DISCUSSION PAPER SERIES

No. 4786

PORTFOLIO DIVERSIFICATION, PROXIMITY INVESTMENT AND CITY AGGLOMERATION

William Goetzmann, Massimo Massa
and Andrei Simonov

FINANCIAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP4786.asp**

ISSN 0265-8003

PORTFOLIO DIVERSIFICATION, PROXIMITY INVESTMENT AND CITY AGGLOMERATION

William Goetzmann, Yale University
Massimo Massa, INSEAD, Fontainebleau and CEPR
Andrei Simonov, Stockholm School of Economics

Discussion Paper No. 4786
December 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 4786

December 2004

ABSTRACT

Portfolio Diversification, Proximity Investment and City Agglomeration*

We study the puzzle of portfolio underdiversification and proximity investment from a novel perspective, linking it to the process of urbanization. We find that urban portfolios are more focused – i.e., less diversified and more concentrated in 'close' stocks. We explain it in terms of the process of 'professional specialization' that characterizes urban environments. We test this against a number of alternative theories: financial sophistication, social competition and hedging non-financial risk. We show that the very same factors behind the drive to city agglomeration also affect both the degree of portfolio diversification and proximity investing by influencing investor information and risk.

JEL Classification: G11 and G14
Keywords: city agglomeration, portfolio choice, professional specialization, proximity investment and under-diversification

William Goetzmann
School of Management
Yale University
Box 208200
New Haven CT 06520-8200
USA
Tel: (1 203) 432 5950
Fax: (1 203) 432 3003
Email: william.goetzmann@yale.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=115358

Massimo Massa
Finance Department
INSEAD
Boulevard Constance
77305 Fontainebleau Cedex
FRANCE
Tel: (33 1) 6072 4481
Fax: (33 1) 6072 4045
Email: massimo.massa@insead.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=144755

Andrei Simonov
Department of Finance
Stockholm School of Economics
Sveavägen 65, Box 6501
SE-113 83 Stockholm
SWEDEN
Tel: (46 8) 736 91 59
Fax: (46 8) 31 23 27
Email: andrei.simonov@hhs.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=153466

* We are grateful to Sven-Ivan Sundqvist for numerous helpful discussions and for providing us with the data. Andrei Simonov acknowledges financial support from the Stockholm Institute for Financial Research and Jan Wallander och Tom Hedelius Stiftelse. All remaining errors are ours.

Submitted 29 November 2004

1 Introduction

In recent years, the economic literature has uncovered two features of investor behavior at odds with standard portfolio theory: the lack of portfolio diversification and proximity investment. Contrary to what standard theory would require, investors hold very undiversified portfolios made up of a limited number of stocks (Barber and Odean 2000, 2003, Goetzmann and Kumar, 2002), and they mostly select these stocks on the basis of geographical or professional proximity to them (Coval and Moskowitz, 2001, Zhu, 2003).

Different theories can explain the lack of portfolio diversification (standard portfolio theory in the presence of non-financial risk, financial sophistication and limited information and social competition), or proximity investment (financial sophistication and limited information and social competition). However, the lack of good quality data has prevented a test of the competing explanations in a framework that *simultaneously* accounts for both phenomena.

We bridge this gap by providing empirical evidence that can help to explain both phenomena and reconcile different theories. We analyze the problem from a different perspective, considering the link between portfolio choice and the process of professional specialization related to city agglomeration. The economic influence of cities on both enterprise and human capital is so important that it is natural to ask whether its effects can be documented in the financial portion of the household investment portfolio. Our claim is that the process of professional specialization that is behind the drive to city agglomeration also affects both the degree of portfolio diversification and proximity investing. In particular, it does this by influencing the investor cost of processing information and the availability of "superior" information.

The process of professional specialization taking place in urban agglomerates changes the way in which investors process financial information. On the one hand, as professional specialization increases and labor market competition rises, the time investors can allocate to search for new stocks decreases. Investors are more focused on their profession and have less time to allocate to financial search. If we assume some sort of bounded rationality (Simon, 1986, Sargent, 1993) or limited processing capacity (Sims, 2000, Peng and Xiong, 2002), an increase in specialization and competition in the investor professional life may reduce the ability to process financial information and to invest in financial assets in general.

On the other hand, professional specialization increases the knowledge of a particular sector/industry. This knowledge spill-over from the professional to the financial life of the investor effectively reduces his search cost for assets correlated to his profession – in a sense, the

investor is free-riding on the informational spillover arising from his professional knowledge. The value-relevant part of this information tends to be concentrated in stocks closer – professionally or geographically – to the investor. Urban households exposed to profession-related knowledge spillover will use their investment portfolio to take advantage of such knowledge, by concentrating their portfolios in close stocks.

These two effects – the relatively higher opportunity cost to "digest" (Peng and Xiong, 2002) financial information and the relatively higher exposure to value-relevant information that business interaction yields in the city – generate "rational inattention" (Sims, 2000) for assets not closely related to the investor's professional life. This *reduces* portfolio diversification and *increases* the investment in closer stocks. Given that knowledge spillover is a major characteristic of cities, one might expect to find portfolio concentration in urban, rather than rural settings. "Closer" investment opportunities may be expected to appear more frequently in an urban environment, and to yield economic returns to those who seize them. This implies that the very factors that make a city dynamic – professional specialization and knowledge spillover – should also affect investor portfolio diversification and decision to invest in closer stocks.

However, cities also increase the degree of *financial sophistication* of the investor – providing the availability of financial services – and change the *degree of non-financial risk* – providing more and better employment opportunities. In this context, if the city is a technology that provides better financial sophistication and facilitates the diversification of human capital and non-financial income risk, it should also *increase* portfolio diversification and *reduce* the investment in closer stocks.[2]

Therefore, the cross-sectional differences afforded by urban versus rural locations as well as the cross-sectional differences among urban areas themselves may help us to shed new light on how people invest in financial assets and to distinguish different theories such as professional specialization, financial sophistication, hedging non-financial risk and social competition.

The major challenge is the availability of comprehensive and detailed data. Indeed, testing requires not only a detailed knowledge of the assets in the financial portfolio but also information about employment, housing, geographical location, and even demographic characteristics, to be able to control for the other confounding effects. We deal with this issue by using a new and unique dataset that contains information on individual investors, traced over time. In fact, we have access to a very detailed geo-coded panel dataset over the period 1995 to

[2] Indeed, if cities allow a better hedging of non-financial risk, they should reduce the need to use the financial portfolio to hedge non-financial income risk. With non-traded human capital, the *financial* portion of the investment portfolio is used to hedge *non-financial* income risk (Heaton and Lucas, 2000, a & b, and Vissing-Jørgensen, 2002a).

2000 for a large and representative sample of households in Sweden. Constructed from several sources, this dataset allows us to explore the geographical determinants of investment. We have available information on investors' wealth, broken down at the asset level (cash, stocks and mutual funds, real estate, loans, bonds and other assets). We also have their income and tax position as well as their demographic characteristics. We use the panel data to test a number of implications about both the risk and return stories presented above.

For each investor, we construct measures of portfolio diversification and proximity investment, as well as proxies that account for the degree of professional specialization in the district in which he lives, isolation, local prosperity, financial services availability, non-financial risk, as well as a set of control variables meant to control for investors' idiosyncratic wealth, income and demographic characteristics and for geographic and country shocks.

Our results are consistent with the professional specialization hypothesis. In contrast, we find scarce evidence that investor behavior can be explained in terms of financial sophistication, social competition or need to hedge non-financial risk. In particular, we show that, all else equal, city investors are more focused (less-diversified), more concentrated in (professionally and geographically) close stocks, and that the tendency to focus the portfolio increases with the degree to which the city itself is industry-focused. An increase in the determinants of professional specialization – i.e., industrial specialization, competition for employees and concentration of the enterprises operating in the district – reduces the degree of portfolio diversification and increases proximity investment. In contrast, all else equal, economic and geographic isolation increase diversification and reduce proximity investment. The economic success of the district – both in terms of growth and profitability of the industries that are located there – as well as the availability of financial services increase portfolio diversification, but have mixed effects on proximity investment.

These results are robust to the inclusion of controls for non-financial risk and other factors affecting portfolio choice. We are able, for the first time to our knowledge, to control for almost all the alternative factors and determinants of portfolio choice brought forward by theory or uncovered before, such as investor's income (level and variance of labor and entrepreneurial income of the investor and the correlations between them and financial and real estate income, capital gains and losses), wealth (financial and real estate wealth), employment risk, momentum (return on financial portfolio, return on the market portfolio), investor's demographic characteristics (high-school and university education, age, the degree of stock market participation, size of the household, ability of the investor in his profession) and macroeconomic

and social effects (immigration, capital city effect, local tax rate, size of local administration, business cycle, consumer confidence, local background risk).

In this paper, we make five contributions. First, we add to the existing literature on investment and portfolio choice, showing that financial diversification and proximity investment are strongly determined not only by non-financial risk (Merton, 1971), social interaction (Hong, Kubik and Stein, 2004) and information (Merton, 1987, Coval and Moskowitz, 1999, 2001, Shapiro, 2002), but also by the degree of professional specialization that characterizes the urban and productivity context in which the investor lives. We also provide a direct link between portfolio choice and urban growth, suggesting that the same economic forces that determine the process of agglomeration are also affect individual portfolio choice.

Second, and more broadly, we address the question of the role that a city plays in the definition of the financial markets. If the very process that drives city agglomeration also stifles diversification and increases proximity investing, this may have important consequences in terms of the development of a viable, deep and efficient financial market. Depending on whether local professional specialization or local prosperity prevails, the development of cities may harm or spur the development of financial markets.

Third, we contribute in terms of the methodology we employ and the data we use. In particular, we are able to separately identify financial diversification from hedging. The fact that investors hold a very undiversified portfolio is, in general, assumed to imply that they do not hedge their non-financial income risk. However, this is not necessarily the case. Few stocks negatively correlated to non-financial income risk may be sufficient to diversify it away fully. Conversely, a well-diversified financial portfolio does not necessarily provide hedging against non-financial income risk. Therefore, financial portfolio diversification needs to be analyzed jointly with the degree to which the investor uses his financial portfolio to hedge non-financial risk. This has been impossible up to now, due to the lack of data that allows us to separately identifying financial diversification from hedging. In this paper, we bridge this gap.

Fourth, our results provide some new useful evidence for the theories of agglomeration. Cities appear to enhance risk-taking behavior based on the economic opportunities they offer – or appear to offer – to their inhabitants. The fact that the very factors economists have associated with urban growth – particularly industry scale, diversity and concentration – are associated with investor risk-taking suggests the existence of "financial channel" that magnifies the effects of agglomeration. Agglomeration, by reducing diversification and spurring financial investment in local assets, increases the size of the local economy, stimulating further agglomeration. This self-

reinforcing process may help to explain the accelerated growth rate of many urban agglomerates around the world.

Finally, we are able to shed some light on the home bias puzzle. Indeed, in an international setting, the problem of under-diversification, coupled with the tendency to invest in "closer" stocks, shows up as a home bias puzzle – i.e., the preference to invest in stocks of the home country. The two puzzles – under-diversification and home bias – are strongly related. The analysis of the under-diversification puzzle in a domestic context allows us to say something on the home bias, without the confounding effect of currency fluctuations. Our results would suggest that the process of economic development that leads to professional specialization, city agglomeration and development of human capital, would also, all else equal, worsen the home bias.

The remainder of the paper is structured as follows. In the next section, we discuss the literature related to our work, with particular focus on the various theories that might explain investor focus, and the testable implications, given our dataset. Section 3 discusses the data, Section 4 describes the construction of the variables. Section 5 describes the estimation procedure. Section 6 reports the empirical results. A brief conclusion follows.

2 Background and testable hypotheses

2.1 Portfolio choice and professional specialization

We consider an approach that links portfolio underdiversification and proximity investment to the process of professional specialization driving urban growth.

Standard information models (e.g., Grossman and Stiglitz, 1980) show that, in general, investors tend to invest more in the assets in which they have better information (Frieder and Subrahmanyam, 2004). More information reduces uncertainty and therefore the perceived riskiness of the asset. A discrete (information or transaction) cost can induce investors to focus on few or even just one asset only. Similarly, if investors are allowed to optimally purchase information and information has increasing returns, investors would exploit the increasing returns to specialization by specializing in learning and investing in few assets (Van Nieuwerburh and Veldkamp, 2004). Alternatively, consider a search model (Carlson and McAfee, 1983). Investors are aware that different stocks have different risk-return profiles, but are uncertain about the one that represents the next best addition to the portfolio, and face a search cost to investigate it. It can be shown (Hortacsu and Syverson, 2004) that they will adopt the following investment rule: keep

on investigating one additional stock if the cost of searching is less than the benefit, otherwise stop and invest in the already investigated stocks.

Both these approaches – i.e., the standard learning one and the search one – explain portfolio focus (i.e., proximity investment and underdiversification) by assuming the existence of discreteness in the information cost that effectively segments the market and generates a discontinuity in the investment process. The cost can be seen as an amount that has to be paid to enter a specific market. It is positively related to the cost of purchasing information as well as the opportunity cost of the time spent doing so. The latter increases with the number of alternative opportunities the investor faces and, therefore, with the profitability of his professional activity.

