Investor Activity and Returns in the Swedish Premium Pension

Jasmine Michelsson (23656)* Josefin Klarin (23490)†

May 14, 2018

BSc Thesis in Finance at the Stockholm School of Economics

ABSTRACT

By comparing individuals from a large sample of pension savers in the Swedish Premium Pension System, this study investigates what effect activity has on absolute returns within the programme. Further, we examine if there are differences in returns and activity across groups of different characteristics. We find that activity has a negative effect on the investor returns. The most profitable group is found to be the one with completely inactive investors, with a 100 percent allocation of their capital in the default fund. Moreover, we find a negative effect of using premium pension advisors. Finally, the results imply that income is positively related to returns while age, on the other hand, is negatively correlated to it and that returns to male versus female investors does not differ significantly.

Keywords: premium pension, investor activity, investor characteristics, defined contribution

Acknowledgements: We would like to thank our tutor Research Fellow Håkan Thorsell for valuable input and the Swedish Pensions Agency for insight and guidance. Especially we would like to thank Bengt Norrby at the Swedish Pensions Agency for providing us with data.

*23656@student.hhs.se

†23490@student.hhs.se

Table of Contents

1 Introduction

In 2000 the Premium Pension System was introduced in Sweden in order to enable investors to be more autonomous in allocation decisions in a small part of their pension portfolios. Large campaigns encouraging savers to take an active role in their pension savings were rolled out at the introduction of the programme, but as time has passed young people becoming enrolled in the programme are exhibiting a more passive stance. This inactive behaviour entails more funds being allocated to the state default fund AP7 Såfa, which forms an essential tool to impact national pension. At the launch in 2000, 67 percent of all individuals made an active allocation choice while since then only 2-3 percent of all new savers make an active allocation decision within their first year in the programme[1]. Also, out of those who made an active choice at the launch of the programme 40 percent have remained with their initial choices and have thus been inactive since which can be an expected consequence of the massive campaigns rolled out in parallel with the introduction of the programme but which have been scaled down since then. Overall 50 percent of the savers with a corresponding 70 percent of the capital have made active choices which leave 30 percent of the capital in the default fund[2], which is a number expected to grow for each year if the trend continues. The system as it looks today has been criticised since it assumes that investors are rational and will actively engage to allocate their funds to mirror their personal situation and preferences, which has been rebutted by the high fraction of inactive investors (Pilot study S2017/03516/SF). With 18 years since the introduction of the Premium Pension System, there has been much public discussion regarding the programme, but there is yet a shortfall of evaluation regarding the returns to the savers.

[1] Presentation "Premiepension höst 2017", Swedish Pensions Agency
[2] Presentation "Premiepension höst 2017", Swedish Pensions Agency

This study aims to examine the relationship between absolute returns and investor activity within a defined contribution scheme, using data covering a large sample of Swedish pension savers in the Premium Pension System from the time of the introduction of the plan in the year 2000 up to present day. By comparing groups with different levels of investor activity and their respective returns from a sample of 96 918 investors in the Swedish Premium Pension System, we illustrate the relative effectiveness of choosing to stay in the default fund compared to either using a financial advisor or choosing to make active investment choices which entail alternative allocations. Further, we examine whether the return and activity are different for investors with different characteristics with regards to age, income and gender. This boils down to our research questions:

Does investor activity affect absolute returns? Are there differences in investor activity and absolute returns across individual investor characteristics?

We find that activity has a negative effect on returns. Inactive investors staying with all of their capital in the default allocation, which 11.6 percent of investors in our sample have done, are on average seeing 2.38 percent higher absolute returns than investors choosing to manage their funds actively. Activity has the most substantial negative effect for investors in the middle segment of activity, making changes to their allocation on average between once every two years down to every ninth month. They have seen on average 1.54 percent lower returns than other investors. Using an independent financial advisor has a small negative effect on returns with those choosing to do so on average getting 0.25 percent lower returns. We have estimated that 12 percent of our sample use a financial advisor. While age is negatively correlated to return, income correlates

positively to it. We find that whether the investor is female or male has no significant effect on the return, while male investors seem to be slightly more active.

This study contributes to existing literature about investor activity in different defined contribution pension solutions by focusing on Sweden in particular. While several countries have similar solutions, Sweden distinguishes itself by the high number of funds available for pension savers to choose from (Study Ds 2013:35, 2013), which might be a contributing factor to the low levels of activity. While most literature on the area focuses on risk-adjusted returns, we have chosen to direct our focus to absolute returns since they comprise the actual return of the programme to savers, and not focus on whether excess return comes from taking more risk or making better investment decisions. With regards to the Swedish Premium Pension System, no research has been done recently, to our knowledge, and thus covering the same exhaustible sample period.

We are looking to evaluate if returns to activity are positive or negative in order to conclude how the default alternative stands in contrast to the other funds available in the system. We also want to see whether investors with certain characteristics have shown different historical activity and returns in the programme. Government investigations aimed at evaluating the Premium Pension System have pointed to the passiveness of savers as a problem (2013)(Lundbergh, 2017) but if the default alternative has been beneficial for investors the current design of the programme might not be as problematic, since inactive savers can feel confident that the default alternative maintains high standards while simultaneously allowing those who want to be active a wide array of funds to choose from.

2 The Swedish Pension System

The Swedish national retirement pension system consists of two parts which are the income pension and the premium pension. 18.5 percent of the yearly salary and other taxable benefits make out for the yearly contribution to the national pension system. There is an income ceiling which restricts the amount of money that is set aside for you of 7.5 income base amounts, which amount varies each year[3]. The cap is compensated in the occupational pension which is progressive and lies outside the national retirement pension, paid by employers.

The national retirement pension is divided into two parts. 16 percentage points are allocated to the income pension which is a system on a pay-as-you-go basis that follows earnings trends in Sweden where the government is responsible for capital management. The remaining 2.5 percentage points go to the Premium Pension System which is a version of a defined contribution plan that was designed to enhance the return on pension savings by allocating a small portion of the individual's pension funds in the capital market. It is a fully funded individual account where pension savers can allocate money according to their personal preferences.

The focus of this paper is on the Premium Pension System. As of 31 December 2017 there were 7 378 231 persons in the Premium Pension System, with an average portfolio worth of 154 218 SEK[4]. The programme was introduced in 2000 and for three reasons. The first one is that people are able to make riskier investments in this part than within the income pension, and thus generate higher returns on their funds. Secondly, the programme allows for further diversification since the income pension is mainly affected by the income development in Sweden, and thus puts a lot of dependence on the state of the Swedish economy, and one can invest in foreign assets in

[3] Which amounts to 468 759 SEK in year 2018 and corresponds to an income of 39 963 SEK/month
[4] Statistics from the Swedish Pensions Agency

the Premium Pension System. The last reason is that it gives the pension savers the opportunity to affect the allocation as one prefers. The pension savers can in this part of the system allocate funds as they wish, choosing from more than 800 funds approved by Finansinspektionen, Sweden's financial supervisory authority.

If one does not actively choose funds, the default allocation is the AP7 Såfa fund. It is managed by the state pension fund AP7 with the objective to achieve higher returns than the average of the privately managed funds offered in the system. It is designed to complement the income pension by exposure to global equity markets. Since May 2010 it is a life cycle fund which means that it adjusts the allocation between the AP7 Equity fund and the AP7 Fixed Income fund based on the age of the pension saver, in order to create the appropriate risk profile. Until the age of 55, there is a 100 percent allocation to the AP7 Equity fund which then from the age of 56 increases the share in the AP7 Fixed Income fund with 3-4 percent each year. At the age of 75, the allocation change stops with two-thirds of fixed income and one third in equities. The default option is designed to complement the income pension which depends heavily on the state of the Swedish economy and thus allocates to funds globally. In order to increase the equity exposure, the AP7 Equity fund invests in derivative contracts. The amount of leverage can be altered by the AP7 board to fit current market conditions. While the previous normal state of leverage was 35 percent, the fund has recently stated that it will be lowering the risk level and thus the new normal state is 25 percent leverage. Due to current high valuations in the equity market the fund has adjusted the leverage down to 15 percent[5]. The ultimate goal of the new strategy, which also includes investing part of the capital in private equity and emerging market equities, is to ensure a higher lowest level by reducing the risk of the fund through diversification (ibid).

