.0.0171	0.0000	0.0000	0.0215	0.0367
6	.0.0159	0.0000	0.0000	0.0217	0.0284
7	.0.0170	0.0000	0.0038	0.0239	0.0295
8	.0.0171	0.0000	0.0022	0.0235	0.0366
9	.0.0290	0.0000	0.0011	0.0254	0.0824
10	.0.0347	0.0000	0.0157	0.0504	0.0495

Table 4: Summary Statistics of Explanatory Variables

variable	nim net interest margin	niish non-interest income	ilev2 capital ratio
mean	3.69	.235	.092
std dev	.909	.160	.034
corr w/ nim	1.000		
corr w/niish	-0.094	1.000	
corr w/ilev2	0.049	-0.059	1.000

Table 5: Linear Models

Dependent variable	niseq	mxlrrentseq	trentseq	niseq	mxlrrentseq	trentseq	niseq	mxlrrentseq	trentseq
at	0.000	0.000***	0.000***	0.000	0.000	0.000	0.000	0.000	0.000
	(1.26)	(4.95)	(3.13)	(0.66)	(1.54)	(0.51)	(0.76)	(1.43)	(0.42)
at10	0.027***	0.029***	0.029***	0.063**	0.014	0.043***	0.076***	0.009	0.037***
	(3.87)	(11.20)	(9.27)	(2.19)	(1.42)	(3.54)	(2.67)	(0.89)	(3.06)
nim				0.036***	0.007***	0.025***	0.035***	0.007***	0.026***
				(11.37)	(6.19)	(19.03)	(11.24)	(6.39)	(19.27)
nimat10				-0.015**	-0.012***	-0.017***	-0.018***	-0.011***	-0.015***
				(-2.28)	(-5.19)	(-5.81)	(-2.74)	(-4.73)	(-5.35)
niish				0.077***	0.042***	0.084***	0.082***	0.039***	0.082***
				(4.02)	(6.06)	(10.34)	(4.29)	(5.70)	(10.11)
niishat10				0.045	0.144***	0.110***	0.036	0.148***	0.114***
				(1.23)	(11.16)	(7.06)	(0.98)	(11.59)	(7.33)
ilev2							0.442***	-0.162***	-0.188***
							(5.67)	(-5.94)	(-5.70)
cons	0.133***	-0.002	1.525***	-0.030*	-0.035***	1.403***	-0.064***	-0.016*	1.418***
	(14.54)	(-0.22)	(373.24)	(-1.82)	(-3.92)	(203.50)	(-3.71)	(-1.72)	(193.16)
yr dummy	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-sq	0.200		0.285	0.230		0.383	0.236		0.388
Nobs	4077	4077	4077	4077	4077	4077	4077	4077	4077

The dependent variables are return on equity (niseq), bankers' rent as a per cent of equity (mxlrrentseq), and total rent (trentseq). The explanatory variables are total assets (at), a dummy variable if an observation is in the 10th size decile (at10), net interest margin (nim), nim interacted with at10 (nimat10), per cent of non-interest income in total revenues (niish), niish interacted with at10 (niishat10), ratio of book equity to total assets (ilev2), and year dummies. The regressions of niseq and trent are estimated by OLS. The mxlrrentseq model is estimated by Tobit regression. T-ratios are reported in parentheses. *, **, and *** indicates significant at the 10%, 5% and 1% levels respectively.