We argue that this information cost is directly related to professional activity of the investor. The more professionally involved the investor is and the more "remunerative" his profession is, the less time he will have to devote to portfolio allocation. We can think of this in terms of bounded rationality (Simon, 1986, Sargent, 1993) or limited processing capacity (Sims, 2000). In a similar context, building on Sims' results in terms of limited processing capacity, Peng and Xiong (2002) show how limited capacity, defined in terms of time to process information, affects portfolio choice. As professional specialization increases, and labor market competition rises, the time investors can allocate to search for new stocks decreases. Investors are more focused on their profession and have less time to allocate to financial search. The increase in the (psychic) cost of processing information induces the investor to focus on the few assets is more comfortable with.

However, this increase in the information cost, by itself, would not explain proximity investment. Indeed, all else equal, a rational investor with no additional information would invest only in an index fund. In fact, professional specialization also endows the investor with additional information, by-product of his labor. This additional information effectively reduces the investor's search cost for assets correlated to his job – in a sense, the investor is free-riding on the informational spillover arising from his professional knowledge. Therefore, professional specialization increases the search costs for the non-professionally related assets and reduces it for the professionally related ones. This increases the incentives to invest in "close" stocks.

Our claim is that one of the main factors that both increase the cost of information on financial assets as a whole and provide privileged information on a selected subset of them is the phenomenon of professional specialization. This allows us to formulate the following hypothesis:

H1: Portfolio underdiversification and proximity investment are positively related to the degree of professional specialization of the investors.

Where does professional specialization come from and how can we quantify it? To answer these questions we borrow from urban economics. What we have loosely defined up to now as "professional specialization" has been analyzed by urban economists in terms of the process of knowledge creation and knowledge focus that leads to urban agglomeration and city growth (Glaeser *et al*., 1992 and Glaeser and Mare, 2001).[3] Agglomeration economies highly specialized in a specific industry increase the returns to investment in high labor-intensive technologies and induce professional specialization. For example, Marshall (1890), Arrow (1962) and Romer (1986) [MAR] argue that the higher the degree of industrial specialization of the district the higher is the intra-industry knowledge spillover and specialized professional knowledge. Porter (1990) claims that the main determinant of growth is *industrial specialization.* The archetypal MAR region is Silicon Valley. Concentration of the high-tech industry around San José, California, generates knowledge spillover between firms in the same industry and knowledge focus.[4]

At the same time, industrial specialization spurs innovation if there is enough competition between firms in the industry. For example, Jacobs (1969) suggests that local *competition between firms* induces knowledge creation and innovation. The given example is the Italian ceramics industry, in which technical and aesthetic innovation is driven by intense, local competition among many similar firms.[5] Moreover, the presence of intense competition between firms operating in the same industry allows the worker to appropriate part the benefits of his professional specialization. Indeed, professional specialization and focus pay off for a local worker if he can "auction off" his knowledge to many competing firms in the industry.

These theories suggest that industrial specialization and competition are among the main drivers behind the process of knowledge creation and focus that generates professional specialization and leads to city agglomeration. We argue that these very same variables are also driving portfolio focus (portfolio diversification and proximity investment). As an example, let us consider, again, the case of Silicon Valley – i.e., a very specialized and competitive economy, concentrated in high-tech activities with many competing firms. In such an economy, there is a knowledge spillover from professional to financial expertise, as investors will tend to extrapolate

---

[3] See, Ellison and Glaeser, 1999, Glaeser *et al*. 1992, 1995, 1997 and 2002, Glaeser and Mare, 2001.

[4] Jacobs (1969) describes the development of industries that grow out of specialization of a particular process. The example is the brassiere industry, which evolved from the dressmakers' industry as opposed to the lingerie industry.

[5] In tests of these off-setting propositions, Glaeser *et al.,* 1992 find that "important knowledge transfers come from outside the core industry"'. Cross-industry spillover, as opposed to within-industry spillover, is found to be the main driver to city growth.

financial performance from the technological knowledge derived from the companies they work for. Time constrained and highly professionally specialized investors will tend to make the best use of their relative informational advantage based on professional expertise. Industry specialization and concentration will implicitly make investors coordinate on the same stocks. In other words, everybody will tend to evaluate stocks by using the same criteria. This "bias" will be stronger, the higher the degree of industrial knowledge creation in the district and its spillover. This approach is consistent with recent findings that explain interpersonal effects on consumer purchases between neighbors on the basis of information sharing (Grinblatt, Keloharju, and Ikaheimo, 2004).

This allows us to find proxies for professional specialization. Indeed, we can directly focus on the factors that the literature has identified as drivers of the process of knowledge creation and focus – i.e., industrial specialization and competition – and see how they affect the investment focus of the local investors. We will elaborate more on this topic in Section 4. We now consider the alternative hypotheses.

2.2 Alternative hypotheses

One explanation for the lack of portfolio diversification and proximity investment is related to the degree of financial sophistication of the investors. The lack of good-quality financial information – for example due to the geographical isolation of the investor – may make investors unaware of the existence and characteristics of all the stocks. They would, therefore, just invest in the few of which they are aware (Merton, 1987 and Shapiro, 2002) or with which they are familiar (Grinblatt and Keloharju, 2000, 2001). Superior information about a few stocks implies a portfolio concentrated in those stocks (Grinold and Kahn, 1999). Notice that this hypothesis assumes that what is relevant is the availability of raw financial data – i.e., financial analysts, banks, financial services. [6] The professional specialization hypothesis is, instead, more related to the psychic cost of processing information given the limited capability of the investor fully involved in his professional activity. This will provide different testable restrictions.

The effect of informational awareness is compounded in the case of the preference for skewness. Shore and White (2002) and Polkovnichenko (2003) point out that those investors with a taste for low-probability high-stakes gambles will invest only in the very few stocks from which they expect – because of better information or familiarity – higher returns, forfeiting the benefits

[6] Additional evidence of this information effect is the fact that less financially sophisticated investors seem to be less diversified (Goetzmann and Kumar, 2002).

of diversification in return for a chance at great wealth. We will group these theories under the label of the financial sophistification hypothesis. This hypothesis posits:

H2: Portfolio underdiversification and proximity investment are negatively related to the degree of financial sophistication of the investors.

A second explanation is a theory of social competition and community effects. Bakshi and Chen (1996) for example, build an economy in which investors are motivated by the social status of wealth – presumably with respect to their geographical neighbors. DeMarzo, Kaniel and Kremer (2004) argue that investors – competing for local resources within their district – have utility that depends on both their own wealth and the aggregate district wealth. Investors, in order to keep up with their neighbors, invest in the same stocks in which their neighbors invest, creating un-diversified portfolios.[7] This theory postulates that the degree of social competition in a district, the more the investors concentrate on the same stocks. This allows us to posit:

H3: Portfolio underdiversification and proximity investment are positively related to the degree of social competition of the investors.

Another reason why we observe financially under-diversified portfolios may be related to hedging non-financial risk. In a Merton (1971) model, for example, investors will hold the market portfolio and some additional position constructed to hedge the changes in other non-financial sources of uncertainty. An investor, subject to non-financial income risk, may try to diversify it away by holding, in addition to the market portfolio, a portfolio negatively correlated to his non-financial income. The net effect may be a *financial* portfolio that appears undiversified, when, in fact, the investor has reduced the *overall* (financial and non-financial) risk of his portfolio. That is, *portfolio underdiversification is positively related to the need of the investors to hedge their non-financial income risk.* However, given that the size of the effect depends on the overall financial and non financial portfolio of the investor, we consider this as an alternative explanation to control for, but not necessarily an alternative hypothesis. Moreover, given that professionally or geographically close stocks tend to be more correlated with investor non-financial income, standard hedging arguments are not well suited to explain proximity investment. We need,

[7] It is important to note that this theory assumes away the mere possibility of migration across municipalities, while we will argue that this is one of the main components of the story.

however, to properly control for it is that urban agglomeration allows a better hedging of non-financial risk. This reduces the need to use the financial portfolio to hedge non-financial risk.

Finally, it is important to notice that in our sample – i.e., Sweden – investors are not forced or induced to own company stocks. This is not only confirmed by casual evidence collected by the authors, but also by the data. This suggests that investor loyalty and own company ownership do not provide a valid alternative explanation to portfolio under-diversification or proximity investment.

To operationalize the tests of these alternatives approaches, we have first to construct proxies for the main hypotheses and controls for the other alternative explanations. This will be done in Section 4. Then, in Section 5, we will proxies and controls to show how the hypotheses laid out in this Section generate testable restrictions. Before moving there, however, we first describe the data.

3 The data

We collect data from different sources. For each investor, we have detailed information about his individual holdings of stocks (broken down at the stock level), mutual funds, bank accounts, real estate and other types of wealth. Fiscal authorities provided us with information on the different sources of investor income, as well as demographic and family characteristics. This information is matched at the individual level, so as to construct a time series of investment and income for each investor. For each stock, we have detailed information on the company and the price, volume and volatility at which it trades. We also use aggregated data on Swedish macro-economic conditions and on the indexes of the real estate market. We now explain the data sources in more detail.

3.1 Individual stockholding

We use the data on individual shareholders collected by Vardepappererscentralen (VPC), the Security Register Center. The data contain both stocks held directly and in a street name, including holdings of US-listed ADRs. In addition, SIS Ägarservice AB collects information on the ultimate owners of shares held via trusts, foreign holding companies and the like (for details, see Sundin and Sundquist, 2002). Our data cover the period 1995-2000. Overall, the records provide information about the owners of 98% of the market capitalization of publicly traded Swedish companies. For the median company, we have information about 97.9% of the equity, and in the worst case, we have information on 81.6% of market capitalization of the company.

The data provided by SIS Ägarservice AB were linked by Statistics Sweden with the LINDA dataset described below.

3.2 LINDA

LINDA (Longitudinal INdividual DAtaset for Sweden) is a register-based longitudinal data set and is a joint endeavor between the Department of Economics at Uppsala University, The National Social Insurance Board (RFV), Statistics Sweden, and the Ministries of Finance and Labor. It consists of a large, representative panel of households for the population over the period 1960 to 2000. For each year, information on all family members of the sampled individuals is added to the data set. The sampling procedure ensures that the data are representative for each year. Moreover, the same family is traced over time. This provides a real time series dimension which, in general, is lacking in surveys based on different cohorts polled over time.

The variables include individual characteristics (gender, age, marital status, country of birth, citizenship, year of immigration, place of residence detailed at the parish level, education, profession, employment status), housing information (type and size of housing, owner, rental and occupation status, one-family or several-family dwelling, year of construction, housing taxation value) and tax and wealth information. In particular, the income and wealth tax registers include information on labor income, capital gains and losses, business income and losses, pension contributions, taxes paid and taxable wealth. A detailed description of the dataset is provided by Edin and Fredriksson, (2000) and is available on the web site http://linda.nek.uu.se/.

The tax aspect deserves more detailed discussion. In Sweden, in addition to income taxation, there exists an additional wealth tax which is paid by every investor with net worth in excess of 900,000 SEK (about US$90,000). The taxable wealth includes the tax-assessed value of real estate, market value of publicly listed securities, balance of bank accounts and fair value of valuable possessions (including jewelry, cars, antiques, etc.).

For the purpose of this paper, we compute the current market value of housing using the tax-assessed value provided by LINDA. We evaluate it at current prices by using the average ratio of market value to tax-assessed value that is provided for each year and county by Statistics Sweden. There is no estimate of the market value of privately held companies. However, the data contain an indicator variable for owners of privately held companies and entrepreneurs who file their business tax return along with their personal tax return. For privately held unlimited liability companies, the value of the assets is included in the tax return. For privately held unlimited liability companies, that are not listed, the value of assets held is generally missing. However, the size of the group is rather small (1.74%-1.91% of the sample depending on a year) and is unlikely

to affect our estimates in a significant way. Moreover, for the members of the wealthiest 5,000 families, we have been able to reconstruct their values and to correctly impute it by using information from SIS Ägarservice AB (Sundin and Sundquist, 2002).

The combined LINDA/Shareholding dataset covers the period 1995-2000. The overall sample we use contains 1,757,406 observations. In addition, we also use 1990-1994 data from LINDA in the implementation of the Carroll and Samwick (1997) procedure to construct the moments of conditional non-financial income. In Table 1, we report some descriptive statistics. In particular, Panel A contains the general demographic characteristics (number of households for each year, members in household, adults in household, age of the oldest member of household, percentage of the sample with secondary and higher education, percentage of immigrants) as well the proxies for diversification (D_1, D_2, D_3 and D_4). Panels B report the characteristics of the local municipalities in terms of the main variables we will focus on (i.e., competition, specialization, concentration). Panels C and D report, respectively, the age and gender distribution of the sample and their wealth and income characteristics.

3.3 Firm-level information and other data

For individual security returns (including dividends) and the overall market index (SIX market index), we use the SIX Trust Database. For information on firm-level characteristics, we use the Market Manager Partners Databases. These two databases are the equivalent of, respectively, CRSP and COMPUSTAT for the US. In addition, Market Manager Partners Databases contain information at the plant level, including district location of the plant. The consumer confidence index is provided by Statistics Sweden. Geographical coordinates are supplied by the Swedish Postal Service and Cartesia Informationsteknik AB and contain latitude and longitude of Swedish Postal Offices.

4 Construction of variables

4.1 Measures of portfolio diversification

Finding a proper measure of portfolio diversification is a hard task as there is no unique measure for this. We therefore consider four alternative measures that capture different facets of portfolio diversification. These are derived from Goetzmann and Kumar (2002). We refer to their paper for a proper description and the rationale of their use.