[5] The Seventh AP Fund

Within the Premium Pension System, the saver can allocate funds as one prefers between up to five funds. There are no direct costs associated with changing funds since the Swedish Pensions Agency has negotiated with all the funds not to charge buying or selling fees. Additionally, heavy discounts on fund management fees have been negotiated by the authority which means that on average the funds return two-thirds of the management fees and thus effectively only charge one third within the system compared to their fees in the retail market outside the system. Since the premium pension is only a smaller part of the portfolio of an individual, where one cheaply can access high-risk funds, the expectation would be that rational investors choose to allocate this part of their funds into more risky assets.

There are private premium pension advisors who people can hire who will manage the portfolio within the Premium Pension System on behalf of the customers. A fee is paid to the financial advisors outside of the system in addition to the management fee which is in line with retail prices of asset management[6]. It was usual that the financial advisors made coordinated changes among all or large parts of their customers. This simultaneous activity meant that a lot of people were allocating their funds identically at the same point in time. These coordinated activities increased transaction costs for all investors, including the ones who were not a part of the coordinated changes, to the extent that they were prohibited starting from the 1st of December 2011. What many advisors did instead was to create their own funds to which clients could allocate their money within the system.

[6] "Fondrådgivare och förvaltningstjänster", The Swedish Pensions Agency, 2017

3 Literature Review and Theoretical Framework

3.1 Portfolio theory

Modern portfolio theory (MPT) is based on the idea that investors will maximise expected returns given a certain level of risk, and not only the expected returns (Markowitz, 1952). The model assumes that investors are risk-averse, which means that for a given level of return an investor would prefer a less risky portfolio to a riskier portfolio. This requisite implies that an investor should hold a diversified portfolio. The idea is that in a perfect market, volatility has a positive linear relationship with returns. The implication is that in order to achieve higher expected returns, additional risk must be taken on. Despite investors being aware of the diversification benefits, studies have made it evident that investors remain under-diversified as a consequence of not properly taking into account the correlation between the assets they hold (Goetzmann and Kumar, 2001).

Although portfolio theory is very straightforward, theoretically there are many instances where investors do not behave rationally as modern financial economics assume, which leads to investors deviating from what theoretic models predict. Moreover, while keeping these theories in mind, it is important to note that we only have information about the investor's partial portfolio in our sample which is what they have in the Premium Pension System. It is a reasonable assumption to make that only a fraction of an investors total funds are allocated within the system. Thus, we cannot apply these theories unquestioningly to our data and do not necessarily expect to see rational investment behaviour within our sample. However, we believe that they are relevant to mention since they can help us understand and explain some of our findings and serve as a useful reference point.

Since there are discounts on fund management fees in the premium pension system, a rational investor will allocate to the most expensive funds in the system, which implies that we should see an unproportionally large share of equity investments in the premium pension since fixed income investments are less expensive. Cheap access to high-risk investments should imply savers to be prone to invest in riskier assets in accordance with MPT. Expected returns are therefore higher within the Premium Pension System than the overall portfolio of the corresponding investor.

Several studies have shown women to be more risk-averse than men regarding their entire portfolio of assets. In general, women invest less of their capital in risky assets. Of their pension assets, the share invested in bonds relative to stocks tend to be larger compared to those of men (Jianakoplos and Bernasek, 1998). Women, being less prone to risk-taking, are consequently expected on average to generate lower returns, since a lower level of risk is associated with lower expected return on investment (Bernasek and Shwiff, 2016). Stronger risk-aversion for long-term investors may thus result in substantially lower accumulation of retirement income for women (Hira and Loibl, 2008). Despite PPM being only a fraction of the individual portfolio, if women are generally more inclined to invest in less risky assets, it might reflect on their behaviour in the Premium Pension System as well. The anticipated result will plausibly be a negative relation between return for women versus that of men.

Some additional explanations for differing behaviour include the level of earnings, financial knowledge, comfort with math or the size of retirement benefits. Additional causes for dissimilar practices may be access to information and the ability or inclination to use available information (Bajtelsmit and Bernasek, 1996). According to traditional life-cycle investment theories, average risk aversion is constant with age, implying that an appropriate portfolio consists

of a growing share in bonds and decreasing the share in equities as the investment period shortens when the individual grows older (Hana and Wang, 1997). These findings contrast those of Morin and Suarez (1983), which instead imply that risk aversion increases with age. It is worthy of notice that high net worth individuals are found to experience a decrease in risk aversion with age. It has also been argued that risk-aversion decreases as wealth rises (Arrow, 1971). Thus both net worth and age may affect the individuals risk tolerance. In contrast to the statement above, Hana and Wang conclude that risk aversion, when seen as the relative share of net worth in stocks, in general decreases as people age all else equal. Thus, tolerance of risk in fact increases with age and the life-cycle theories are rejected. Instead, it is advocated to focus less on subjective opinion on risk and more on objective matters, as the investment horizon. Research into the relationship between risk aversion and age provide ambiguous conclusions. What correlation to expect between age and return is thus unclear. Since research has shown risk-aversion to decrease with age wealthier investors should have higher expected returns. We will use the labour income as a proxy for wealth and thus the hypothesis is that income will be positively correlated with returns.

3.2 Current literature

Defined contribution plans for pension savings are often opposed due to considerable responsibility being transferred to the individual in contrast to the alternative of defined benefit plans where the administration of capital is instead exercised by the plan sponsor. Within defined contribution plans, great financial risk lands upon the individual as the final pension depends on the administrative and financial skills of the individual. Given a generally low engagement with the management of funds, this type of pension program might not provide in the best returns to savers (Battocchio and Menoncin. 2003).

A paper which subject is highly related to ours is "Individual Investor Activity and performance" by Dahlquist et al. (2016) which concludes that due to better investment decisions, mainly assigned to superior timing of asset classes and funds, active investors earn both higher absolute returns and higher risk-adjusted returns. This study is conducted using data from the Swedish Pensions Agency on the Premium Pension System regarding the first ten years since the implementation of the scheme, years 2000 to 2010. A large focus of the paper is that trading costs generated by a small portion of very active investors burden the whole system when trading in mutual funds and not individual stocks and that this tendency is exacerbated by financial advisors who used to make large, coordinated fund changes before December 2011 when it was banned. During those initial years, the advisors were responsible for an estimated 80 percent of fund changes made according to the study.

Graham, Harvey and Huang (2009) find that investors who consider themselves more competent in making investments are more likely to trade more frequently on average. Excessive trading in stocks hurts net returns due to higher transaction costs (Barber and Odean, 2000). The study showed that there was no difference in gross returns between households trading frequently and who was inactive, but that active households saw lower net returns. This result is consistent with the hypothesis that trading derives from overconfidence but rejects the hypothesis that investors are rational and will trade when the marginal benefit of doing so is equal to or higher than the cost of doing so (2000). Further, it is found that men are generally overconfident relative to women and thus trade more excessively than a rational investor which hurt their returns (Barber and Odean, 2001). These studies are made on common stock investments which entail a commission fee when trading. Since there are no direct costs associated with trading within PPM the hypothesis based on this literature is that there should be no difference in returns between

investors with different activity. On the other hand, since Dahlquist et al. (2016) found that highly active investors earn higher returns the previous literature on investor activity and returns is ambiguous, and thus we do not know what result should be expected from our study.

4 Data

We have obtained data from The Swedish Pensions Agency, complemented with information from Statistics Sweden (SCB), on a random sample of 96 918 individuals in the Premium Pension System. The sample period ranges from the start of the programme, which was in September 2000, until January 2018. The data presents initial fund choices of each individual and following changes of that allocation on a daily basis. The data also includes information regarding gender, age, income, municipality and return on invested capital for all individuals. We do not have information about how much capital is in the accounts. With respect to the individual accounts continuously receiving new capital, we use the capital weighted return, which is also called the internal rate of return (IRR), as the dependent variable[7].

$$\sum_{t=0}^{T-1} C_t * (1+r)^{\frac{T-t}{365}} = C_T$$

It is the measurement the pension authority uses to compare return rates across investors (Berglöf and Birkholz, 2013). The equation looks like above and solving for r gives the internal rate of return. T is the number of days since the start, C_t is cash-flow at time t, and C_T is the current account/portfolio value.

Since the IRR is capital weighted based on the available capital in the accounts, the return for an individual will be affected by withdrawals and thus not fully comparable to others.

[7] From here on "internal rate of return", "absolute return" and "return" will be used interchangeably

Therefore, those individuals who have begun withdrawing funds from their premium pension accounts are excluded from the sample.