Table 6: Panel and Instrumental Variables Estimation

	Random effect model of bank returns			IV regressions		
Dependent variable	niseq	mxlrrentseq	trentseq	niseq	mxlrrentseq	trentseq
at	0.000	-0.000***	0.000	-0.000	-0.000	-0.000
	(0.97)	(-4.05)	(0.04)	(-0.21)	(-0.99)	(-0.68)
at10	0.080**	0.015*	0.015	0.189**	0.012	0.070*
	(2.35)	(1.82)	(1.01)	(2.21)	(0.60)	(1.93)
nim	0.048***	0.006***	0.032***	0.032***	0.011***	0.031***
	(12.35)	(5.30)	(18.12)	(5.72)	(8.32)	(12.84)
nimat10	-0.021***	-0.005**	-0.009**	-0.027*	-0.011***	-0.015**
	(-2.66)	(-2.51)	(-2.45)	(-1.91)	(-3.48)	(-2.56)
niish	0.088***	0.037***	0.095***	0.306**	0.213***	0.293***
	(4.22)	(7.84)	(10.69)	(2.58)	(7.93)	(5.80)
niishat10	0.044	0.034***	0.097***	-0.227	0.051	-0.034
	(1.00)	(3.14)	(4.99)	(-1.45)	(1.43)	(-0.51)
ilev2	0.838***	-0.216***	-0.042	1.115***	-0.337***	-0.263***
	(8.91)	(-9.08)	(-1.00)	(6.42)	(-8.58)	(-3.56)
cons	-0.268***	-0.004	1.309***	-0.289***	-0.044***	1.287***
	(-9.71)	(-0.67)	(112.73)	(-5.03)	(-3.36)	(52.70)
yr dummy	yes	yes	yes	yes	yes	yes
firm re	yes	yes	yes	no	no	no
iv for niish and niishat10	no	no	no	yes	yes	yes
R-sq (within)	0.298		0.406	0.237		0.293
Nobs	4077	4077	4077	2117	2117	2117

Table 7: Wholesale Bank Measures

Variable	Obs	Mean	Std. Dev.	Min	Max
repo	428	.056	.070	0	.460
sec lending	919	.123	.479	0	4.713
futures	791	.093	.276	0	2.193
forwards	791	.354	1.004	0	8.798
options	791	.078	.296	0	2.935
otc options	791	.280	.879	0	7.961
tier1 cap/ total assets	930	.087	.020	.043	.215

Table 8: Wholesale Measures Correlations

Variable	repo	sec lending	futures	forwards	options	otc options	tier1/at
repo	1.00						
sec lending	0.25	1.00					
futures	0.06	-0.01	1.00				
forwards	0.26	0.71	0.48	1.00			
options	0.05	-0.04	0.66	0.22	1.00		
otc options	0.02	-0.01	0.84	0.48	0.61	1.00	
tier1 / at	-0.26	-0.22	-0.07	-0.19	-0.03	-0.07	1.00

Table 9: Linear model: Bank returns with Repo

Dependent variable	niseq	mxlrrentseq	trentseq
at	-0.000	0.000	-0.000
	(-0.37)	(1.23)	(-0.46)
at10	0.106*	0.043**	0.032
	(1.77)	(2.40)	(1.32)
nim	0.035***	0.011***	0.030***
	(6.46)	(6.49)	(13.91)
nimat10	-0.022*	-0.020***	-0.015***
	(-1.73)	(-5.20)	(-2.85)
niish	0.059*	0.021**	0.060***
	(1.82)	(2.21)	(4.60)
niishat10	0.057	0.153***	0.143***
	(0.93)	(8.21)	(5.81)
repo	0.340***	0.008	0.170***
	(4.21)	(0.29)	(5.25)
repoat10	-0.210	-0.024	-0.026
	(-1.43)	(-0.50)	(-0.44)
ilev2	1.152***	-0.396***	-0.197***
	(6.93)	(-7.57)	(-2.95)
cons	-0.245***	-0.009	1.341***
	(-5.44)	(-0.61)	(74.22)
yr dummy	yes	yes	yes
R-sq	0.262		0.393
Nobs	2140	2140	2140