Diversification is a concept that relates to the portfolio variance. The variance can be reduced either by increasing the number of shares in the portfolio (N), or by selecting the stocks so as to reduce the average correlation among the stocks in the portfolio ($\overline{Corr}$). Therefore, the simplest measure (D_1) is just the number of positions in the portfolio (i.e., $D_1 = N$). The second measure (D_2) also accounts for the degree of correlation of the stocks in the portfolio. It is constructed as:

$$D_2 = -\left[\frac{1}{N} + (1 - \frac{1}{N}) * \overline{Corr}\right]$$

where N is the number of positions and $\overline{Corr}$ is the average correlation of the stocks in the portfolio. D_2 is a normalized version of the portfolio variance multiplied by –1 and, as such, it increases when the variance is reduced. This happens either when the number of shares increases – in the limit as N goes to infinity the portfolio variance converges to the covariance among the stocks in the portfolio – or when the covariance among the stocks decreases.

Alternatively, diversification can be calculated as a measure of distance from a particular benchmark portfolio – the market portfolio. Therefore, the third and fourth measures (D_3 and D_4) are constructed as:

$$D_3 = -\sum_{i=1}^{N}\left(w_i - w_{mkt_3}\right)^2, \quad \text{and} \quad D_4 = -\sum_{i=1}^{N}\left(w_i - w_{mkt_4}\right)^2 \quad \text{for each i} = 1, ..., \text{N},$$

where w_i is the weight of the stock in the portfolio of the investor and w_{mkt} is the weight that the same stock would have in the market portfolio. D_3 and D_4 differ in the way w_{mkt} is constructed. In the case of D_3, w_{mkt3} is constructed by using the overall capitalization of the company, while in the case of D_4, w_{mkt4} is constructed by using the free float.[8] Both D_3 and D_4 are defined in such a way that, as diversification decreases, the measures become more negative. Therefore, these measures increase as diversification increases.

4.2 Measures of proximity investment

We also construct measures of *professional proximity* and *geographical proximity* to a particular stock. *Professional proximity* is a dummy taking the value 1 if the investor's profession is in the same area of activity as the company whose stock is under consideration and zero otherwise. We use the one digit SNI92 codes (similar to SIC codes) to identify the areas of

[8] This is all the more relevant in a country like Sweden where a sizable fraction of the companies' shares are in the hands of few blockholdrs.

activities. For example, for an investor working in the mining sector who holds the stock of a mining firm, the dummy would be equal to 1.

Geographical proximity is the proximity between the residence of the investor and the place where the company is located. We consider two different measures: the first one is the logarithm of the inverse of the distance between the ZIP code of the investor and the ZIP code of the closest branch/subsidiary of the company whose stock we consider. As an alternative measure, we use the logarithm of the inverse of the distance between the ZIP code of the investor and the ZIP code of the company headquarters. Given that the results do not differ, and that the variables are highly collinear, we report only the first specification. These measures are analogous to those proposed by Coval and Moskowitz (1999, 2001) in a study of geographical preferences in mutual fund investment. The greater the value of the variable, the closer the investor is to the stock. These measures are constructed at the stock level and then aggregated at the investor level, across all the stocks of his portfolio, weighting them by their share in the portfolio. This procedure delivers two measures of proximity investment for each investor and time.

4.3 Measures of professional specialization

Our unit of local district is the municipality. As we explained in Section 2, we can use the variables that the literature on urban economics has identified as the determinants of the process of professional specialization related to city agglomeration and growth. They are: *industrial specialization*, *competition* and *concentration*. In constructing them, we follow a methodology similar to Glaeser *et al*. (1992).[9] The *industry specialization* variable is the ratio of the share of the main industry in local employment to the share of this same industry in national employment. This provides a measure of industry specialization at the municipality level. In particular, it represents how specialized a municipality is in its main industry relative to what one would expect if employment in this industry were scattered randomly across the country. Both MAR and Porter predict that higher industry specialization should increase the process of knowledge creation that leads to city growth.

We also consider an alternative measure of specialization based on the degree of industrial *concentration* (or industrial underdiversification) of the municipality. The intuition behind this variable is that less diversity and higher concentration fosters focus and therefore professional specialization. This measure is the negative of the measure of diversity defined by

[9] However, while their study is conducted on industry-level data, we focus on district-level data. This implies that the definition of our variables and the analysis is econometrically different.

Glaeser *et al*. – i.e., the negative of the share of the next top (after the main one) five industries in municipal employment (1992). Municipalities with less diversity in employment will have a higher value. This variable also captures an alternative intuition. Jacobs (1969) posits that diversity is conducive to growth because "in more diversified cities, there is more interchange of ideas." Bringing together "people from different walks of life fosters transmission of ideas." This would suggest a negative relation between this variable and the rate of local agglomeration.

Competition is the number of firms per employee in a municipality relative to the number of firms per employee in Sweden. This variable represents the degree of firm competition for employees. A value greater than one means that the municipality has more firms per employee than the overall country has. As we mentioned, higher competition spurs innovation as well as allows workers to appropriate part of the benefits of professional specialization, increasing their incentive to focus.

In order to assess whether these variables do, indeed, explain growth at the local (municipality) level, we regress the latter on these measures. We find that competition, industry specialization and concentration are all positively associated with the growth rate of the municipality.[10] This confirms our hypothesis that competition, industrial specialization and concentration are among the main factors behind the process of professional specialization that drives local (municipality) growth.

It is important to note that these variables have been constructed so as to be positively related to professional specialization and negatively related to social competition. Indeed, one of the testable restrictions of the social competition hypothesis is that "in areas where there is a dominant company or sector of companies, unconstrained investors – i.e., investors who do not work for local firms – will be less diversified (and biased towards local firms)" (DeMarzo, Kaniel and Kremer, 2004). That is, the mechanism of social competition mostly affects the investors who are *not* working for the local industry and therefore need to invest in it to achieve the same consumption paths as their neighbors. Therefore, portfolio focus is *lower* for the investors who do work for the dominant firm/industry. In the case of the professional specialization hypothesis, instead, portfolio focus is *higher* for the investors who work in the dominant industry.

Our measures of professional specialization (and especially *industrial specialization*) increase with the representation of the dominant industry in the district and therefore with the probability that the investor works in such industry. As the dominance of an industry increases, social competition should decrease. In the case all the investors in a district worked for the same

[10] These results, that use the some functional specification as Table 3 in Glaeser *et al*. (1992) are available upon request from the authors

industry, the underdiversification (and proximity investment) would be maximum according to the professional specialization hypothesis and null according to the social competition hypothesis.

4.4 Measures of local prosperity

The main (first) empirical implication of the theory of social competition is a positive correlation between the degree of economic prosperity of the municipality and the degree of portfolio focus of the investors living in it. "Competition for local resources will be more intense in areas where wealth has increased substantially relative to the supply of local resources … in "boom town" economies" (DeMarzo, Kaniel and Kremer, 2004). We therefore consider four variables that measure the degree of local prosperity in a given year: the percentage of new start-ups (*start-ups*), the percentage of bankruptcies (*bankruptcies*), the employment growth (*employment growth*) and a measure of profitability of the companies operating in the municipality (i.e., *return on capital*). These variables are constructed at the firm level and aggregated at the municipality level. As an alternative measure of local prosperity, we also consider local consumption price indexes. This would be the variable required by the municipality risk hypothesis. However, this variable is very highly correlated to the other variables of local prosperity. We therefore omit it in the reported specification.

It is also worth noting that two variables of economic prosperity (i.e., number of local start-ups and bankruptcies) also capture (at least partly) the volatility of the local economy. The theory of social competition posits that portfolio focus should be higher in areas characterized by high volatility of the local sector – i.e., more "hi-tech towns than in towns dominated by less volatile industries." (DeMarzo, Kaniel and Kremer, 2004). Therefore, the social competition hypothesis suggests that the number of local start-ups should be positively related to both underdiversification and proximity investment.[11]

4.5 Measures of financial services availability and degree of isolation

The empirical implication of the theory of financial sophistication is a negative correlation between the availability of financial information and degree of financial sophistication of the investor and his degree of portfolio focus. The availability of financial services of a municipality provides a good proxy for both the availability of financial information (Garmaise and Moskowitz, 2003). In order to proxy for the availability of financial services, we include the degree of bank coverage. *Bank coverage* is constructed as follows. We first identify the list of

[11] In the case of bankruptcies, the effect is more mixed as an increase in bankruptcies both increases volatility and reduces prosperity.

credit institutions as reported by the Finance Inspection (Swedish equivalent of SEC) and then we calculate, for each municipality, the number of branches that the credit institutions have for a given year. The resulting variable is the logarithm of 1 + the number of branches per municipality.

Another way of quantifying the availability of financial information is to look at how isolated the municipality is with respect to major financial and information centers. The measures of isolation are: the *index of rural areas*, the *distance from airport* and the *population density*. The *index of rural areas* is the Urban Code as reported by Statistics Sweden. It ranges from 1 in the case of a Metropolitan area to 9 in the case of the countryside. The *distance from airport* is the distance from the closest civilian airport measured as the logarithm of the distance between airport and central post office in the municipality. *Population density* is the population (in tens of thousands) per square kilometer.

4.6 Measures of non-financial risk.

We define as non-financial risk variables those that allow us to test for portfolio choice in the presence of non-financial income risk (Heaton and Lucas, 2000a,b). We consider three measures. The first is an index of investor hedging (Γ_i). It captures the extent to which the investor's portfolio differs from the market portfolio in terms of correlation with investors' non-financial risk. It is constructed as:

$$\Gamma_i = corr(Y_i, r_m) - corr(Y_i, r_{i,port}),$$

where r_m is the return on the market portfolio, $r_{i,port}$ is the return on the financial portfolio of the ith investor and Y_i is the investor non-financial (labor and entrepreneurial) income. In the Appendix, we provide a detailed description of how the correlations between financial and non-financial income are constructed. Γ_i proxies for the change in correlation between financial and non-financial income induced by the investors' portfolio choice and quantifies the extent to which the investor deviates from a passive (holding the market portfolio) strategy. It is positive in the case of active hedging and captures the contribution of portfolio choice to the reduction of the overall investor's risk. It represents the extent to which investors actively pursue a negative correlation between financial and non-financial income that differs from the one embedded in the correlation between their non-financial income and the market portfolio. The *active hedging index* quantifies the extent to which the investor's portfolio differs from the market portfolio in terms of correlation with his non-financial risk.

The second proxy for non-financial risk is the percentage of the population in the municipality that is enrolled in welfare programs (*% of population on welfare assistance*). It

captures the risk of unemployment that comes from the fact of living in a depressed area. It is worth stressing that this is above and beyond the risk of unemployment that each individual investor perceives, and for which we construct a specific variable (unemployment risk), which we will describe later among the control variables. Finally, we also consider a measure of *debt exposure*. This is the ratio of investor debt to total assets. It is constructed at the investor level at time *t*.

4.7 Control variables

We consider the following sets of control variables: measures of income and wealth, momentum variables, demographic variables, and macroeconomic and social variables.

The *wealth variables* contain the overall level of wealth of the investor, defined as the sum of financial and real estate wealth. The *income variables* include the level and variance of labor and entrepreneurial income of the investor and the correlations between them and financial and real estate income. In order to make the results comparable with the standard literature on portfolio choice in the presence of non-financial income risk, we construct measures of the permanent (expected) non-financial income, following the approach of Carrol and Samwick (1997) and Vissing-Jørgensen (2002a, b). [12] In the Appendix, we provide a brief description of the methodology. We consider non-financial income to be labor income and entrepreneurial income. We also consider separately measures of financial gain/loss.

We also include a measure of unemployment risk that proxies for the probability of being unemployed in the following year. It is the one year-ahead forecast of a linear probability model where the unemployment status (i.e., 1 if unemployed and zero otherwise) is regressed on demographic variables, measures of income and wealth and regional, geographic and professional dummies.

The *momentum variables* include the return of the portfolio of the investor and of the market portfolio in the previous 12 months. These variables are meant to control for the possibility that the change in the degree of portfolio diversification or proximity investment is due to momentum, that is, to changes to the variation in the value of the stock market or in the value of the portfolio holdings.

[12] As an additional robustness check, we also replicated our results by using the actual levels of non-financial income, their volatilities and the correlation of financial and non-financial incomes. This replaces the measures of permanent income, volatility of income and their correlations with portfolio returns that had been constructed according to the Carrol and Samwick (1997) methodology. Given that the results are consistent, we will report only those based on the Carrol and Samwick methodology.

The *demographic variables* include: the level of education of the investor, broken down into high-school and university level, the age of the oldest member of the family of the investor and its value squared. This latter variable is consistent with standard results (Guiso and Jappelli, 2002, Vissing-Jørgensen, 2002a) which find a non-linear relation between age and the degree of stock market participation.

We also include a Stockholm and an immigration dummy. The Stockholm dummy takes the value of 1 if the investor lives in the capital and 0 otherwise. The immigration dummy takes the value 0 if all the members of the household are native Swedes, and 1 if at least one member of the household has immigrated. Furthermore, we construct a variable to proxy for the ability of the investor in his occupation. This is based on the difference between his income and the average income of his profession. The assumption is that the higher the income of the investor relative to the average income of the other investors in the same area, the higher his ability should be.