We differentiate between allocation changes made actively by the individual and by a financial advisor on behalf of the individual in the period from the start until December 2011. Since there has been no feasible way of gathering information on which individuals have hired an advisor we have created a dummy variable for all investors who have made more than ten changes in the period ranging from the start until the coordinated changes were banned and who have made no changes since then. This estimation will be used as a proxy for those who had a financial advisor before the period and have decided not to start to manage the funds actively themselves since December 2011 either, since they have remained in the same fund since before then. Using this strategy produces a category for the use of advisory including 8 785 individuals, representing 12 percent of the sample. In the study by Dahlquist et al. a similar but not identical method was used to create a proxy for the use of advisors, which accounted for 11.5 percent of their sample individuals (Dahlquist et al. 2016). Thus we consider our estimate to be reasonable since it is in line with other literature.

In order to handle investor activity, we create a variable for the average number of changes per year made by each individual. The average activity is the accumulated activity divided by the number of years the individual has been in the system. From here on, activity refers to average activity. By looking at an average over time, we account for the amount of time spent in the pension programme and receive a measurement comparable across investors of different ages. Five approximately equally large activity groups are generated, which are those who are “Inactive”, having made changes “Less than every 2 years”, “Every 2 years - 9 months”, “Every 9 months - 5 months” and “More than every 5 months” since the individual entered the system. Those in the

first group, “Inactive”, are the ones who automatically end up in the default fund and have never been active and those who have actively made an investment allocation decision before entering the system and have never made any changes since that first allocation. Additionally, we create one dummy variable for those who have automatically ended up in the default allocation and who have never made any changes, in order to distinguish the return for this specific group.

The individual characteristics age, income, gender and municipality are included in the regression and can be considered to be used as control variables. The income and municipality variables are not available for all individuals, and we have thus excluded individuals from the sample who have missing values in either one or both of them. The data is originally from SCB, and we can assume that the municipality information is missing at random, and thus dropping individuals missing a municipality will not cause any bias. The income data consists of the pensionable income for the assessment year 2016 and is reported in Swedish Krona (SEK). Regarding the income data, we have chosen to drop individuals who miss the data on the basis that the variable can be assumed to correlate with investor activity, through channels such as education and financial literacy for example. Causes for missing values can include those who have not reported any income such as students, people on parental leave, and those who have retired for example. These occurrences are not at random, and one has to take into account the fact the effect that such individuals are excluded from our sample when drawing conclusions from this paper. To account for income in our regression, five approximately equally large groups have been created; less than 225 000 SEK, 225 000-335 000 SEK, 335 000-410 000 SEK, 410 000-520 000 SEK and more than 520 000.

Once these restrictions have been imposed, we are left with 73 434 individuals in our sample. Table I illustrates the sample distribution across the individual characteristics gender, age

and income. The sample is evenly distributed between the genders, with 53.4 percent male individuals and 46.6 percent female. The age of the pension savers ranges from 17 to 79 years, as seen in Figure I. The income distribution, presented in Figure II, is more centred around the median and lower values, featuring a long right tail since there is no upper boundary to income. Table II illustrates the sample distribution across activity and return rates. We find that 17 percent of pension savers in our sample are inactive, of which 11.6 percent have been in AP7 Såfa for the whole period, and 5.4 percent have stayed in their initial allocation outside the default fund. There are a few individuals with extremely high incomes as well as a few with extremely high levels of activity. In order to avoid outliers to drive the regression disproportionately, we create categorical variables for both income and activity, grouping them according to approximately equally large groups. The distribution of returns by the different activity groups is displayed in Figure III. It is apparent that returns for inactive savers are concentrated around higher numbers, with a majority of observations within a range of returns between 5 and 15 percent. This category of investors also displays a wider dispersion of returns than all other groups, which are far more concentrated around their mean values. Mean returns for active savers are overall located around 6-7 percent, except for the most active group which has a peak shifted to the right indicating slightly higher mean returns.

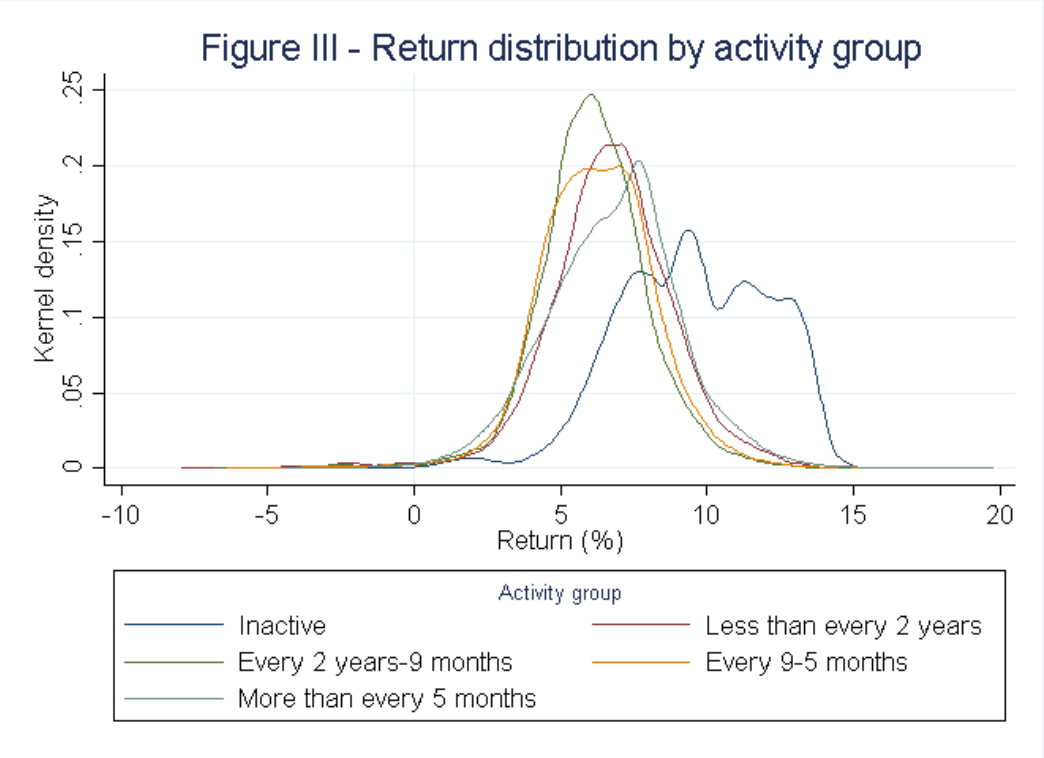

Figure III illustrates the kernel return rate distribution for the sample by dividing the sample into activity groups.

5 Method and Results

$$\boldsymbol{Return = \beta_0 + \beta_1 activity\ group_i + \beta_2 default\ dummy}$$

$$\boldsymbol{+ \beta_3 proxy\ advisory\ dummy + u}$$

The initial regression on the return by activity presented above and in Table III, produce significant negative coefficients for all levels of activity, indicating that active management of the portfolio has not generated increased returns for the savers compared to those who have been inactive investors. Although, those that have been highly active have performed better than those in the middle and lower parts of the activity range. On average those that have stayed in the default option have been better off than active investors, and those that have had their funds managed by a