Table 10: Partial Correlations of Returns and Wholesale Indicators

Dependent variable	niseq	mxlrrentseq	trentseq
repo	0.340***	0.008	0.170***
	(4.21)	(0.29)	(5.25)
repoat10	-0.210	-0.024	-0.026
	(-1.43)	(-0.50)	(-0.44)
sec lending	0.161	0.295***	0.250***
	(0.90)	(5.86)	(3.47)
sec lending at10	-0.155	-0.271***	-0.227***
	(-0.86)	(-5.35)	(-3.14)
forwards	-0.105	0.468***	0.178**
	(-0.54)	(8.61)	(2.28)
forwards at10	0.117	-0.451***	-0.160**
	(0.60)	(-8.27)	(-2.04)
otc options	-0.018	0.087***	0.036*
	(-0.36)	(6.05)	(1.76)
otc opt at10	0.001	-0.069***	-0.036
	(0.02)	(-4.29)	(-1.59)
futures	-0.038	0.090***	0.023
	(-0.70)	(5.90)	(1.06)
futures at10	0.013	-0.054***	-0.018
	(0.19)	(-2.72)	(-0.63)
options	3.134	0.107	0.584
	(0.84)	(0.10)	(0.39)
options at10	-3.175	-0.071	-0.585
	(-0.85)	(-0.07)	(-0.39)

Table 11: Linear Model with All Wholesale Indicators

Dependent variable	niseq	mxlrrentseq	trentseq
at	-0.000	-0.000	-0.000
	(-0.28)	(-0.74)	(-0.62)
at10	0.108*	0.009	0.015
	(1.69)	(0.49)	(0.59)
nim	0.035***	0.009***	0.030***
	(6.38)	(5.90)	(13.60)
nimat10	-0.022	-0.010**	-0.008
	(-1.52)	(-2.27)	(-1.40)
niish	0.061*	0.003	0.052***
	(1.83)	(0.33)	(3.89)
niishat10	0.057	0.134***	0.123***
	(0.80)	(6.62)	(4.30)
ilev2	1.168***	-0.349***	-0.183***
	(6.89)	(-6.96)	(-2.69)
repo	0.338***	-0.034	0.153***
	(4.03)	(-1.26)	(4.56)
repoat10	-0.227	-0.000	-0.030
	(-1.48)	(-0.00)	(-0.50)
sec lending	0.043	0.235***	0.166**
	(0.21)	(4.20)	(2.04)
sec lending at10	-0.068	-0.209***	-0.155*
	(-0.33)	(-3.68)	(-1.87)
futures	-0.014	0.028	0.004
	(-0.20)	(1.44)	(0.13)
futures at10	0.007	-0.022	-0.001
	(0.07)	(-0.79)	(-0.02)
forwards	-0.142	0.248***	0.038
	(-0.54)	(3.39)	(0.36)
forwards at10	0.170	-0.246***	-0.028
	(0.64)	(-3.36)	(-0.27)
options	2.650	0.266	0.435
	(0.71)	(0.26)	(0.29)
options at10	-2.677	-0.233	-0.434
	(-0.72)	(-0.23)	(-0.29)
otc options	0.001	0.045***	0.029
	(0.02)	(2.62)	(1.15)
otc opt at10	-0.018	-0.036*	-0.029
	(-0.24)	(-1.82)	(-1.02)
cons	-0.248***	0.001	1.345***
	(-5.44)	(0.04)	(73.92)
yr dummy	yes	yes	yes
R-sq	0.262		0.397
Nobs	2117	2117	2117