We also consider *macroeconomic and social variables*. The macroeconomic variable is the Index of Consumer Confidence. The social variables include the local tax rate and the percentage of foreign-born households. The local tax rate controls for disparities across municipalities merely due to tax treatment. The percentage of foreign-born in the population of the municipality controls for the welfare assistance provided to the immigrants. It can be particularly relevant in some areas of the country. Among these variables, we also include some variables (return on sales and number of active enterprises) that account for the local characteristics not captured by our measures of local prosperity.

Also, we use the percentage of the population that is employed by the municipality (*% of municipal employees*). The percentage of municipal employees proxies for other factors such as political attitudes, degree of local government intervention, political patronage. It also proxies for the representation and power projection of the municipality at the national level that determines the local attribution of resources.

Finally, to control for outside "background risk" (Gollier and Pratt, 1996), we also include a variable that reports the number of injured in auto accidents per km of the roads, as reported by the Ministry of Industry, Employment and Communications.

5 Testable restrictions and econometric methodology

The testable restrictions are based on the hypotheses laid out in Section 2 and the variables described in Section 4. We summarize them below. D represents portfolio

diversification and *PI* represents proximity investment. We consider only the *direct* effect of each variable on diversification and proximity investment, all else equal.

Empirical Restrictions

	Alternative Hypotheses		
Variables	***Specialization***	***Financial Sophistication***	***Social Competition***
Professional Specialization	D decreases; PI increases	-	D increases; PI decreases
Local Prosperity	-	D increases; PI decreases	D decreases; PI increases
Financial Service Availability	-	D increases; PI decreases	-
Isolation	D increases; PI decreases	D decreases; PI increases	-

Our hypothesis posits that an increase in professional specialization raises proximity investment and reduces portfolio diversification. As we mentioned before, the social competition hypothesis reaches an opposite conclusion. The proxies of local prosperity provide a nice way of separating the hypothesis of financial sophistication from that of social competition. An increase in local prosperity should increase diversification and reduce proximity investment in the case of the financial sophistication hypothesis and reduce diversification and increase proximity investment in the case of the social competition hypothesis.

The availability of financial services is positively related to financial sophistication. Therefore, an increase in the availability of financial services should increase diversification and reduce proximity investment if the financial sophistication hypothesis is correct. Isolation, on the contrary, is negatively related to financial sophistication, so it should reduce diversification and increase proximity investment according to the financial sophistication theory. However, higher isolation also implies lower professional focus and therefore would imply higher diversification and lower proximity investment according to the professional specialization hypothesis. Finally, even if we did not report it among the competing hypotheses, it is worth remembering that an increase in the non-financial risk may lower the diversification if investors want to hedge non-financial risk.

To implement these restrictions, we concentrate on the following specification:

$$Y_{it} = \alpha + \beta A_{it} + \gamma C_{it}^{Y} + \varepsilon_{it}, \qquad (1)$$

where Y_{it} is alternatively one of our measures of portfolio diversification (*D*) or proximity investment (*PI*). A_{it} contains the sets of variables we are interested in (i.e., professional specialization, local prosperity, financial services availability, isolation, non-financial risk).

C_{it}^{Y} contains all the control variables (i.e., income, wealth, momentum, demographic and macroeconomic variables).

The econometric estimation of equation (1) has to account for the selection bias due to the fact that we do not observe the investment decision of the investors who do not participate in the financial market. Given that the participation decision is endogenous, the standard estimates of equation (1) would be biased. To address this issue, we use a Heckman (1979) two-stage procedure and separately estimate the decision to enter the stock market and the portfolio choice. The decision to enter the market can be represented as:

$$P_{it} = \alpha + \beta A_{it} + \gamma C_{it}^{P} + \varepsilon_{it}\,, \qquad (2)$$

where P_{it} is a dummy that takes the value of 1 if the investor participates in the financial market and zero otherwise, A_{it} is defined as before, while C_{it}^{P} contains the vector of control variables.[13] The probability that the investor enters the financial market (P_{it}) is modeled as a normal c.d.f., defined on an expanded dataset that includes the totality of the households traced over time over each of the sample years 1995 through 2000, regardless of whether they invested in the stock market. It totals 1,757,406 households-year observations.

From the estimation of equation (2), we derive a variable (λ_{it}) that is employed in the second stage to control for the selection bias. The significance of the estimate of λ_{it} provides a test of the null of no sample selection bias. The results show that in all the specifications λ_{it} is always strongly significant, suggesting that self-selection is indeed important in the sample. We will therefore estimate:

$$Y_{it} = \alpha + \beta A_{it} + \gamma C_{it}^{Y} + \delta\lambda_{it} + \varepsilon_{it}\,, \qquad (3)$$

Some of the explanatory variables are potentially endogenous (e.g., the proxy for hedging). Moreover, the very location of the investor (and therefore his proximity to a stock) may be considered as an endogenous choice (even if taking place way before the portfolio choice

[13] $\mathbf{C_{it}^{P}}$ differs from $\mathbf{C_{it}^{Y}}$ only in terms of some variables that provide the identification restriction in the Heckman specification. In particular, $\mathbf{C_{it}^{P}}$ also contains time dummies, macro-regions, industry dummies and the correlations between non-financial income and the market portfolio. That is, the correlations between labor income and entrepreneurial labor income and the market portfolio and the correlation between the investor's real estate and the market portfolio. $\mathbf{C_{it}^{Y}}$ also contains the prior 12-month returns and volatility of the investor's portfolio, his prior capital gains and losses separately considered, and his tax rate. The main implicit assumption is that the participation decision is a function of the market portfolio (i.e., its correlations), while the portfolio decision is a function of the individual stocks.

decision). To address these issues, we use an instrumental variable estimation.[14] In order to control for heteroskedasticity, we correct the standard errors in the second stage regression, using a two stage least squares with a consistent variance-covariance matrix. We employ data disaggregated at the individual investor level.

Given that equation (2) is just an auxiliary regression only needed for the proper estimation of the second stage, but out of the scope of this paper, for brevity we will not report the results and we will focus on specification (3). A few results are, however, worth mentioning. Professional specialization always increases stock market participation. Isolation reduces stock market participation, while local prosperity – and in particular higher company profitability – reduces stock market participation. This is consistent with the fact that if the industrial area is thriving, local investors are more likely to invest in their own business, and to start their own enterprises and ventures, rather than to invest in the stock market. It is important to note that these results obtain after we control for individual wealth (both financial and real estate) and for the level and volatility of income (labor and entrepreneurial).

6 Main results

We proceed in two steps. First, we consider the determinants of portfolio diversification. This first part of the analysis, while it addresses the question of *why* investors might hold few stocks in their portfolio, however, does not explain *which* stocks they invest in. The second part of the analysis addresses the issue of stock-selection that leads to proximity investment. That is, we relate the drive to invest in professionally and geographically close stocks to the degree of professional specialization.

6.1 Determinants of portfolio under-diversification

We regress our measures of portfolio diversification on the proxies for professional specialization, isolation, local prosperity, financial services availability, non-financial risk, as well as the set of control variables (wealth, income, momentum, demographic, macroeconomic

[14] In particular, we instrument the potentially endogenous variables using as instruments a combination of strictly exogenous variables (i.e., demographic variables, industry and time dummies as well as variables proxying for the wealth and income of the family of the investor years before) and the lagged values of the main variables in the different specifications. Alternatively, we modify the estimation of the second stage of Heckman's procedure and perform a robustness check based on the estimation of a system of simultaneous equations. That is, we re-estimate equation 4 as part of a two-equation system where also the potentially more endogenous variables (proxies for hedging or location) are jointly determined. The results (not reported) do not differ from those derived from the instrumental variable estimation (reported).

and social variables). The results are reported in Tables 2-6. In particular, Table 2 contains the main specification for the four measures of portfolio diversification (respectively, for the different measures of portfolio diversification, D_1, D_2, D_3, and D_4) for the overall sample. Tables 3-6 report the results with the sample broken down into low-wealth and wealthy investors for the different measures of portfolio diversification.

A classification based on wealth is an indirect way of grouping investors on the basis of information. Indeed, a wealthy investor would be willing to spend more to purchase information on a particular stock than a less wealthy one, because the relative cost of investing in information decreases with the level of wealth (Calvet *et al*. 2000, Peress, 2002). We define as high wealth investors all the investors who, in the previous year, paid wealth tax. We define as low-wealth investors all the others. The high wealth investors represent approximately 10% of the overall sample.

For each class of investors, we consider four alternative specifications differing on the basis of the control variables that are used. We will mostly focus on the complete specification with all controls – Specification 1 in the tables. The results broadly support the professional specialization hypothesis, while they fail to support both the financial sophistication hypothesis and the social competition hypothesis and provide mixed and mostly negative evidence for the possibility that investors under-diversify in order to hedge non-financial risk.

The first finding is a significant negative correlation between portfolio diversification and our measures of professional specialization. Industry specialization, concentration and competition are all negatively related to portfolio diversification. That is, the more specialized and concentrated in a particular area the industry is, and the higher the degree of local competition for employees is, the lower the degree of portfolio diversification. This holds overall, and for the different classes of investors and is robust across all specifications for different measures of portfolio diversification. These results are striking in that they provide a direct link between urban growth and portfolio choice. They support the professional specialization hypothesis and fail to support the local competition hypothesis.

As the tables indicate, diversification decreases with the growth of competition, industry specialization and concentration. Thus, a one standard deviation increase in our measure of competition leads to a decrease in diversification of 3.4%, 3.3%, 2.3% and 3.0% for D_1, D_2, D_3 and D_4, respectively. Similarly, the effect of concentration is 3.3%, 4.1%, 3.9% and 4.7% for the corresponding measures of diversification. The effect of industry specialization is way stronger. One standard deviation increase in industry specialization leads to a 19%, 20.4%, 8% and 10.3%

decrease in diversification. These effects are stronger for high wealth investors (with an impact of competition and concentration on average around 4% and industry specialization 25%).

It is interesting to compare these findings to the results from the first stage: professional specialization increases stock market participation, but reduces portfolio diversification. This suggests that the process of professional specialization generates more – perhaps illusory – investment opportunities for the investors. The investors, attracted by these investment opportunities, rush to invest in them. This simultaneously increases stock market participation and reduces portfolio diversification.

The results regarding the measures of local prosperity are also interesting. Local prosperity is positively related to the degree of portfolio diversification. There is, indeed, a positive relation between portfolio diversification and employment growth, the number of new start-ups and the measure of profitability of local establishments (Return on Capital). An increase in bankruptcies, on the other hand, reduces diversification. These results hold overall, and for the different classes of investors, and are particularly strong for the high-wealth investors. It is also worth remembering that if we consider local consumption prices as a proxy for local prosperity, there is a strong positive correlation between prosperity and portfolio diversification.

The fact that portfolio diversification increases with the prosperity of the local area fails to support the social competition hypothesis. However, it supports the financial sophistication hypothesis. Indeed, it is likely that more financial information is available in more prosperous municipalities. For example, brokers and financial analysts may flock in, banks may increase the number of their branches and the number of financial services provided. This would reduce the search cost and therefore increase portfolio diversification.

A further element to support the financial sophistication theory is the strong, positive and statistically significant relation between portfolio diversification and availability of financial services. In other words, the higher the degree of bank coverage, the more investors are likely to diversify their portfolios. This holds across all the measures of diversification and for both classes of investors. A one standard deviation increase in the number of banking branches in a municipality leads to an average diversification increase between 3.8% and 7.4%. Moreover, the effect is stronger for high-wealth households (between 6% and 12% for high wealth investors, compared to a range between 2.4% and 5% for the low wealth investors). It is worth stressing that in these specifications, we are also controlling for the wealth and income of the investors. Therefore, these effects are "municipality effects" that act over and above the individual wealth and income effects.

Is the financial sophistication fully supported by the data? To answer this question, let us focus on the isolation variables. We find that, *all else equal*, portfolio diversification increases with isolation. That is, the further away an investor is from an urban center, the lower his connection to the rest of the world is and the less densely populated the municipality in which he lives is, the more diversified is his portfolio. In particular, one standard deviation increase in population density leads to a decrease of diversification equal to 14%, 9%, 24% and 31% for D_1, D_2, D_3 and D_4, respectively. These results are contrary to what the financial sophistication hypothesis would predict[15], while they are consistent with the professional specialization hypothesis.

Finally, what do these results tell us in terms of the hedging non-financial risk? They suggest that diversification increases at the very time when either non-financial income risk decreases (reduction in the number of bankruptcies, increase in growth and profitability) or its effects are alleviated by the existence of public services or of a financial network. This provides some evidence that underdiversification may be related to the need to hedge non-financial risk. Indeed, a financial portfolio used to hedge non-financial risk would appear to be less diversified. Areas characterized by a high employment growth rate, high start-up rate, and low bankruptcy rate, should be areas characterized by low non-financial risk, and should therefore display higher financial diversification. Also, the percentage of people on welfare assistance impacts investors differently, depending on their level of wealth. It is always negative and statistically significant for the high-wealth investors and either not significant (for D_1 and D_2) or positive for the low-wealth investors (for D_3 and D_4).