Table III

OLS Regression	(1)	(2)	(3)	(4)	(5)
Explanatory variables	**Baseline**	**Add age**	**Add income**	**Add gender**	**Add municipality**
Activity					
Less than every 2 years	-0.874***	-1.027***	-1.033***	-1.033***	-1.028***
	(0.0281)	(0.0272)	(0.0271)	(0.0271)	(0.0272)
Every 2 years-9 months	-1.396***	-1.539***	-1.547***	-1.547***	-1.544***
	(0.0278)	(0.0271)	(0.0270)	(0.0270)	(0.0272)
Every 9-5months	-1.275***	-1.333***	-1.345***	-1.345***	-1.353***
	(0.0269)	(0.0263)	(0.0262)	(0.0262)	(0.0266)
More than every 5 months	-0.751***	-0.767***	-0.793***	-0.793***	-0.809***
	(0.0288)	(0.0285)	(0.0285)	(0.0285)	(0.0290)
Investors remained in default fund	2.875***	2.356***	2.385***	2.385***	2.378***
	(0.0341)	(0.0313)	(0.0312)	(0.0312)	(0.0314)
Proxy for advisory	-0.187***	-0.280***	-0.270***	-0.271***	-0.248***
	(0.0231)	(0.0222)	(0.0222)	(0.0222)	(0.0225)
Age					
30-39 years		0.665***	0.617***	0.617***	0.620***
		(0.0601)	(0.0599)	(0.0599)	(0.0599)
40-49 years		-0.422***	-0.515***	-0.515***	-0.504***
		(0.0597)	(0.0596)	(0.0596)	(0.0596)
50-59 years		-0.688***	-0.777***	-0.777***	-0.764***
		(0.0597)	(0.0595)	(0.0595)	(0.0595)
60-69 years		-0.808***	-0.848***	-0.848***	-0.836***
		(0.0606)	(0.0605)	(0.0605)	(0.0605)
70-79 years		-0.983***	-0.900***	-0.900***	-0.908***
		(0.0884)	(0.0897)	(0.0897)	(0.0896)
Income (SEK)					
225 000 - 335 000			0.0655*	0.0655*	0.0673**
			(0.0259)	(0.0259)	(0.0260)
335 000 - 410 000			0.156***	0.156***	0.154***
			(0.0253)	(0.0254)	(0.0255)
410 000 - 520 000			0.262***	0.262***	0.254***
			(0.0256)	(0.0259)	(0.0260)
520 000 -			0.351***	0.351***	0.312***
			(0.0256)	(0.0262)	(0.0268)
Gender				0.000818	-0.00618
				(0.0158)	(0.0158)
Constant	7.588***	8.128***	8.034***	8.033***	8.046***
	(0.0213)	(0.0623)	(0.0643)	(0.0652)	(0.145)
Municipality dummies	No	No	No	No	Yes
Observations	73434	73434	73434	73434	73434
Adjusted R-squared	0.278	0.316	0.318	0.318	0.321

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table III reports the results from the main regression using OLS. The dependent variable is return rate for all regressions. Regression no. (1) is the baseline regression from the equation on page 16. One by one, the control variables age, income, gender and municipality are added. Regression no. (5) is illustrated by the equation on page 18.

financial advisor have been worse off on average than those who have remained in the default option or who have actively managed their funds by themselves.

$$\boldsymbol{Return = \beta_0 + \beta_1 activity\ group_i + \beta_2 default\ dummy} \\ \boldsymbol{+ \beta_3 proxy\ advisory\ dummy + \beta_4 age\ group_i + \beta_5 income\ group_i} \\ \boldsymbol{+ \beta_6 female\ dummy + \beta_7 municipality_i + u}$$

Lastly, we run the multiple regression illustrated above, also presented in Table III, also including indicator variables for each of the age groups, income groups, municipalities and a dummy variable for gender. The estimates for the coefficients indicate how each respective group on average differs from not being included in that specific group. The first group is omitted for each of these variables in order to avoid multicollinearity. It is important to note that the activity-group that has been omitted covers those that have either never made any choice and thus have always been in the default fund and those who made one initial allocation change away from the default fund before they entered the system but have never made any changes after that.

Another point one should keep in mind is that the regressed return variable is not risk-adjusted, and thus an absolute return measure. All activity coefficients are negative and significant through all alterations made to the regression, predicting a negative effect from being an active investor on average, while holding all other variables constant. Worth noting is that although the investors in the dataset with the highest returns have been active, so have those with the lowest returns. Thus, the high dispersion in returns for highly active investors is worth keeping in mind when interpreting the results.

Those who made one initial choice and no more since that are 17-29 years old, had an income below 225 000 in 2016, and that are male are represented in the constant and has on average had a return of 8.05 percent in the final model including all the control variables. All other

coefficients indicate how much the return on average differs from that of the constant in each one of the regressions.

The coefficient for the dummy variable accounting for investors who have stayed in the default fund is also highly significant, implying a positive effect of allocating all funds in AP7 Såfa of on average 2.38 percent more in return than alternative allocations. This positive result indicates that the average investor would receive the highest return by staying in the default fund. Using a financial advisor, on the other hand, appears to have had a small but significant and negative impact on returns by on average -0.25 percent. Moreover, we have not taken into account the fees charged by advisors outside the system, meaning that the final net return to the investor will be even lower. Among active individuals, those performing best on average is the most active group, generating 0.81 percent lower return than that of investors in other activity groups. Following comes the group of low activity with changes less frequent than every second year with 1.03 percent lower return and 1.35 percent for those with average changes once every 9 - 5 months and then those who make allocation changes most infrequently, who generate 1.03 percent lower return on average. The worst performing group is those who have been somewhat active, having made on average one change in every second year to very nine months during their participation in the Premium Pension System, with on average 1.54 percent lower rate of return.

5.1 Sensitivity analysis

A set of control variables are added to the regression stepwise in order to perform a robustness check. As all our core variables of interest (activity, the default dummy and proxy for advisory) remain highly significant as control variables are added, structural validity can be inferred from coefficient robustness. The added variables are not the main focus for the regression at this stage,

which is what effect the activity, staying in the default fund and using a financial advisor has had on returns. The variables for age, income, gender and municipality are added in this order to the regression in order to exempt their effect on both activity and return. We thereby control the results for these individual-level characteristics not to drive the correlation between activity and return by being an underlying source of interaction. Since we do not have a randomised experiment and activity is actively chosen by the individuals in the sample it may vary systematically across the sample and thus cause a biased output. Including variables likely to be correlated to individual activity in the regression, and hence excluding them from the error term will make the results less biased. By looking at how the coefficients in the regression changes in each step, we get a sense of how these factors affect the results. As more control variables are accounted for, no coefficients change sign, increase or decrease drastically. As activity stays highly significant through all regressions and the adjusted R^2 increases by every added variable, except for that of gender, the results imply that we capture more of the explanation for differences in return by controlling for age, income and municipality. When including the control variables age, income and municipality we use the same strategy of recoding them into indicator variables as with activity. This strategy generates significant coefficients for all variables except the one for the female dummy.

Furthermore, we use robust standard errors to account for the possible occurrence of heteroskedastic standard errors within any of the variables. Since the sample size is large, this should pose no problem.

Due to the presence of outliers regarding both activity and return, a Least Absolute Deviation regression, also referred to as a median regression is run to make sure that these few observations do not have an unproportionally strong effect on the results. The Least Absolute Deviations regression is more robust than the Ordinary Least Squares regression in the sense that

it is more resistant to outliers in the data. This strategy conveys how the conditional median of returns changes with activity, instead of the conditional mean as in OLS, and provides a fairly similar output presented in Table IV. The evident alterations in the results are larger standard errors and a smaller coefficient for the group of highest activity, which is to be expected since the outliers should be less driving for the estimates. Other coefficients change somewhat but keep the same sign and significance. Thus, the main results remain robust to this alternate method of estimating the model.

5.2.1 Investor activity and performance

We do not look at variation within factors for each individual, but cross-sectional variation across the individuals. We do this due to the available measurement of return for each person is one single covering the whole sample period, the internal rate of return measured since the start. We thus compare the difference in investor activity between individuals, since it is beyond the scope of this paper to compare the investor activity difference over time for each individual with a varying rate of return. To exclude reverse causality we ideally would have wanted data where we could have separated the investor activity and the return in two periods where the return follows previous investor activity. This would have allowed us to draw conclusions in the causal direction of interest and avoid a regression where returns drive investor activity. That is unfortunately not the case, and thus it is only possible to determine whether there is a correlation between return and activity with certainty within the scope of this study. It is evident from the output that there is, in fact, a significant correlation between the two. In order to identify a causal relationship, there are three important components. First of all is the cause and effect component which means in this case that activity should come before return and not the other way around. This possible scenario seems

believable due to the fact that if everybody would see their return before engaging with their funds, all investors would start in the default fund and then based on that return choose to stay or to switch away from it. Since it is not the case that all investors do, it is more likely that activity precedes return. The second component is a correlation which has already been established. Lastly, is the credibility component. Overall it seems credible that investors would choose funds based on expected returns in the future. Since it is evident from our sample data that investor activity is overall quite low, with one-fifth of investors not having made any changes to their allocation at all since they entered the Premium Pension System, the first premise appears reasonable. Despite all components pointing to a direction of causality from activity to returns, we can not entirely rule out the possibility that the return rate might spur one to engage more or less in the pension programme.

Differences in returns in an individual's account are explained by specific fund choices in combination with the timing of the change. What factors that affect the activity of an individual are on the other hand numerous and in principle impossible to fully incorporate into a regression. The theory suggests that activity should tend towards riskier investment and produce high expected returns for all savers. It is evident that since activity is overall low and a large fraction stays in AP7 Såfa, investors do not act accordingly, and there are notable differences in behaviour across the sample. The inclusion of control variables is an effort to take account for and sort out these systematic differences in activity and return. Although it is important to keep in mind that there are other potential sources of variation in returns such as financial knowledge, mathematic-skills, education, amount of spare-time etc., we do not know if these variables are correlated to investor activity since we do not have a natural experiment and there is no such data available.