Appendix

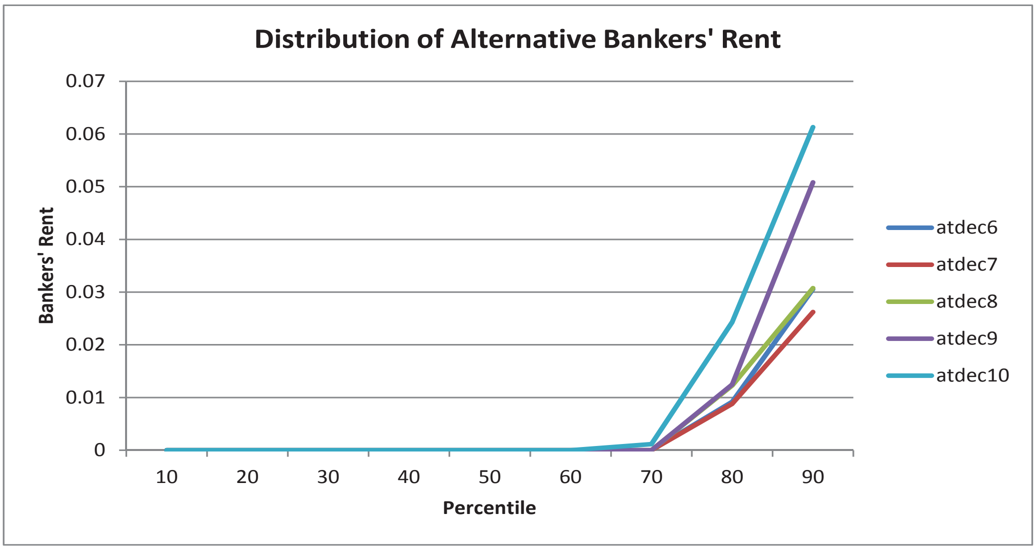

Figure A: Alternative Rent Measure, Sample Distributions

Table I: Return on Equity (*niseq*) versus Return on Assets (*niat*)

Dependent variable	*niseq*	*niat*
at	0.000	0.000
	(0.76)	(0.29)
at10	0.076***	0.002
	(2.67)	(1.05)
nim	0.035***	0.003***
	(11.24)	(17.14)
nimat10	-0.018***	-0.001*
	(-2.74)	(-1.86)
niish	0.082***	0.006***
	(4.29)	(5.09)
niishat10	0.036	0.008***
	(0.98)	(3.63)
ilev2	0.442***	0.052***
	(5.67)	(11.57)
cons	-0.064***	-0.008***
	(-3.71)	(-7.67)
yr dummy	yes	yes
R-sq	0.236	0.336
Nobs	4077	4077

Table II: Measuring size by the logarithm of total assets (*lnat*)

Dependent variable	niseq	rent	trent
lnat	0.000	0.007***	0.003**
	(0.09)	(5.97)	(2.56)
at10	0.076**	-0.009	0.028**
	(2.56)	(-0.89)	(2.19)
nim	0.035***	0.007***	0.026***
	(11.24)	(6.42)	(19.28)
nimat10	-0.019***	-0.010***	-0.015***
	(-2.77)	(-4.46)	(-5.23)
niish	0.082***	0.033***	0.079***
	(4.25)	(4.88)	(9.71)
niishat10	0.040	0.143***	0.111***
	(1.10)	(11.29)	(7.16)
ilev2	0.441***	-0.162***	-0.188***
	(5.66)	(-5.95)	(-5.71)
cons	-0.066**	-0.066***	1.392***
	(-2.28)	(-5.24)	(112.82)
yr dummy	yes	yes	yes
R-sq	0.236		0.389
Nobs	4077	4077	4077

Table III: Using alternative measures of bankers' rents

Dependent variable	mxlrrentseq	mxlrrentseq2	mxlrrentseq3	trentseq	trentseq2
at	0.000	0.000	0.000***	0.000	-0.000**
	(1.43)	(1.56)	(2.62)	(0.42)	(-2.28)
at10	0.009	0.007	-0.019	0.037***	-0.025
	(0.89)	(0.65)	(-1.55)	(3.06)	(-0.99)
nim	0.007***	0.007***	0.006***	0.026***	0.011***
	(6.39)	(6.02)	(5.05)	(19.27)	(4.12)
nimat10	-0.011***	-0.011***	-0.015***	-0.015***	0.001
	(-4.73)	(-4.55)	(-5.53)	(-5.35)	(0.15)
niish	0.039***	0.036***	0.019**	0.082***	0.027
	(5.70)	(5.25)	(2.57)	(10.11)	(1.62)
niishat10	0.148***	0.152***	0.179***	0.114***	0.122***
	(11.59)	(11.73)	(11.51)	(7.33)	(3.79)
ilev2	-0.162***	-0.151***	-0.115***	-0.188***	-1.323***
	(-5.94)	(-5.44)	(-3.76)	(-5.70)	(-19.36)
cons	-0.016*	-0.016*	-0.033***	1.418***	0.196***
	(-1.72)	(-1.68)	(-3.10)	(193.16)	(12.94)
yr dummy	yes	yes	yes	yes	yes
R-sq				0.388	0.119
Nobs	4077	4077	4077	4077	4077