In order to further examine this issue, we need to consider the relation between portfolio diversification and our measure of active non-financial income risk hedging (Γ_i). The results are very striking. They show no correlation between hedging and portfolio diversification for the low-wealth investors and a negative correlation for the high-wealth investors. These results are very robust across different specifications and alternative measures of portfolio diversification. They suggest that the reason why low-wealth investors have an underdiversified portfolio is not related to hedging non-financial risk. For the high-wealth investors, rather, it appears that those who hedge (i.e., $\Gamma_i>0$) tend to have a less diversified portfolio. This is consistent with hedging. Indeed, high-wealth investors, more sophisticated and capable of appreciating the benefits of hedging, tilt their financial portfolio in a direction required to hedge their non-financial risk. This

[15] They would actually lend some support to standard portfolio theory. Indeed, the more isolated the investor is, the higher is the probability that he will not be able to hedge away his income or professional idiosyncratic shock and, therefore, the more diversified he should be.

shows up as a less diversified *financial* portfolio. Therefore, except for the limited subset of the high-wealth investors, the relative lack of diversification cannot be attributed to non-financial risk hedging.

It is interesting to note the role played by debt exposure. It scarcely affects the low-wealth investors, while it increases portfolio diversification for the high-wealth ones. At the aggregate level, debt exposure is positively related to D_1 and D_2 and not related to the other proxies of diversification. This apparently counterintuitive result may be explained as follows. For the low-wealth investors, debt exposure mostly affects the decision to participate in the stock market. And indeed, from the estimation of equation (2),[16] we find that the debt exposure negatively affects stock market participation for the low- wealth investors. For the high-wealth investors, on the other hand, debt exposure, while not sufficient to prevent stock market participation, induces the investors to reduce the financial exposure of their portfolio. This additional prudence increases the degree of financial diversification. Of course the causality may be reversed. Lenders may extend credit to diversified investors.

To sum up, the process of professional specialization reduces the drive to diversification. Only the professional specialization hypothesis is supported by the data, why the results fail to support the other alternative hypotheses. We now move on to directly study the portfolio choice.

6.2 Determinants of proximity investment

We regress our measures of proximity investment – i.e., professional and geographical proximity – on the proxies for professional specialization, isolation, local prosperity, financial service availability, as well as a set of control variables (wealth, income, non-financial risk, momentum, demographic, debt exposure and percentage of the population under welfare macroeconomic and social variables).

As before, we consider a specification based on the aggregated sample and one where the sample has been broken down into high- and low-wealth investors. The results are reported in Tables 7 and 8, respectively for geographical proximity and professional proximity. We consider both the entire sample ("all households") and high- and low-wealth investors separately. For each class of investors, we consider three alternative specifications differing on the basis of the control variables that are used. The results show that that the same variables that determine the degree of portfolio under-diversification also affect the proximity investment bias.

The main finding is a significant and positive correlation between both professional proximity and geographical proximity and our measures of professional specialization. If we

[16] Not reported, but available upon request.

consider the overall sample, competition, industry specialization and concentration are all positively related to both geographical and professional proximity. That is, the more the local economy is industrially specialized, concentrated in few and competitive industries, the more the local investors tend to invest in assets of companies located close to them or operating in the same area. In particular, one standard deviation increase in concentration (competition, industry specialization) results in the portfolio being 11% (16%, 3%) geographically closer to the investor and increases probability to invest into professionally close companies by 8.6% (6.5% and 3.6%).[17] If we consider the different classes of investors, we see that, while competition and concentration equally affect low- and high-wealth investors, industry specialization does not seem to affect the low-wealth investors. This relation is strong, statistically significant and holds across specifications. They support the professional specialization hypothesis and fail to support the local competition hypothesis.

It is interesting to compare these findings to those based on local isolation. If we consider the variables that proxy for isolation, we find that, all else equal, the more isolated the investors are, *the less* they tend to invest in stocks professionally and geographically close to them, even controlling for the opportunity set. For the overall sample, both professional and geographical proximity are *negatively* related to the distance from the airport and the degree of "rurality" of the municipality in which the investor lives, and *positively* related to its density of population. Living in a municipality with Index of Rural Areas one standard deviation lower than the national mean results in the portfolio being 24% (6%) geographically (professionally) closer to the investor.

If we break down the sample for different investors, we see that high-wealth investors are mostly affected by the degree of rurality and population density in the case of geographical proximity and by the distance from the airport in the case of professional proximity. The low-wealth investors, instead, are affected by the degree of rurality in the case of geographical proximity and by distance from the airport in the case of professional proximity.

If we then consider the measures of local prosperity, we find that proximity investment in general decreases with local prosperity. An increase in the return on capital of the local companies reduces the incentive to invest in geographically and professionally close stocks. Employment growth, instead, reduces the incentive to invest in geographically close stocks, while it increases the incentive to invest in professionally close stocks. Moreover, the results are quite

[17] According to our definition of geographical and professional proximity spelled out in section 4.6. For example, an increase of 100% of geographical proximity means that the investor is shifting towards a portfolio twice as much closer (in terms of miles) to where he lives. An increase of 100% of professional proximity, instead, can be interpreted as the investor shifting his portfolio towards an allocation with twice as much weight given to stocks belonging to industry in which he works.

different when we consider start-ups and bankruptcies. Investment in close stocks increases with the number of new start-ups and decreases with the number of local bankruptcies. This holds for both geographical and professional proximity.

Overall, these findings fail to support the financial sophistication hypothesis and are consistent with both the professional specialization hypothesis and the social competition hypothesis. However, if we consider the impact of the availability of financial services, we see that this always reduces proximity investment. This holds both overall and for the different classes of investors and across most specifications. These findings are consistent with the financial sophistication hypothesis. Therefore, overall there is mixed evidence for the two main alternative hypotheses – social competition and financial sophistication. The professional specialization hypothesis is instead fully supported by the data.

These results are consistent with the previous findings on portfolio under-diversification and provide a direct link between professional specialization and proximity investment. They suggest that the tension between professional specialization and prosperity also affects the exposure to the proximity investment bias. The more professionally specialized the investors are and the more numerous the local investment opportunities where they are locally involved (i.e., local entrepreneurial dimension), the more they will invest in closely related stocks. In contrast, the richer and more prosperous the area in which they live, the greater and the easier the access to information, the lower is the search cost and therefore the lower is the impact of the proximity investment bias.

Conclusion

In this paper, we address the puzzle of portfolio focus (under-diversification and tendency to invest in close stocks) from a broader perspective, linking it to theories of labor and the urban economy. In particular, we argue that portfolio focus is closely related to factors linked to professional specialization, knowledge creation and spill-over in the urban environment. We show that the processes of professional specialization that characterizes city agglomeration on the one hand reduces the availability of time to collect and analyze financial information and by the same token increases the relative information that an investor has with respect to stocks professionally or geographically closer to him. The net effect of these two factors appears to be that investors in an information-rich urban professional environment – educated though they may be – appear to diversify less.

Appendix: Construction of income-related variables

Here, we briefly describe the methodology we follow to construct proxies for permanent non-financial income, its volatility and its correlation to financial and real estate income. We follow the approach of Carroll and Samwick (1997) and Vissing-Jørgensen (2002a). We consider the total non-financial income (i.e., labor income and entrepreneurial income). In particular, we define the relevant moments of long-term investor's non-financial income:

$$E(\omega_{it}|\omega_{it-1}, X_{it-1}),\ Var(\omega_{it}|\omega_{it-1}, X_{it-1}) \text{ and } \rho_{it},$$

where ω_{it} is the non-financial income of investor *i* at time *t*, X_{it-1} are the variables that can be used to predict income next period and ρ_{it} is the conditional correlation between shocks to log non-financial income and the log stock return. We assume that non-financial income follows:

$$\ln\omega_{it} = p_{it}+\varepsilon_{it},$$

where:

$$p_{it} = g_{it}+p_{it-1}+\eta_{it},\ \varepsilon_{it}\sim N(0,\sigma_{\varepsilon t}^2),\ \eta_{it}\sim N(0,\sigma_{\eta i}^2),$$

and

$$cov(\varepsilon_{it}, \varepsilon_{is}) = 0,\ cov(\eta_{it}, \eta_{is}) = 0,\ cov(\varepsilon_{it}, \eta_{is}) = 0 \text{ for each } t, s.$$

The variable p_{it} represents the permanent income component of non-financial income. It has a drift term (g_{it}) that is known and based on the information available at *t*-1. This allows us to write:

$$\ln\omega_{it}-\ln\omega_{it-1} = p_{it}-p_{it-1}+\varepsilon_{it}-\varepsilon_{it-1} = g_{it}+ \varepsilon_{it}-\varepsilon_{it-1}+\eta_{it} ,$$

or

$$\ln\omega_{it} = \ln\omega_{it-1} + g_{it}+ \eta_{it} + \varepsilon_{it} - \varepsilon_{it-1}$$

This implies:

$$E(\omega_{it}|\omega_{it-1}, X_{it-1}) = \ln\omega_{it-1} + g_{it} = \omega_{i,t-1}\, G_{it} \exp\{J_{it} /2\}$$

$$Var(\omega_{it}|\omega_{it-1}, X_{it-1}) = (\omega_{i,t-1}\, G_{it})^2 \exp(J_{it})\ \{ \exp(J_{it}) - 1\},$$

where:

$$G_{it} = \exp(g_{it}),\ J_{it} = \sigma_{\eta i}^2+2\sigma_{\varepsilon i}^2$$

and $X_{i,t-1}$ is the set of variables usable to predict g_{it}.

In order to estimate $E(\omega_{it}|\omega_{it-1}, X_{it-1})$ and $Var(\omega_{it}|\omega_{it-1}, X_{it-1})$, we use data for the period 1990-2000, with a 5-year lagged rolling window. Following the Carrol and Samwick (1997) and Vissing-Jørgensen (2002a) methodologies, we regress $\ln\omega_{it} - \ln\omega_{it-1}$ on the set of explanatory variables $X_{i,t-1}$ and use the predicted values of such a regression as an estimate of g_{it} and the residuals as an estimate of $\eta_{it}+\varepsilon_{it}-\varepsilon_{it-1}$. The correlation between financial and non-financial income (ρ_{it}) is constructed as the conditional correlation between shocks to log non-financial income

($\eta_{it}+\varepsilon_{it}-\varepsilon_{it-1}$) and the log gross stock returns (i.e., $\ln(1+R_t)$). We use rolling five-year windows to estimate the parameters.

The set of variables contained in $X_{i,t-1}$ are: demographic variables (secondary education, higher education, age, age squared, marital status, size of the household, number of adults belonging to the household), changes in the demographic variables, industry dummies for the company the investor is working for (e.g., oil industry), dummies for the type of profession of the investor (e.g., doctor), emigration status. Following Vissing-Jørgensen (2002a), given the potential inaccuracy of estimates based on few observations, we calculate the correlation over the entire sample.

References

Arrow, K.J., 1962, "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 29(3), June, pp.155-73.

Bakshi, G. and Z.Chen, 1996, "The Spirit of Capitalism and Stock Market Prices," *American Economic Review*, 86(1), March, pp. 133-157.

Barber, B.M. and T.Odean, 2000, "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 55(2), April, pp. 773-806.

Barber, B.M. and T.Odean, 2003, "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." Working Paper, University of California, Berkeley.

Calvet, L., Gonzalez-Eiras and P. Sodini, 2000, Financial Innovation, Stock Market Participation and Asset prices, SSE/WFI Working Paper 464.

Carlson, J.A. and R.P.MacAfee, 1983, R.P., "Discrete Price Equilibrium Dispersion," *Journal of Political Economy*, 91(3), June, pp. 480-493.

Carroll, C.D. and A.A., Samwick, 1997, "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), pp. 41-71.

Carroll, C.D. and M.S. Kimball., 2001, "Liquidity Constraints and Precautionary Saving," NBER Working Paper 8496.

Coval, J.D. and T.J. Moskowitz., 1999, "Home Bias at Home: Local Equity Preference in Domestic Portfolios," *Journal of Finance*, 54(6), December, pp. 2045-2073.

Coval, J.D. and T.J. Moskowitz, 2001, "The Geography of Investment: Informed Trading and Asset Prices," *Journal of Political Economy*, 109, pp. 811-841.

DeMarzo, P.M., R.Kaniel and I.Kremer., 2002, "Diversification as a Public Good: Community Effects in Portfolio Choice," *Journal of Finance*, 59, pp. 1677-1716.

Edin, P.A., and P. Fredriksson, 2000, "LINDA: Longitudinal Individual Data for Sweden 1960-1997" Working Paper 2000:19, Department of Economics, Uppsala University.

Ellison, G. and E.L., Glaeser, 1999, "The Geographic Concentration of Industry: does Natural Advantage Explain Agglomeration?" *American Economic Review*, 89(2), May, pp. 311-16.

Frieder, L. and A. Subrahmanyam, 2004, "Brand perception and the market for common stock", *Journal of Financial and Quantitive Analysis*, forthcoming.

Garmaise, M. and T.J. Moskowitz. 2003, "More Banks Less Crime. The Real and Social Effects of Bank Competition," Working Paper, The Wharton School and University of Chicago.

Glaeser E.L., H.D. Kallal, J.A. Scheinkman and A. Shleifer, 1992, "Growth in Cities," *The Journal of Political Economy*, 100(6), December, pp. 1126-1151.

Glaeser E.L., J.A. Scheinkman and A. Shleifer, 1995, "Economic Growth in a Cross-Section of Cities," *Journal of Monetary Economics*, 36(1), August, pp. 117-143.

Glaeser, E.L., J.Kolko and A. Saiz, 2002, "Consumer City," Harvard Institute of Economic Research Working Papers.