Criticism towards defined contribution plans arguing that the pressure to actively manage the funds in order to generate returns is too high given the low level of engagement savers generally present, do not align with the findings in this study where inactive savers earn the best returns. The fact that the coefficient is negative and highly significant for all activity groups is consistent with the hypothesis that trading derives from overconfidence. Another rational explanation for this discrepancy from common opinion is that the Swedish default alternative AP7 Såfa has achieved higher returns to its investors compared to its counterparts in other countries, where inactive investors might have suffered due to a lower quality of default funds. Worth mentioning is that AP7 Såfa is a leveraged fund, of 23.67 percent since 2018 and 35 percent in previous years, which may not be the case for comparable default funds. The leverage in combination with a strong development of stock markets might also explain why the Swedish default alternative outperforms many active investors, which might not have been the case in another market situation and does not ensure future performance. This result indicates that activity within the Premium Pension System has been less effective as a tool to generate returns due to the superior return offered by the default alternative.

It is important to bear in mind that the measure of return used is not risk-adjusted. Thus, it is not possible to draw any conclusion on the quality of the investment decisions investors have made. Higher returns can be a result both from taking on higher risk and making better investment decisions, and it is not within the scope of this paper to decompose the higher returns.

The main result of a positive correlation between activity and returns in the study by Dr Dahlqvist et al. (2016) also differs from the result of this study. Although the most active investors in both studies earn high returns, the conclusion made by us is that the ones in the default fund are the winners on average and not the ones with higher activity. Important to take note of is that there

are differences in the execution between the two which could contribute to the different outcomes. Given the small fraction of the population that constitutes both samples, differences could be pushed by unalike samples. Another probable driver of differences is the sample period, where our sample cover all 18 years while the earlier conducted study covers the first 11 years. Considering the development of the financial markets during the two different time periods, the other study ended its sample period shortly after the last financial crisis while this paper includes following years of recovery and well performing financial markets. Thus, returns will differ, and the effect activity will appear to have on the returns will do so as well. The default fund previous to AP7 Såfa which was replaced in 2010 did most likely not provide great returns as the study by Dahlqvist et al. (2016) reveals, in contrast to ours including the later positive development of the markets, which AP7 Såfa has incorporated with leverage and thus generated highly positive returns. Moreover, while Dahlquist and his colleagues in their conclusion focus on mass changes being harmful to the entire Premium Pension System, our result highlight that savers who have not been active since they left their advisors in 2011 have not benefited from staying with their advisor in the fund they have offered where they have remained.

5.2.2 How individual characteristics relate to activity and return

In this section, we will go through how the variables age, income and gender relate to the level of activity and absolute returns within the Premium Pension System. Evident from the main regression in Table III is that these variables are important in explaining the returns within the system. As age is added to the baseline regression the adjusted R^2 increases from 0.278 to 0.316, suggesting that the age group an investor belongs to explain a significant part of the return an individual will on average get. Our data shows that older investors have performed worse than

younger investors, and the result is highly significant for all age groups. The coefficient for the oldest age group in the fifth regression including all control variables can be interpreted as an investor being 70-79 years old being predicted to get on average a 0.91 percent lower return than those not belonging to that specific age group. These findings contributes to the previous ambiguous literature on age and risk-aversion by supporting the theory that risk-aversion increases with age.

The income shows that investors with higher incomes on average have seen higher returns. The coefficient is positive and significant for all income groups, and the positive effect increases for every group. Investors with a higher income than 520 000 SEK for the year 2016 have on average seen 0.31 percent higher returns than those with a lower income than that when including all the control variables. This result is consistent with previous theories stating that high net worth individuals experience an increase in risk tolerance. Since our return measurement is absolute, we cannot draw any conclusion on whether the higher returns that high-income individuals see are due to taking a higher risk or making better investment decisions, but we can presume that both factors are important.

The gender coefficient is very small in absolute value, and switches sign when adding municipality dummies and is insignificant. The result indicates that whether the investor is male or female on average has no effect on the return which contrasts theories about gender differences in investments such as women being more risk-averse than men. Although unknown within the scope of this study whether the theories apply to the investors when considering their whole portfolios, it can be established that any potential differences are not reflected in this pension programme.

Figure IV-VI plots the kernel distribution of return by age groups, income groups and gender. Figure VII-IX shows histograms illustrating the level of activity by age groups, income groups and gender. From Figure IV it is evident that the two youngest age groups, 17-29 and 30-39, have distributions that differ from the other age groups. There is overweight on higher returns, and the distribution is a lot more dispersed than for other age groups, for which distributions are more concentrated around a return of 8 percent. There are two factors that might be contributing to these results; first of all many of the young investors were not in the system when the global financial crisis hit in 2008, or those who were had just entered the system. Since the return measurement is capital weighted the drawdown did not have as much effect on their weighted returns. A second explanation is that those who have been in the system since the start (i.e. the older groups) have more concentrated results due to the fact that they have been through exactly the same market conditions as each other for almost 17 years and thus their results have had more time to converge, with different portfolio allocations weighting out each other over time. Figure VII shows that young investors have an overweight in the inactive group. This result might indicate that young investors are relatively less interested in their pension savings than those getting closer to their retirement. The fact that young people have been less active but have on average received higher returns than the older age groups is consistent with our finding that investors who have remained in the default fund have on average received 2.38 percent higher returns than those who have left the default option.

Figure V shows a lower concentration for the lowest income group around a return of 7 percent than the higher income groups which have a higher concentration around the 8 percent point. Despite that, the lines switch places in the right tail with the higher returns, with the highest income group displaying a lower dispersion of returns than any other group. Figure VIII illustrates

clearly that income correlated positively with activity. Almost 30 percent of investors in the lowest income group has remained in the default fund or made one initial choice before entering the system, while the corresponding number for the highest income group is 10 percent. On the other hand, a quarter of investors in the highest income group have made changes on average more frequently than every five months.

Figure VI shows that there are almost no differences in the return distribution between the two genders, which is consistent with our finding that there is no significant difference between male and female investors in our sample. This result is inconsistent with previous research stating that females receive lower returns due to a higher risk aversion than male investors. From Figure IX we can see that the only clear difference between men and women is that men tend to be more active than women, but the activity has almost no impact on the return rates, which is consistent with previous findings that men tend to be overconfident which leads to excessive trading hurting their returns.

Table IV-X illustrate various statistical tests that have been carried out in order to see if the means in average activity and return are statistically different from each other by age groups, income groups and gender. To test if the means are different for age groups and income groups the one-way analysis of variance (ANOVA) has been used since there are more than two independent groups. The output for one-way ANOVA only states if at least two groups are significantly different from each other or not. In order to specify what specific group means are different from each other the Tukey post-hoc test is used, which does pairwise comparisons. It should be noticed that these tests do not include control variables but take into account only the two relevant variables of interest. Thus they might exhibit some bias, and the tests should be interpreted as nothing more than a sanity check. Table V tests for the difference in average activity

mean by age groups. There was a statistically significant difference between groups as determined by one-way ANOVA (F(5,73428)=416.8, p=.000). A Tukey post hoc test revealed that activity was statistically significantly different for all age groups at the 5 percent level except for the groups 70-79 vs 40-49 and 70-79 vs 50-59. Between all other age groups, the older group was more active than the younger group, except for 70-79 vs 60-69 where the younger group was more active. Table VI tests for the difference in return rate mean by age groups. There was a statistically significant difference between groups as determined by one-way ANOVA (F(5,73428)=2049.98, p=.000). A Tukey post hoc test revealed that returns were statistically significantly different for all age groups at the 5 percent level except for the groups 70-79 vs 50-59. Between all other age groups, the older groups have seen lower returns than the younger group except for 70-79 vs 60-69 where the older group has performed better. Table VII tests for the difference in average activity mean by income groups. There was a statistically significant difference between groups as determined by one-way ANOVA (F(4,73429)=213.07, p=.000). A Tukey post hoc test revealed that activity was statistically significantly different for all age group at the 5 percent level except for the groups 225,000-335,000 vs -225,000. Between all other income groups, the group with the highest income was the most active. Table VIII tests for the difference in return rate mean by income groups. There was a statistically significant difference between groups as determined by one-way ANOVA (F(4,73429)=39.07, p=.000). A Tukey post hoc test revealed that returns were statistically significantly different for all age groups at the 5 percent level except for the groups 410,000-520,000 vs. 335,000-410,000, 520,000- vs. 335,000-410,000 and 520,000- vs. 410,000-520,000. Between all other income groups, the group with the highest income has seen lower returns. Since this test does not control for activity while the main regression in Table III does, the positive correlation between income and return is more likely to be the true correlation.