mxlrrentseq is bankers' rent as a per cent of equity, where the competitive wage used in calculations is based on banks with more than 50 employees and less than \$1 billion in total assets. mxlrrentseq2 is bankers' rent as a per cent of equity, where the competitive wage used in calculations is based on banks with more than 50 employees and less than \$2 billion in total assets. mxlrrentseq3 is bankers' rent as a per cent of equity, where the competitive wage used in calculations is based on size decile the bank belongs to. trentseq is total rent (trentseq), where the competitive wage used in calculations is based on banks with more than 50 employees and less than \$1 billion in total assets. trentseq2 is bankers' rent as a per cent of equity, where the competitive wage used in calculations is based on banks with more than 50 employees and less than \$2 billion in total assets. The explanatory variables are total assets (at), a dummy variable if an observation is in the 10th size decile (at10), net interest margin (nim), nim interacted with at10 (nimat10), per cent of non-interest income in total revenues (niish), niish interacted with at10 (niishat10), ratio of book equity to total assets (ilev2), and year dummies. The regressions of niseq and trent are estimated by OLS. The mxlrrentseq model is estimated by Tobit regression. T-ratios are reported in parentheses. *, **, and *** indicates significant at the 10%, 5% and 1% levels respectively.

Table IV: Sensitivity to too-big-to-fail

	Excluding top 1% of banks			SIFI dummy		
Dependent variable	niseq	mxlrrentseq	trent	niseq	mxlrrentseq	trent
at	0.000	0.000***	-0.000	0.000	0.000	-0.000
	(0.04)	(2.69)	(-0.05)	(0.35)	(1.06)	(-0.11)
at10	0.078***	0.003	0.037***	0.075***	0.009	0.036***
	(2.61)	(0.25)	(2.94)	(2.63)	(0.89)	(2.99)
nim	0.035***	0.007***	0.026***	0.035***	0.007***	0.026***
	(11.13)	(6.45)	(19.17)	(11.24)	(6.39)	(19.27)
nimat10	-0.019***	-0.009***	-0.015***	-0.018***	-0.011***	-0.015***
	(-2.69)	(-3.90)	(-5.04)	(-2.68)	(-4.70)	(-5.24)
niish	0.081***	0.039***	0.082***	0.082***	0.039***	0.082***
	(4.25)	(5.68)	(10.07)	(4.29)	(5.70)	(10.11)
niishat10	0.039	0.137***	0.113***	0.036	0.148***	0.115***
	(1.02)	(10.38)	(6.98)	(0.99)	(11.58)	(7.34)
ilev2	0.443***	-0.160***	-0.186***	0.442***	-0.162***	-0.188***
	(5.63)	(-5.82)	(-5.60)	(5.67)	(-5.94)	(-5.68)
Too-big-to-fail				0.011	-0.001	0.010
				(0.27)	(-0.06)	(0.59)
cons	-0.167***	-0.016*	1.360***	-0.064***	-0.016*	1.417***
	(-6.21)	(-1.70)	(119.67)	(-3.71)	(-1.72)	(193.04)
yr dummy	yes	yes	yes	yes	yes	yes
R-sq	0.236		0.383	0.236		0.388
Nobs	4014	4014	4014	4077	4077	4077