Glaeser, E.L. and D.C. Mare, 2001, "Cities and Skills," *Journal of Labor Economics*, 19(2), April, pp. 316-42.

Glaeser, E.L. 1997, "Learning in Cities," *Journal of Urban Economics*, 46(2) September, pp. 254-277.

Goetzmann, W.N. and A. Kumar, 2002, "Equity Portfolio Diversification," National Bureau of Economic Research Working Paper # 8686.

Gollier, C. and J.W. Pratt, 1996, Risk Vulnerability and the Tempering Effect of Background Risk, *Econometrica*, 64, 1109-1123.

Grinblatt, M. and M. Keloharju, 2000, "The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set*," Journal of Financial Economics*, 55(1), January, pp. 43-67.

Grinblatt, M. and M. Keloharju, 2001, "How Distance, Language, and Culture Influence Stockholdings and Trades," *Journal of Finance*, 56(3), June, pp. 1053-1073.

Grinblatt, M., M. Keloharju, and S. Ikaheimo, 2004, "Interpersonal effects in consumption: evidence from the automobile purchases of neighbours" *NBER Working Paper 10226*.

Grinold, R.C. and R.N. Kahn, 1999, *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Selecting Superior Returns and Controlling Risk*, McGraw-Hill.

Grossman, S. J. and J. E. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70(3), pp. 393-408.

Guiso, L. and T. Jappelli, 2002, "Household Portfolios in Italy," in Luigi Guiso, Michael Halaiassos and Tullio Jappelli (eds.) *Households Portfolios*, MIT Press, Cambridge, pp. 250-289.

Heaton, J.B. and D. Lucas, 2000, "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance*, 55(3), pp. 1163-1198.

Heaton, J.B. and D. Lucas, 2000, "Portfolio Choice in the Presence of Background Risk," *Economic Journal*, 110(460), January, pp. 1-26.

Hong, H., J. D. Kubik, and J. C. Stein, 2004, "Social Interaction and Stock-Market Participation, " *Journal of Finance*, 59(1), pp. 137-163.

Heckman, J., 1979, "Sample Selection Bias as a Specification Error," *Econometrica* 47(1), January, pp. 153-161.

Hortacsu, A. and C. Syverson, 2004, "Search Costs, Product Differentation, and Welfare Effects of Entry: A Case Study of S&P 500 Index Funds, " *The Quarterly Journal of Economics,* 119(2), pp. 403-456.

Jacobs, J., 1969, *The Economy of Cities*, New York, Vintage.

Marshall, A., 1890, *Principles of Economics*, London, Macmillan.

Merton, R.C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," Journal of Economic Theory, 3(4), December, pp. 247-413.

Merton, R.C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information," Journal of Finance, 42(3), July, pp. 483-510.

Peng, L. and W. Xiong, 2002, "Time to Digest and Volatility Dynamics," Working Paper, Duke and Princeton Universities.

Peress, J., 2002, Wealth, Information Acquisition and Portfolio Choice, Mimeo.

Porter, M.E., 1990, *The Competitive Advantage of Nations*, New York, Free Press.

Polkovnichenko, V., 2003, "Household Portfolio Diversification: A Case for Rank-Dependent Preferences", Mimeo.

Romer, J, 1986, "Increasing Returns and Long-Run Growth", *Journal of Political Economy*, 94, 1002-1037.

Shapiro, A., 2002, "The Investor Recognition Hypothesis in a Dynamic General Equilibrium: Theory and Evidence," *Review of Financial Studies*, 15(1), Spring, pp. 97-141.

Shore, S.H. and J.S. White, 2002, "External Habit Formation and the Home Bias Puzzle," Working Paper, Harvard University.

Sargent, T.J., 1993, *Bounded Rationality in Macroeconomics*, Oxford University Press, Clarendon Press.

Simon, H.A., 1986, *Theories of Bounded Rationality*, University of Minnesota Press, Minneapolis.

Sims, C., 2000, "Implications of Rational Inattention," Working Paper, Princeton University.

Sundin, A. and S.I. Sundqvist, 2002, "Owners and Power in Sweden's Listed Companies," SIS Ägarservice AB, 1986-2002.

Van Nieuwerburh S. and L. Veldkamp, 2004, "Information immobility and the home bias puzzle", Mimeo.

Vissing-Jørgensen, A, 2002, "Towards an Explanation of Household Portfolio Choice Heterogeneity: Non-Financial Income and Participation Cost Structures," NBER Working Paper No. w8884.

Vissing-Jørgensen, A., 2002, "Limited Stock Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 110(4), August, pp. 825-53.

Zhu, N., 2003, The Local Bias of Individual Investors," Working Paper, Yale School of Management.

Table 1: Descriptive statistics

This table contains the descriptive statistics of the sample. Panel A reports the general demographic characteristics (number of households for each year, members of the household, adults in the household, age of the oldest member of the household, percentage of the sample with secondary and higher education, percentage of immigrants). Panel B describes the local characteristics. We use 289 municipalities for the period 1999-2000 and 288 for the period 1995-1998. We also report descriptive statistics for our measures of diversification $D_1 = N$, $D_2 = -1/N - (1-1/N)\ \overline{Corr}$, $D_3 = -\sum_{i=1}^{N}(w_i - w_{mkt})^2$, and $D_4 = -\sum_{i=1}^{N}(w_i - w_{ffl})^2$, where N is number of positions in the portfolio, $\overline{Corr}$ is the average correlation of stocks in the portfolio, w_i is the weight of the stock in the portfolio of the investor, w_{mkt} is the weight that the same stock would have in the market portfolio, and w_{ffl} is the weight that the same stock would have in the free float market portfolio. *Industry specialization* is the share of the main industry in local employment to the share of the same industry in national employment (source: Statistics Sweden). *Competition* is the number of firms per employee incorporated in a municipality relative to the number of firms per employee in Sweden (sources: Statistics Sweden, MM Partners). *Concentration* is the negative of the share of the next (after the main one) top 5 industries in local employment (source: Statistics Sweden). *Active enterprises* is the number (in thousands) of active firms in a municipality in a given year (source: MM Partners). *Start-ups* and *Bankruptcies* measure the percentage of failed (i.e., ceased business activities) and started firms in a municipality in a given year as a percentage of active firms (source: MM Partners). *Profitability* and *Return on Capital* are sales-weighted average of profitability and return on capital of the enterprises in a municipality (source: MM Partners). Employment growth is the growth of employment in a municipality w.r.t. 1985 (source: Statistics Sweden). *Index of Rural Areas* is a code set by Statistics Sweden that ranges from 1 (Metropolitan area) to 9 (countryside). *Distance from Airport* is measured as the logarithm of the distance between the closest civilian airport and the central post office in the municipality (source: Cartesia Informationsteknik AB, Swedish Civil Aviation Board). *Population Density* is the population (in tens of thousands) per square kilometer (source: Statistics Sweden). *Bank Coverage* is the logarithm of the number of credit institutions' branches (plus one) in the municipality (source: MM Partners). *Percentage of population on welfare assistance* and *percent of municipal employees* are provided by Statistics Sweden. *Index of investor hedging* is constructed as $\Gamma_i = corr(Y_i, r_m) - corr(Y_i, r_{i,port})$, where r_m is the return on the market portfolio, $r_{i,port}$ is the return on the financial portfolio of the ith investor and Y_i is the investor non-financial income. Debt Exposure is the ratio of investor' debt to investor' total assets. Panel C reports the age and gender distribution of the sample. Panel D reports the percentage of the households paying wealth tax, having labor income, having entrepreneurial income and having real estate wealth. We report mean, standard deviation, median and inter-quartile range (IQR). They have been calculated over the whole sample (i.e., across-investors and time). The column "Representation in the sample" reports the fraction of households in the sample that pay wealth tax, earn labor or entrepreneurial income or hold real estate wealth. The other columns report statistics (Mean, Median, Standard Deviation, IQR) of, respectively, the value wealth, labor and entrepreneurial income gross yearly income) and real estate. All monetary values are in Swedish krones (SEK).

Panel A: General demographic characteristics

Variable	*Mean*	*Median*	*Std.Dev.*	*IQR*
Number of households	292,901	291,913	647	686
Number of members in the household	2.67	2.00	1.51	3.00
Number of adults in the household	1.77	2.00	0.69	1.00
Age of oldest household member	49.28	47	17	24
% with secondary education	43.5%	43.5%	0.6%	0.5%
% with higher education	31.4%	31.2%	1.4%	1.4%
% of immigrants	16.4%	16.3%	2.7%	4.6%
D_1	1.63	1.00	1.78	0.00
D_2	-0.87	-1.00	0.23	0.00
D_3	-0.98	-0.88	0.60	0.61
D_4	-0.96	-0.95	0.50	0.61

Panel B: Characteristics of the local district

Variable	*Mean*	*Median*	*Std. Dev.*	*I. Q. R.*
Active Enterprises	0.795	0.239	1.947	0.406
Competition	1.934	1.874	0.466	0.708
Industry specialization	4.590	1.957	7.811	2.666
Concentration	-0.292	-0.293	0.031	0.038
Bankruptcies	0.717	0.746	1.457	0.299
Start-ups	0.501	0.512	1.427	0.263
Sale Profitability	0.103	0.091	0.154	0.040
Return on Capital	0.046	0.001	0.107	0.008
Employment Growth	-0.094	-0.097	0.118	0.133
% of Population on Welfare Asst.	0.075	0.073	0.031	0.038
% of Municipal Employees	0.065	0.065	0.011	0.015
Index of Rural Areas	3.544	3.000	2.104	3.000
Distance from Airport	8.425	8.909	1.538	2.622
Population Density	0.609	0.078	1.159	0.516
Bank Coverage	1.768	1.792	0.711	0.811
Active Hedging Index	-0.036	-0.040	0.883	1.306
Debt Exposure	0.496	0.561	0.237	0.342

Panel C: Age and gender distribution of the sample

Age	*Males*	*Females*	*Age of oldest household member*
0-19	18.2%	17.2%	0.5%
20-29	4.8%	4.9%	10.7%
30-39	7.1%	8.2%	21.7%
40-49	7.4%	7.4%	23.6%
50-59	5.9%	5.3%	17.9%
60+	6.6%	7.2%	25.8%
Total	49.9%	50.1%	100%

Panel D: Wealth and income characteristics of the households

Variable	*Representation in the sample*	*Mean*	*Median*	*Std. Dev.*	*I. Q. R.*
Wealth-Tax Payers	7.9%	359,592	102,700	2,648,521	353,400
Real Estate Holders	54.6%	449,400	387,000	348,736	340,000
Labor Income Earners	100.0%	321,489	287,722	237,526	276,190
Entrepr. Income Earners	9.8%	88,114	43,268	172,565	111,726

Table 2: Measures of diversification for overall sample

We report the results for the full specification where the dependent variables are measures of diversification $D_1 = N$, $D_2 = -1/N - (1-1/N)\ \overline{Corr}$, $D_3 = -\sum_{i=1}^{N}(w_i - w_{mkt})^2$, and $D_4 = -\sum_{i=1}^{N}(w_i - w_{ffl})^2$, where N is number of positions in the portfolio, $\overline{Corr}$ is average correlation of stocks in the portfolio, w_i is the weight of the stock in the portfolio of the investor, w_{mkt} is the weight that the same stock would have in the market portfolio, and w_{ffl} is the weight that the same stock would have in the free float market portfolio. The main variables are as described in Table 1, while the control variables are described in the text. We also control in each specification for consumer confidence and local tax rate (not reported). *t*-statistics are reported in parentheses. We also report the *Adjusted* R^2. All the coefficients are multiplied by 100.

	D_1		D_2		D_3		D_4	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
			Local Professional Specialization					
Competition	-12.80	(-14.07)	-1.64	(-13.87)	-2.84	(-9.64)	-3.22	(-13.44)
Industry specialization	-0.74	(-7.55)	-0.12	(-9.05)	-0.30	(-9.49)	-0.30	(-11.67)
Concentration	-1088.10	(-22.70)	-151.29	(-24.35)	-147.52	(-9.52)	-165.57	(-13.10)
			Degree of Local Isolation					
Population Density	-50.00	(-3.04)	-3.83	(-1.85)	-29.00	(-5.59)	-31.00	(-7.30)
Index of Rural Areas	6.58	(19.03)	0.82	(18.19)	0.28	(2.52)	0.44	(4.83)
Distance from Airport	1.34	(4.09)	0.07	(1.64)	0.32	(3.06)	0.23	(2.62)
			Local Prosperity					
Bankruptcies	-51.84	(-34.24)	-5.86	(-29.86)	-4.22	(-8.62)	-5.57	(-13.95)
Start-ups	46.33	(32.52)	5.24	(28.37)	3.64	(7.89)	4.91	(13.06)
Return on Capital	43.20	(8.60)	7.80	(11.98)	20.10	(12.34)	20.50	(15.45)
Employment Growth	93.46	(21.48)	12.60	(22.35)	8.81	(6.26)	14.09	(12.28)
			Financial Services Availability					
Bank Coverage	16.75	(20.44)	2.42	(22.77)	3.19	(12.03)	3.29	(15.23)
			Non-Financial Risk					
Active Hedging Index	-9.90	(-11.81)	-1.52	(-13.98)	-1.43	(-5.28)	-1.50	(-6.77)
Debt exposure	0.15	(3.18)	0.02	(4.06)	0.00	(-0.11)	0.00	(0.18)
% of population on welfare asst.	229.64	(11.63)	20.78	(8.12)	-46.55	(-7.29)	-41.49	(-7.97)
Wealth Variables	Yes		Yes		Yes		Yes	
Income Variables	Yes		Yes		Yes		Yes	
Momentum Variables	Yes		Yes		Yes		Yes	
Demographic Variables	Yes		Yes		Yes		Yes	
Macro and Social Variables	Yes		Yes		Yes		Yes	
Lambda	-294.33	(-39.15)	-28.92	(-29.67)	-14.90	(-6.13)	-23.64	(-11.92)
Constant	465.49	(21.81)	-73.81	(-26.68)	-97.77	(-14.16)	-89.49	(-15.90)
Adj R^2	0.070		0.064		0.115		0.178	

Table 3: Dependent variable D_1

We report the results for the specification where the dependent variable $D_1 = N$,where N is number of positions in the portfolio. We report results for Low and High wealth households. In each case, four different specifications are reported. The main variables are as described in Table 1, while the control variables are described in the text. We also control in each specification for consumer confidence and local tax rate (not reported). *T*-statistics are reported in parentheses. We also report the adjusted R^2. All the coefficients are multiplied by 100.