To test if the means are different for gender an independent t-test has been used since there are only two unrelated groups. It determines whether the mean difference between the two groups if different to zero. Table IX tests for the difference in average activity mean by gender. The value 0 represents male investors, and 1 is female investors. The results show that male investors had statistically significantly higher activity (1.65+-0.01) than female investors (1.38+-0.01) at the 5 percent level, (t(73432)=17.5087, p=0.000). Table X tests for the difference in return rate mean by gender. The results show that there is no statistically significant difference in return rate for male investors (7.00+-0.01) and female investors (6.97+-0.01) at the 5 percent level. (t(73432)=1.6037, p=0.1088).

6 Concluding remarks

6.1 Answering the research question

By comparing savers within the premium pension system our study suggest that investor activity has a negative effect on absolute returns, although the most active investors generate higher returns than those less active. The use of financial advisors previous to the ban of mass changes has not provided better returns for those who stayed in the final allocation given compared to other savers. The completely inactive savers with all capital allocated in the default fund form the group with highest absolute returns to their premium pension savings. The results also suggest that there are differences in activity and returns for investors across individual investor characterics, with regards to age and income while gender has no impact on returns although men on average display higher level of activity.

6.2 Limitations of our study and suggestions for future research

Why AP7 Såfa has outperformed alternative allocations be due to several causes and should be subject to further research in order to determine whether the fund can be assumed to continue to do so in the future. Bearing in mind that much of AP7 Såfa's absolute return can be attributed to the bull market in previous years, it can be considered a great risk that a large number of investors are inactive since the high concentration of capital in the same fund may have far-reaching societal implications should the market conditions turn unfavourable for AP7 Såfa, or the fund for some other reason performs badly.

Given that our R^2 is quite low at 32.1 percent there should be additional research on what factors might be the most important for absolute returns. Variables of interest that we did not have access to are for example education level and financial literacy. Regarding activity, it would be interesting to see what factors drive investor activity and specifically the choice to deviate from the default allocation within the Swedish Premium Pension System, more than an individual's willingness to take on risk.

All in all, the inference drawn from this study is no certain truth for the population or pension system as a whole, but bearing in mind the assumptions made we consider it possible to reach reasonable conclusions regarding the effect of activity on returns within the Premium Pension System.

7 References

Books

Arrow, Kenneth J. 1971, "Essays in the theory of risk-bearing", Chicago: Markham.

E-journals

Barber, Brad M. & Odean, Terrance. 2001, "Boys will be boys: gender, overconfidence, and common stock investment", *The Quarterly Journal of Economics*, vol. 116, no. 1, pp. 261-292.

Barber, Brad M. & Odean, Terrance. 2000, "Trading is hazardous to your wealth: the common stock investment performance of individual investors", *The Journal of Finance*, vol. 55, no. 2, pp. 773-806.

Battocchio, Paolo & Menoncin, Francesco. 2004, "Optimal pension management in a stochastic framework", *Insurance: Mathematics and Economics*, vol. 34, no. 1, pp. 79-95.

Hira, Tahira K., Rock, Whitney L. & Loibl, Caezilia. 2009, "Determinants of retirement planning behaviour and differences by age", *International Journal of Consumer Studies*, vol. 33, no. 3, pp. 293-301.

Goetzmann, William N. & Kumar, Alok. 2008, "Equity portfolio diversification", *Review of Finance*, vol. 12, no. 3, pp. 433-463.

Graham, John R., Harvey, Campbell R. & Huang, Hai. 2009, "Investor competence, trading frequency and home bias", *Management Science*, vol. 55, no. 17, pp. 1094-1106.

Markowitz, H. 1952, "Portfolio Selection", *The Journal of Finance*, vol. 7, no. 1, pp. 77-91.

Wang, Hui & Hanna, Sherman D. 1997, "Does risk tolerance decrease with age?", *Financial Counseling and Planning*, vol. 8, no. 2, pp. 27-32.

Government documents

Pilot study S2017/03516/SF. 2017, "Pilot study of possible structural changes for a more secure and more efficient premium pension system". Stockholm: Ministry of Health and Social Affairs.
Available: http://www.government.se/4ad143/globalassets/government/dokument/socialdepartementet/pilot-study-premium-pension_en.pdf [2018-05-09]

Study Ds 2013:35. "Vägval för premiepensionen". Stockholm: Ministry of Health and Social Affairs; Ministry of Finance.
Available: http://www.regeringen.se/49bb49/contentassets/11b0f648e97241858cb3074516931bb7/vagval-for-premiepensionen-ds-201335 [2018-05-09]

Websites

The Seventh AP Fund. "AP7 investerar I faktorpremier och sänker hävstångens normalnivå". *Nyheter 2018-04-25.* Available: https://www.ap7.se/aktuellt/ap7-investerar-i faktorpremier-och-sanker-havstangens-normalniva/ [2018-05-10]

The Swedish Pensions Agency. "Fondrådgivare och förvaltningstjänster". Available: https://www.pensionsmyndigheten.se/forsta-din-pension/valj-och-byt-fonder/fondradgivare-och-forvaltningstjanster [2018-05-01]

Other

The Swedish Pensions Agency. "Premiepension höst 2017". PowerPoint presentation obtained from the agency [2018-01-05]

The Swedish Pensions Agency. "2017.xlsx". Excel document obtained from https://www.pensionsmyndigheten.se/statistik-och-rapporter/statistik/statistik-for-premiepension [2018-04-10]

8 Appendix

8.1 Tables

Table I

		Sample	
		N	%
Gender	Female	34 236	46,6%
	Male	39 198	53,4%
	All	*73 434*	*100,0%*
Age	17-29	3 433	4,7%
	30-39	11 823	16,1%
	40-49	20 691	28,2%
	50-59	22 357	30,4%
	60-69	13 560	18,5%
	70-79	1 570	2,1%
	All	*73 434*	*100,0%*
Income (SEK)	-225 000	14 476	19,7%
	225 000-335 000	14 772	20,1%
	335 000-410 000	15 047	20,5%
	410 000-520 000	14 265	19,4%
	520 000-	14 874	20,3%
	All	*73 434*	*100,0%*

Table I shows descriptive statistics covering the whole sample with regards to the independent characteristics of the sample including gender, age and income.

Table II

		Sample	
		N	%
Inactive investors	Default fund	8 520	11,6%
	One initial choice	3 950	5,4%
	Total	*12 470*	*17%*
Average activity	Inactive	12 470	17,0%
	Less than evry 2 years	14 419	19,6%
	Every 2 years-9 months	16 168	22,0%
	Every 9-5 months	16 575	22,6%
	More than every 5 months	13 802	18,8%
	All	*73 434*	*100,0%*
Return (%)	-0	467	0,6%
	0-3	2 125	2,9%
	3-6	22 925	31,2%
	6-10	39 916	54,4%
	10-15	7 968	10,9%
	15-	33	0,0%
	All	*73 434*	*100,0%*

Table II shows descriptive statistics covering the whole sample with regards to the main variables of interest; the dummy variable for investors in the default fund, the different activity groups and the return rate.

Table IV

LAD Regression	(1)
Explanatory variables	**Least Absolute Deviation**
Activity	
Less than every 2 years	-1.038***
	(0.0407)
Every 2 years-9 months	-1.616***
	(0.0412)
Every 9-5months	-1.420***
	(0.0404)
More than every 5 months	-0.668***
	(0.0409)
Investors remained in default fund	2.309***
	(0.0459)
Proxy for advisory	-0.284***
	(0.0278)
Age	
30-39	-0.158***
	(0.0459)
40-49	-1.423***
	(0.0459)
50-59	-1.727***
	(0.0460)
60-69	-1.916***
	(0.0473)
70-79	-1.972***
	(0.0720)
Income (SEK)	
225 000-335 000	0.0867**
	(0.0272)
335 000-410 000	0.132***
	(0.0275)
410 000-520 000	0.252***
	(0.0282)
520 000-	0.371***
	(0.0286)
Gender	0.00746
	(0.0173)
Constant	9.019***
	(0.0589)
Observations	73434

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table IV reports the results from the main regression using Least Absolute Deviation (LAD). The dependent variable is return rate. Municipality dummies are excluded from this regression.

Table V

Age groups	Summary of avg_activity Mean	Std. Dev.	Freq.
17-29	.44381057	1.6063158	3,433
30-39	1.0910839	1.7516681	11,823
40-49	1.5030478	1.9362112	20,691
50-59	1.718958	2.1700387	22,357
60-69	1.8961217	2.3645823	13,560
70-79	1.6094664	2.7902594	1,570
Total	1.5277945	2.1036493	73,434

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	8968.44667	5	1793.68933	416.80	0.0000
Within groups	315997.572	73428	4.30350237		
Total	324966.019	73433	4.42534036		

Bartlett's test for equal variances: chi2(5) = 2.1e+03 Prob>chi2 = 0.000

Pairwise comparisons of means with equal variances

over : age_group

	Number of Comparisons
age_group	15

avg_activity	Contrast	Std. Err.	Tukey t	Tukey P>|t|	Tukey [95% Conf.	Interval]
age_group						
30-39 vs 17-29	.6472734	.0402189	16.09	0.000	.5326612	.7618855
40-49 vs 17-29	1.059237	.0382303	27.71	0.000	.950292	1.168182
50-59 vs 17-29	1.275147	.0380271	33.53	0.000	1.166781	1.383513
60-69 vs 17-29	1.452311	.039635	36.64	0.000	1.339363	1.565259
70-79 vs 17-29	1.165656	.0632033	18.44	0.000	.9855452	1.345767
40-49 vs 30-39	.4119639	.0239162	17.23	0.000	.3438098	.4801179
50-59 vs 30-39	.6278741	.0235899	26.62	0.000	.5606497	.6950984
60-69 vs 30-39	.8050378	.0261029	30.84	0.000	.7306522	.8794233
70-79 vs 30-39	.5183825	.0557232	9.30	0.000	.3595877	.6771773
50-59 vs 40-49	.2159102	.020012	10.79	0.000	.158882	.2729384
60-69 vs 40-49	.3930739	.0229207	17.15	0.000	.3277568	.458391
70-79 vs 40-49	.1064187	.0543054	1.96	0.366	-.0483357	.261173
60-69 vs 50-59	.1771637	.02258	7.85	0.000	.1128173	.2415101
70-79 vs 50-59	-.1094915	.0541625	-2.02	0.330	-.2638387	.0448556
70-79 vs 60-69	-.2866552	.0553033	-5.18	0.000	-.4442533	-.1290572

Table V shows STATA output for the one-way analysis of variance (ANOVA) and Tukey post-hoc test with pairwise comparisons for difference in average activity mean between age groups.

Table VI

Age groups	Summary of Return rate Mean	Std. Dev.	Freq.
17-29	9.3899105	3.4019897	3,433
30-39	8.3903641	2.7446346	11,823
40-49	6.8090143	2.2516893	20,691
50-59	6.4674016	2.1133071	22,357
60-69	6.3550437	2.0283893	13,560
70-79	6.5562215	2.3014204	1,570
Total	6.9910329	2.4840543	73,434

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	55503.7078	5	11100.7416	2049.98	0.0000
Within groups	397616.499	73428	5.41505283		
Total	453120.207	73433	6.17052561		

Bartlett's test for equal variances: chi2(5) = 2.9e+03 Prob>chi2 = 0.000

Pairwise comparisons of means with equal variances

over : age_group

	Number of Comparisons
age_group	15

rr	Contrast	Std. Err.	Tukey t	Tukey P>|t|	Tukey [95% Conf.	Interval]
age_group						
30-39 vs 17-29	-.9995464	.045115	-22.16	0.000	-1.128111	-.870982
40-49 vs 17-29	-2.580896	.0428843	-60.18	0.000	-2.703104	-2.458689
50-59 vs 17-29	-2.922509	.0426563	-68.51	0.000	-3.044067	-2.800951
60-69 vs 17-29	-3.034867	.04446	-68.26	0.000	-3.161565	-2.908169
70-79 vs 17-29	-2.833689	.0708973	-39.97	0.000	-3.035726	-2.631653
40-49 vs 30-39	-1.58135	.0268276	-58.94	0.000	-1.657801	-1.504899
50-59 vs 30-39	-1.922962	.0264616	-72.67	0.000	-1.99837	-1.847555
60-69 vs 30-39	-2.03532	.0292805	-69.51	0.000	-2.118761	-1.951879
70-79 vs 30-39	-1.834143	.0625067	-29.34	0.000	-2.012268	-1.656017
50-59 vs 40-49	-.3416127	.0224481	-15.22	0.000	-.4055833	-.2776421
60-69 vs 40-49	-.4539706	.0257109	-17.66	0.000	-.5272391	-.3807021
70-79 vs 40-49	-.2527928	.0609162	-4.15	0.000	-.4263861	-.0791994
60-69 vs 50-59	-.1123579	.0253288	-4.44	0.000	-.1845375	-.0401782
70-79 vs 50-59	.0888199	.060756	1.46	0.689	-.0843167	.2619565
70-79 vs 60-69	.2011778	.0620356	3.24	0.015	.0243945	.377961

Table VI shows STATA output for the one-way analysis of variance (ANOVA) and Tukey post-hoc test with pairwise comparisons for difference in return rate mean between age groups.

Table VII

Income groups	Summary of avg_activity Mean	Std. Dev.	Freq.
-225 000	1.3117896	2.1014814	14,476
225 000-3	1.3254682	1.8335371	14,772
335 000-4	1.4435805	1.8569778	15,047
410 000-5	1.6513238	2.14115	14,265
520 000-	1.9056805	2.4650711	14,874
Total	1.5277945	2.1036493	73,434

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	3728.49362	4	932.123404	213.07	0.0000
Within groups	321237.525	73429	4.37480457		
Total	324966.019	73433	4.42534036		

Bartlett's test for equal variances: chi2(4) = 1.8e+03 Prob>chi2 = 0.000

Pairwise comparisons of means with equal variances

over : income_group

	Number of Comparisons
income_group	10

	avg_activity	Contrast	Std. Err.	Tukey t	Tukey P>|t|	Tukey [95% Conf.	Interval]
income_group							
225 000-335 000 vs -225 000		.0136785	.0244615	0.56	0.981	-.053047	.080404
335 000-410 000 vs -225 000		.1317909	.0243506	5.41	0.000	.0653678	.1982139
410 000-520 000 vs -225 000		.3395342	.0246757	13.76	0.000	.2722244	.4068441
520 000- vs -225 000		.5938909	.02442	24.32	0.000	.5272787	.6605031
335 000-410 000 vs 225 000-335 000		.1181123	.024226	4.88	0.000	.0520293	.1841953
410 000-520 000 vs 225 000-335 000		.3258557	.0245527	13.27	0.000	.2588814	.39283
520 000- vs 225 000-335 000		.5802123	.0242957	23.88	0.000	.5139392	.6464855
410 000-520 000 vs 335 000-410 000		.2077434	.0244423	8.50	0.000	.1410704	.2744163
520 000- vs 335 000-410 000		.4621	.024184	19.11	0.000	.3961315	.5280686
520 000- vs 410 000-520 000		.2543567	.0245113	10.38	0.000	.1874953	.321218

Table VII shows STATA output for the one-way analysis of variance (ANOVA) and Tukey post-hoc test with pairwise comparisons for difference in average activity mean between income groups.