	Low-wealth households								***High-wealth households***							
	(1)		(2)		(3)		(4)		(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Local Knowledge Spillover																
Competition	-4.26	(-8.06)	-2.86	(-5.58)	-2.69	(-5.33)	-3.37	(-6.71)	-15.41	(-5.73)	-5.32	(-2.05)	-16.88	(-6.28)	-5.92	(-2.30)
Industry specialization	-0.30	(-5.50)	-0.28	(-5.04)	-0.26	(-4.63)	-0.26	(-4.75)	-1.09	(-3.22)	-0.74	(-2.23)	-1.12	(-3.31)	-0.48	(-1.45)
Concentration	-376.88	(-13.51)	-314.80	(-11.56)	-309.55	(-11.40)	-345.47	(-12.82)	-1718.34	(-11.26)	-918.97	(-6.45)	-1861.86	(-12.23)	-1054.27	(-7.50)
Degree of Local Isolation																
Population Density	-31.03	(-3.24)	-39.20	(-3.70)	-37.17	(-3.60)	-34.09	(-3.50)	-86.10	(-1.01)	-66.81	(-1.33)	-77.19	(-1.92)	-54.91	(-1.23)
Index of Rural Areas	1.99	(10.28)	1.37	(7.44)	1.20	(6.82)	1.48	(8.36)	5.88	(5.69)	0.99	(1.02)	7.08	(6.89)	1.00	(1.05)
Distance from Airport	0.26	(1.38)	0.08	(0.40)	-0.03	(-0.16)	0.08	(0.45)	4.81	(4.53)	0.84	(2.83)	5.71	(5.38)	0.34	(0.34)
Local Prosperity																
Bankruptcies	-13.30	(-15.12)	-7.87	(-11.41)	-6.57	(-12.32)	-9.25	(-16.38)	-78.86	(-14.89)	-25.47	(-6.84)	-90.48	(-17.57)	-22.20	(-7.12)
Start-ups	11.72	(14.24)	6.91	(10.39)	5.73	(10.64)	8.11	(14.36)	74.10	(14.56)	23.86	(6.56)	85.10	(17.18)	20.50	(6.64)
Return on Capital	24.10	(8.05)	22.20	(7.43)	21.90	(7.34)	22.70	(7.61)	57.10	(3.85)	44.70	(3.04)	59.80	(4.03)	44.70	(3.04)
Employment Growth	33.59	(12.84)	26.19	(10.61)	23.95	(9.87)	28.35	(11.80)	192.18	(12.45)	68.93	(5.44)	217.02	(14.25)	57.26	(4.82)
Financial Services Availability																
Bank Coverage	5.96	(12.39)	5.01	(10.57)	4.80	(10.26)	5.34	(11.45)	28.59	(11.02)	18.29	(7.35)	30.74	(11.87)	20.07	(8.20)
Non-financial Risk																
Active Hedging Index	0.87	(1.56)	1.13	(2.01)	0.97	(1.74)	1.10	(1.67)	-17.02	(-8.45)	-15.55	(-7.77)	-17.24	(-8.54)	-15.63	(-7.82)
Debt exposure	0.03	(1.43)	0.03	(1.16)	0.02	(1.03)	0.03	(1.28)	7.73	(7.00)	5.76	(5.32)	7.16	(6.50)	5.82	(5.54)
% of Population on Welfare Asst.	66.52	(5.87)	28.35	(2.64)	22.66	(2.16)	48.29	(4.60)	174.71	(3.04)	-110.11	(-2.03)	229.64	(4.01)	-103.55	(-1.94)
Control Variables																
Wealth Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Income Variables	Yes		No		Yes		Yes		Yes		No		Yes		Yes	
Momentum Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Demographic Variables	Yes		Yes		Yes		No		Yes		Yes		Yes		No	
Macro and Social Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Lambda	-64.01	(-14.78)	-31.39	(-11.29)	-23.33	(-27.27)	-39.19	(-29.23)	-561.85	(-17.20)	-171.09	(-9.84)	-643.99	(-20.48)	-130.32	(-15.50)
Constant	151.73	(12.24)	93.76	(9.02)	68.83	(7.65)	104.60	(11.59)	241.60	(4.68)	146.34	(2.90)	304.90	(5.97)	-33.51	(-0.73)
Adj R^2	0.040		0.035		0.035		0.035		0.053		0.048		0.052		0.048	

Table 4: Dependent variable D_2

We report the results for the specification where the dependent variable is the measure of diversification $D_2 = -1/N - (1-1/N)\ \overline{Corr}$ where N is number of positions in the portfolio and $\overline{Corr}$ is average correlation of stocks in the portfolio. The notations are as in Table 2. All the coefficients are multiplied by 100.

	Low-wealth households								High-wealth households							
	(1)		(2)		(3)		(4)		(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Local Knowledge Spillover																
Competition	-1.22	(-10.88)	-0.86	(-7.90)	-0.78	(-7.27)	-0.95	(-8.90)	-1.53	(-5.67)	-0.86	(-3.32)	-1.56	(-5.80)	-1.05	(-4.08)
Industry specialization	-0.07	(-5.55)	-0.06	(-5.12)	-0.05	(-4.39)	-0.05	(-4.37)	-0.16	(-4.75)	-0.15	(-4.50)	-0.16	(-4.77)	-0.13	(-3.84)
Concentration	-93.43	(-15.79)	-76.62	(-13.28)	-74.52	(-12.97)	-83.59	(-14.63)	-186.57	(-12.21)	-129.97	(-9.11)	-189.77	(-12.48)	-157.45	(-11.17)
Degree of Local Isolation																
Population Density	-1.86	(-0.90)	-2.76	(-1.33)	-2.88	(-1.39)	-2.38	(-1.15)	-4.54	(-1.03)	-1.84	(-0.42)	-4.48	(-1.02)	-2.12	(-0.48)
Index of Rural Areas	0.54	(13.24)	0.38	(9.84)	0.32	(8.59)	0.39	(10.40)	0.75	(7.21)	0.44	(4.51)	0.77	(7.54)	0.52	(5.48)
Distance from Airport	0.03	(0.81)	0.02	(0.41)	0.05	(1.26)	0.02	(0.56)	0.29	(2.75)	0.06	(0.59)	0.31	(2.95)	0.01	(0.08)
Local Prosperity																
Bankruptcies	-3.75	(-20.12)	-2.33	(-15.92)	-1.85	(-16.37)	-2.53	(-21.14)	-6.38	(-12.02)	-3.08	(-8.26)	-6.63	(-12.89)	-3.48	(-11.10)
Start-ups	3.32	(19.05)	2.06	(14.62)	1.63	(14.30)	2.24	(18.68)	5.97	(11.71)	2.87	(7.87)	6.21	(12.55)	3.21	(10.37)
Return on Capital	6.00	(9.46)	5.60	(8.77)	5.40	(8.51)	5.60	(8.81)	8.60	(5.78)	8.00	(5.40)	8.60	(5.82)	7.90	(5.38)
Employment Growth	8.69	(15.66)	6.64	(12.71)	5.97	(11.62)	7.09	(13.92)	17.10	(11.07)	9.19	(7.25)	17.65	(11.60)	9.85	(8.26)
Financial Services Availability																
Bank Coverage	1.63	(15.93)	1.38	(13.69)	1.30	(13.11)	1.45	(14.68)	2.87	(11.04)	2.16	(8.66)	2.92	(11.28)	2.59	(10.54)
Non-financial Risk																
Active Hedging Index	-0.07	(-0.58)	0.00	(-0.03)	-0.04	(-0.34)	0.01	(0.12)	-2.53	(-12.52)	-2.43	(-12.11)	-2.53	(-12.55)	-2.32	(-11.56)
Debt exposure	0.01	(2.95)	0.01	(2.62)	0.01	(2.43)	0.01	(2.74)	13.74	(12.42)	12.64	(11.67)	13.61	(12.37)	15.22	(14.44)
% of Population on Welfare Asst.	15.34	(6.38)	5.27	(2.32)	2.98	(1.35)	9.38	(4.22)	13.78	(2.39)	4.97	(0.92)	15.01	(2.62)	1.83	(0.34)
Control Variables																
Wealth Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Income Variables	Yes		No		Yes		Yes		Yes		No		Yes		Yes	
Momentum Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Demographic Variables	Yes		Yes		Yes		No		Yes		Yes		Yes		No	
Macro and Social Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Lambda	-16.88	(-18.38)	-8.30	(-14.08)	-5.37	(-29.67)	-9.37	(-32.98)	-40.00	(-12.23)	-16.03	(-9.21)	-41.81	(-13.31)	-16.97	(-20.11)
Constant	-81.36	(-30.95)	-97.10	(-44.09)	-104.78	(-55.04)	-96.71	(-50.57)	-83.60	(-16.17)	-88.60	(-17.55)	-82.21	(-16.11)	-111.02	(-24.16)
Adj R^2	0.034		0.030		0.031		0.030		0.040		0.037		0.040		0.036	

Table 5: Dependent variable D_3

We report the results for the specification where the dependent variable is the measure of diversification $D_3 = -\sum_{i=1}^{N}(w_i - w_{mkt})^2$, where w_i is the weight of the stock in the portfolio of the investor and w_{mkt} is the weight that the same stock would have in the market portfolio. The notations are as in Table 2. All the coefficients are multiplied by 100.

	Low-wealth households								***High-wealth households***							
	(1)		(2)		(3)		(4)		(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Local Knowledge Spillover																
Competition	-2.74	(-10.96)	-2.39	(-9.87)	-2.34	(-9.79)	-2.97	(-12.48)	-2.37	(-3.02)	-1.79	(-2.35)	-1.81	(-2.31)	-2.85	(-3.76)
Industry specialization	-0.22	(-8.40)	-0.22	(-8.59)	-0.21	(-7.88)	-0.22	(-8.60)	-0.35	(-3.58)	-0.34	(-3.46)	-0.34	(-3.43)	-0.39	(-3.96)
Concentration	-53.70	(-4.06)	-35.16	(-2.73)	-33.50	(-2.60)	-59.23	(-4.64)	-260.04	(-5.82)	-205.06	(-4.90)	-208.65	(-4.70)	-334.59	(-8.08)
Degree of Local Isolation																
Population Density	-36.00	(-7.80)	-36.00	(-7.82)	-36.00	(-7.84)	-37.00	(-7.95)	-33.00	(-2.57)	-37.00	(-2.87)	-32.00	(-2.51)	-31.00	(-2.42)
Index of Rural Areas	0.22	(2.38)	0.06	(0.71)	0.02	(0.27)	0.34	(4.02)	0.04	(0.12)	0.27	(0.96)	0.40	(1.32)	0.47	(1.69)
Distance from Airport	0.45	(4.97)	0.38	(4.24)	0.38	(4.30)	0.51	(5.74)	0.06	(0.21)	0.33	(1.11)	0.38	(1.24)	0.03	(0.11)
Local Prosperity																
Bankruptcies	-2.35	(-5.65)	-0.78	(-2.40)	-0.71	(-2.83)	-3.37	(-12.58)	-3.51	(-2.27)	-0.12	(-0.11)	0.61	(0.40)	-6.35	(-6.90)
Start-ups	2.02	(5.18)	0.62	(1.98)	0.56	(2.19)	2.91	(10.89)	2.74	(1.84)	-0.46	(-0.43)	-1.16	(-0.80)	5.41	(5.94)
Return on Capital	17.10	(12.03)	16.80	(11.88)	16.60	(11.72)	17.40	(12.28)	23.90	(5.51)	23.40	(5.42)	22.90	(5.27)	25.40	(5.86)
Employment Growth	6.64	(5.35)	3.96	(3.39)	4.18	(3.64)	9.09	(7.99)	5.03	(1.11)	3.38	(0.91)	3.72	(0.84)	12.36	(3.53)
Financial Services Availability																
Bank Coverage	2.04	(8.93)	1.78	(7.91)	1.70	(7.66)	2.11	(9.53)	4.41	(5.80)	3.70	(5.06)	3.62	(4.78)	5.61	(7.77)
Non-financial Risk																
Active Hedging Index	-0.34	(-1.30)	-0.31	(-1.16)	-0.32	(-1.20)	-0.31	(-1.17)	-2.08	(-3.53)	-2.00	(-3.39)	-2.00	(-3.38)	-1.95	(-3.30)
Debt exposure	0.01	(0.74)	0.01	(0.89)	0.01	(0.96)	0.01	(0.55)	14.56	(4.50)	12.82	(4.03)	17.10	(5.32)	20.57	(6.64)
% of Population on Welfare Asst.	40.39	(7.52)	51.34	(10.09)	52.37	(10.55)	32.38	(6.52)	-50.20	(-2.98)	-67.84	(-4.26)	-70.30	(-4.20)	-38.39	(-2.43)
Control Variables																
Wealth Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Income Variables	Yes		No		Yes		Yes		Yes		No		Yes		Yes	
Momentum Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Demographic Variables	Yes		Yes		Yes		No		Yes		Yes		Yes		No	
Macro and Social Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Lambda	-9.17	(-4.47)	-0.56	(-2.43)	-0.64	(-3.57)	-16.18	(-25.47)	-3.72	(-2.39)	21.43	(4.20)	25.41	(2.77)	-22.70	(-9.16)
Constant	-93.98	(-16.00)	-114.73	(-23.30)	-114.57	(-26.89)	-86.07	(-20.14)	-60.29	(-3.99)	-69.61	(-4.70)	-84.59	(-5.67)	-133.33	(-9.88)
Adj R^2	0.227		0.226		0.226		0.225		0.044		0.043		0.040		0.040	

Table 6: Dependent variable D_4

We report the results for the specification where the dependent variable is the measure of diversification $D_4 = -\sum_{i=1}^{N} (w_i - w_{ffl})^2$, where w_i is the weight of the stock in the portfolio of the investor and w_{ffl} is the weight that the same stock would have in the free float portfolio. The notations are as in Table 2. All the coefficients are multiplied by 100.

	Low-wealth households								***High-wealth households***							
	(1)		(2)		(3)		(4)		(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Local Knowledge Spillover																
Competition	-2.81	(-13.96)	-2.45	(-12.59)	-2.34	(-12.20)	-2.93	(-15.30)	-3.45	(-5.42)	-2.63	(-4.27)	-3.15	(-4.96)	-3.40	(-5.54)
Industry specialization	-0.22	(-10.60)	-0.23	(-10.80)	-0.21	(-9.90)	-0.23	(-10.73)	-0.36	(-4.49)	-0.34	(-4.30)	-0.35	(-4.42)	-0.36	(-4.55)
Concentration	-71.26	(-6.71)	-53.45	(-5.16)	-49.86	(-4.83)	-74.68	(-7.28)	-265.74	(-7.36)	-189.81	(-5.61)	-233.52	(-6.50)	-287.59	(-8.60)
Degree of Local Isolation																
Population Density	-34.00	(-9.02)	-33.00	(-8.98)	-33.00	(-8.89)	-34.00	(-9.04)	-39.00	(-3.70)	-43.00	(-4.19)	-37.00	(-3.59)	-38.00	(-3.65)
Index of Rural Areas	0.29	(3.99)	0.13	(1.91)	0.06	(0.92)	0.36	(5.29)	0.31	(1.28)	0.11	(0.47)	0.04	(0.17)	0.40	(1.78)
Distance from Airport	0.31	(4.36)	0.24	(3.35)	0.23	(3.27)	0.35	(4.94)	0.09	(0.36)	0.26	(1.09)	0.11	(0.43)	0.01	(0.04)
Local Prosperity																
Bankruptcies	-3.14	(-9.39)	-1.53	(-5.85)	-1.17	(-5.79)	-3.69	(-17.19)	-6.77	(-5.40)	-2.16	(-2.44)	-4.15	(-3.41)	-6.52	(-8.77)
Start-ups	2.70	(8.62)	1.27	(5.03)	0.95	(4.63)	3.19	(14.83)	6.14	(5.09)	1.79	(2.07)	3.66	(3.13)	5.90	(8.02)
Return on Capital	16.80	(14.73)	16.50	(14.51)	16.20	(14.22)	16.90	(14.89)	24.10	(6.87)	23.40	(6.70)	23.50	(6.71)	24.70	(7.04)
Employment Growth	9.76	(9.80)	7.10	(7.57)	6.89	(7.47)	11.20	(12.26)	16.15	(4.42)	4.79	(1.59)	10.48	(2.92)	16.23	(5.74)
Financial Services Availability																
Bank Coverage	2.18	(11.88)	1.92	(10.63)	1.81	(10.17)	2.20	(12.43)	4.05	(6.59)	3.08	(5.21)	3.58	(5.86)	4.44	(7.61)
Non-financial Risk																
Active Hedging Index	0.02	(0.08)	0.05	(0.22)	0.05	(0.22)	0.01	(0.03)	-2.24	(-4.71)	-2.12	(-4.47)	-2.19	(-4.60)	-2.25	(-4.73)
Debt exposure	0.01	(0.76)	0.01	(0.96)	0.01	(1.08)	0.01	(0.66)	13.27	(5.07)	11.07	(4.31)	14.15	(5.45)	14.01	(5.60)
% of Population on Welfare Asst.	39.74	(9.22)	50.77	(12.43)	53.14	(13.34)	34.91	(8.75)	-38.50	(-2.83)	-62.95	(-4.89)	-50.53	(-3.74)	-40.04	(-3.14)
Control Variables																
Wealth Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Income Variables	Yes		No		Yes		Yes		Yes		No		Yes		Yes	
Momentum Variables	Yes		Yes		No		Yes		Yes		Yes		No		Yes	
Demographic Variables	Yes		Yes		Yes		No		Yes		Yes		Yes		No	
Macro and Social Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Lambda	-15.24	(-9.25)	-5.23	(-4.94)	-3.40	(-10.43)	-19.16	(-37.57)	-25.80	(-3.34)	8.22	(1.99)	-7.26	(-0.98)	-22.85	(-11.41)
Constant	-87.10	(-18.47)	-108.31	(-27.39)	-111.50	(-32.59)	-83.42	(-24.31)	-60.87	(-4.98)	-72.35	(-6.04)	-73.65	(-6.11)	-119.78	(-10.99)
Adj R^2	0.314		0.313		0.313		0.312		0.067		0.066		0.066		0.064	

Table 7: Dependent variable Geographical Proximity

We report the results for the specification where the dependent variable is our measure of geographical proximity. It is measured as value-weighted average of inverse of the logarithm of the distance between the ZIP code of the investor and the ZIP code of the closest branch/subsidiary of the company whose stock we consider. The notations are as in Table 2. Coefficients for Industry specialization are multiplied by 100.

	All households		***Low-wealth households***						***High-wealth households***					
			(1)		(2)		(3)		(1)		(2)		(3)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Local Knowledge Spillover														
Competition	0.138	(4.54)	0.052	(1.98)	0.040	(1.54)	0.023	(0.91)	0.203	(3.54)	0.102	(1.91)	0.029	(0.60)
Industry specialization	0.150	(2.43)	-0.021	(-0.35)	-0.036	(-0.60)	-0.054	(-0.87)	0.245	(2.78)	0.274	(2.09)	0.385	(2.97)
Concentration	1.495	(7.33)	0.490	(2.51)	0.556	(2.85)	0.943	(4.87)	2.567	(6.27)	2.570	(6.64)	2.871	(7.56)
Degree of Local Isolation														
Population Density	2.540	(3.53)	0.520	(0.79)	0.460	(0.70)	-0.600	(-1.01)	3.850	(3.59)	3.610	(3.57)	4.050	(4.35)
Index of Rural Areas	-0.044	(-11.00)	-0.014	(-3.48)	-0.010	(-2.66)	-0.007	(-2.12)	-0.066	(-8.25)	-0.045	(-6.22)	-0.035	(-5.06)
Distance from Airport	-0.014	(-1.50)	0.004	(0.67)	0.003	(0.46)	0.000	(0.03)	0.009	(0.95)	0.005	(0.54)	0.003	(0.35)
Local Prosperity														
Bankruptcies	-0.820	(-5.36)	-0.360	(-2.43)	-0.463	(-3.02)	-0.815	(-5.55)	-0.540	(-1.91)	-1.279	(-4.84)	-1.812	(-7.42)
Start-ups	0.748	(4.60)	0.334	(2.09)	0.473	(3.30)	0.864	(5.50)	0.463	(2.60)	1.338	(5.19)	1.881	(7.95)
Return on Capital	-0.004	(-2.63)	-0.000	(-0.31)	-0.000	(-0.23)	-0.001	(-0.96)	-0.010	(-3.69)	-0.010	(-3.87)	-0.012	(-5.43)
Employment Growth	-0.181	(-2.63)	-0.156	(-2.80)	-0.179	(-3.24)	-0.212	(-3.77)	-0.435	(-3.25)	-0.325	(-2.61)	-0.274	(-2.44)
Financial Services Availability														
Bank Coverage	-0.065	(-5.55)	-0.014	(-1.29)	-0.019	(-1.83)	-0.044	(-4.11)	-0.177	(-7.58)	-0.184	(-8.44)	-0.200	(-10.32)
Control Variables														
Wealth Variables	Yes		Yes		No		Yes		Yes		No		Yes	
Income Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Momentum Variables	Yes		Yes		No		Yes		Yes		No		Yes	
Demographic Variables	Yes		Yes		Yes		No		Yes		Yes		No	
Macro and Social Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Non-financial risk	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Lambda	1.273	(11.48)	0.205	(2.05)	-0.089	(-5.61)	-0.723	(-23.71)	2.295	(9.87)	0.526	(10.51)	-0.534	(-7.36)
Constant	-9.581	(-13.49)	-6.921	(-23.09)	-6.184	(-41.49)	-4.026	(-27.94)	-8.522	(-8.84)	-4.204	(-6.77)	-2.391	(-4.96)
Adj R^2	0.129		0.116		0.116		0.1131		0.111		0.110		0.108	

Table 8: Dependent variable Professional Proximity

We report the results for the specification where the dependent variable is our measure of professional proximity. For each position, it takes value 1 if any of the household members are employed in the same industry of the company they are investing in. For the investor, the measure is computed as value-weighted average over all position in household portfolio. The notations are as in Table 2. Coefficients for Industry specialization are multiplied by 100.

	All households		***Low-wealth households***						***High-wealth households***					
			(1)		(2)		(3)		(1)		(2)		(3)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Local Knowledge Spillover														
Competition	0.007	(3.03)	0.006	(4.09)	0.006	(4.19)	0.009	(5.60)	0.001	(0.15)	0.003	(0.43)	-0.017	(-1.04)
Industry specialization	0.023	(5.21)	0.007	(2.51)	0.008	(2.50)	0.005	(2.35)	0.946	(9.08)	0.939	(9.18)	0.713	(10.22)
Concentration	0.139	(9.61)	0.043	(4.50)	0.051	(5.16)	0.009	(2.78)	0.898	(4.26)	0.976	(4.90)	2.122	(15.93)
Degree of Local Isolation														
Population Density	0.280	(0.69)	0.002	(0.04)	-0.009	(-0.24)	0.041	(1.16)	-0.073	(-0.49)	-0.071	(-0.48)	0.120	(1.15)
Index of Rural Areas	-0.001	(-3.41)	-0.001	(-2.81)	-0.001	(-2.99)	-0.001	(-5.05)	0.001	(0.63)	0.000	(0.17)	-0.003	(-3.71)
Distance from Airport	-0.002	(-5.66)	-0.002	(-6.30)	-0.002	(-6.39)	-0.002	(-5.36)	-0.003	(-2.59)	-0.003	(-2.59)	0.000	(-0.42)
Local Prosperity														
Bankruptcies	-0.041	(-3.88)	-0.057	(-7.77)	-0.052	(-6.88)	-0.078	(-9.23)	-0.313	(-8.79)	-0.312	(-9.03)	-0.244	(-10.77)
Start-ups	0.047	(4.16)	0.063	(7.97)	0.057	(6.96)	0.088	(9.67)	0.264	(8.02)	0.264	(8.27)	0.260	(10.52)
Return on Capital	-0.001	(-7.97)	0.000	(-3.51)	0.000	(-3.32)	0.000	(-3.37)	-0.002	(-5.60)	-0.002	(-5.55)	-0.001	(-2.14)
Employment Growth	0.022	(4.54)	0.013	(4.50)	0.017	(5.64)	0.007	(2.06)	-0.117	(-7.13)	-0.118	(-7.36)	-0.124	(-11.29)
Financial Services Availability														
Bank Coverage	-0.006	(-7.77)	-0.002	(-3.84)	-0.003	(-5.03)	-0.001	(-0.88)	-0.059	(-11.98)	-0.058	(-12.15)	-0.043	(-12.98)
Control Variables														
Wealth Variables	Yes		Yes		No		Yes		Yes		No		Yes	
Income Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Momentum Variables	Yes		Yes		No		Yes		Yes		No		Yes	
Demographic Variables	Yes		Yes		Yes		No		Yes		Yes		No	
Macro and Social Variables	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Non-financial risk	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Lambda	0.024	(3.07)	0.025	(4.88)	-0.009	(-9.46)	0.010	(4.67)	-0.048	(-2.24)	-0.024	(-4.72)	0.021	(4.35)
Constant	0.081	(1.61)	-0.026	(-0.84)	0.130	(5.98)	0.184	(6.84)	0.537	(5.78)	0.494	(6.08)	0.712	(13.35)
Adj R^2	0.011		0.007		0.007		0.004		0.021		0.021		0.011	