Table VIII

Income groups	Summary of Return rate Mean	Std. Dev.	Freq.
-225 000	7.200851	2.8069403	14,476
225 000-3	7.0296551	2.5763757	14,772
335 000-4	6.9225371	2.4403016	15,047
410 000-5	6.9153283	2.339532	14,265
520 000-	6.890369	2.2062083	14,874
Total	6.9910329	2.4840543	73,434

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	962.393742	4	240.598435	39.07	0.0000
Within groups	452157.813	73429	6.15775529		
Total	453120.207	73433	6.17052561		

Bartlett's test for equal variances: chi2(4) = 1.0e+03 Prob>chi2 = 0.000

Pairwise comparisons of means with equal variances

over : income_group

	Number of Comparisons
income_group	10

rr	Contrast	Std. Err.	Tukey t	Tukey P>|t|	Tukey [95% Conf.	Interval]
income_group						
225 000-335 000 vs -225 000	-.1711959	.0290212	-5.90	0.000	-.2503592	-.0920326
335 000-410 000 vs -225 000	-.2783139	.0288896	-9.63	0.000	-.3571183	-.1995095
410 000-520 000 vs -225 000	-.2855227	.0292753	-9.75	0.000	-.3653792	-.2056662
520 000- vs -225 000	-.310482	.0289719	-10.72	0.000	-.3895108	-.2314532
335 000-410 000 vs 225 000-335 000	-.107118	.0287417	-3.73	0.002	-.185519	-.028717
410 000-520 000 vs 225 000-335 000	-.1143268	.0291294	-3.92	0.001	-.1937852	-.0348684
520 000- vs 225 000-335 000	-.1392861	.0288244	-4.83	0.000	-.2179127	-.0606596
410 000-520 000 vs 335 000-410 000	-.0072088	.0289983	-0.25	0.999	-.0863097	.0718921
520 000- vs 335 000-410 000	-.0321681	.028692	-1.12	0.795	-.1104334	.0460971
520 000- vs 410 000-520 000	-.0249593	.0290803	-0.86	0.912	-.1042838	.0543651

Table VIII shows STATA output for the one-way analysis of variance (ANOVA) and Tukey post-hoc test with pairwise comparisons for difference in return rate mean between income groups.

Table IX

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	39,198	1.654556	.0113184	2.24087	1.632371	1.67674
1	34,236	1.382661	.0104007	1.924433	1.362276	1.403047
combined	73,434	1.527795	.0077629	2.103649	1.512579	1.54301
diff		.2718945	.0155291		.2414575	.3023315

diff = mean(0) - mean(1) t = 17.5087
Ho: diff = 0 degrees of freedom = 73432

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 0.0000

Table IX shows STATA output for an independent t-test to test for difference in average activity mean between male (=0) and female (=1) investors.

Table X

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	39,198	7.004771	.0126719	2.508853	6.979934	7.029608
1	34,236	6.975304	.0132697	2.455296	6.949294	7.001313
combined	73,434	6.991033	.0091667	2.484054	6.973066	7.009
diff		.0294675	.0183752		-.0065478	.0654827

diff = mean(0) - mean(1) t = 1.6037
Ho: diff = 0 degrees of freedom = 73432

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.9456 Pr(|T| > |t|) = 0.1088 Pr(T > t) = 0.0544

Table X shows STATA output for an independent t-test to test for difference return rate mean between male (=0) and female (=1) investors.

8.2 Figures

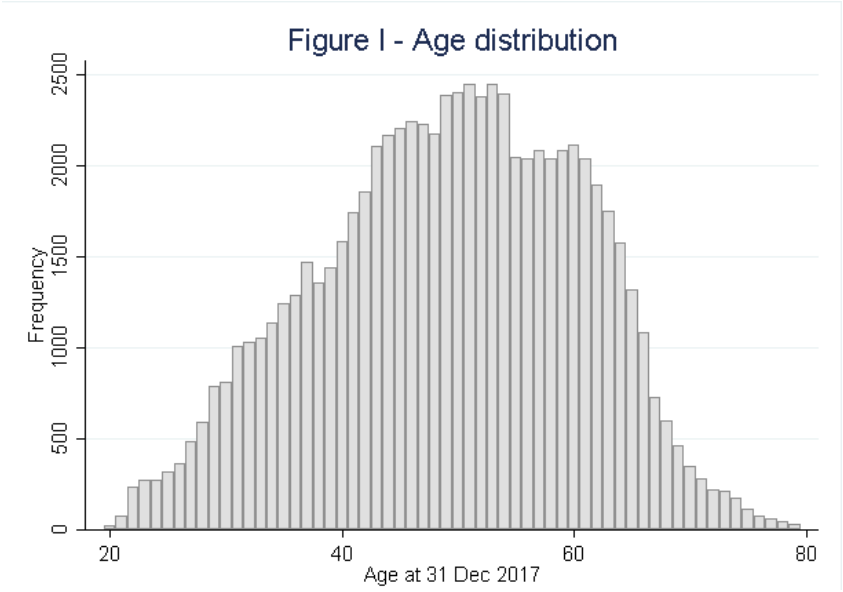

Figure I illustrates age distribution for the final sample.

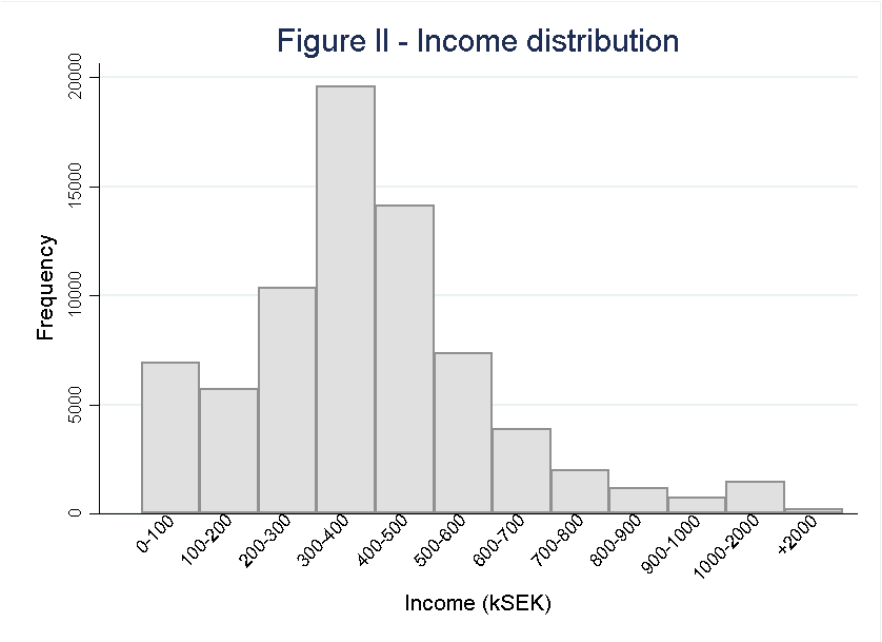

Figure II illustrates income distribution for the final sample.

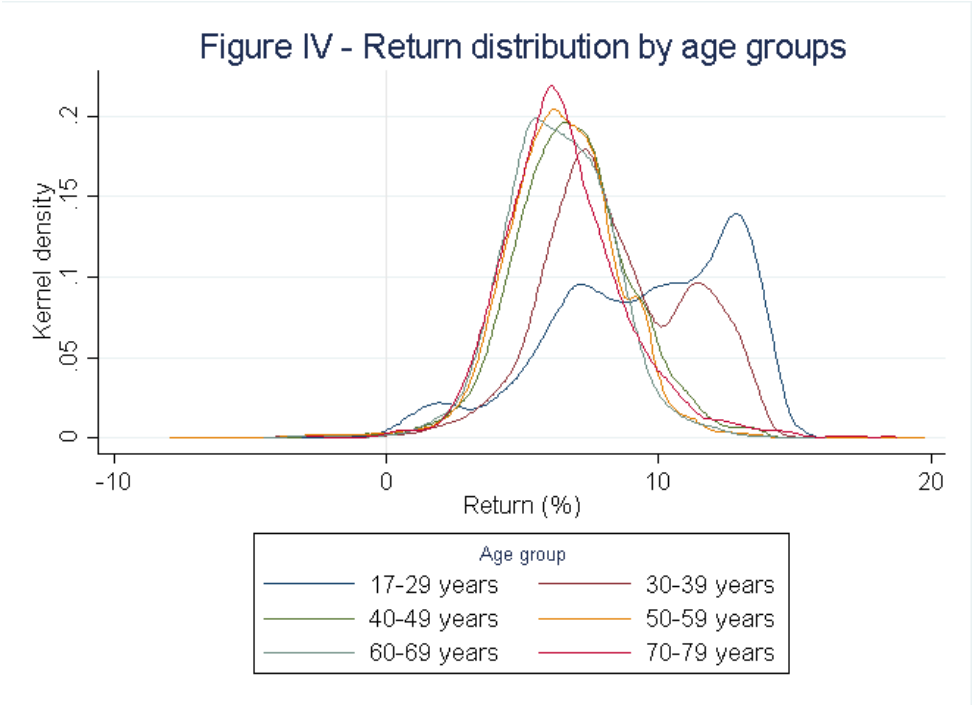

Figure IV illustrates the kernel return rate distribution for different age groups.

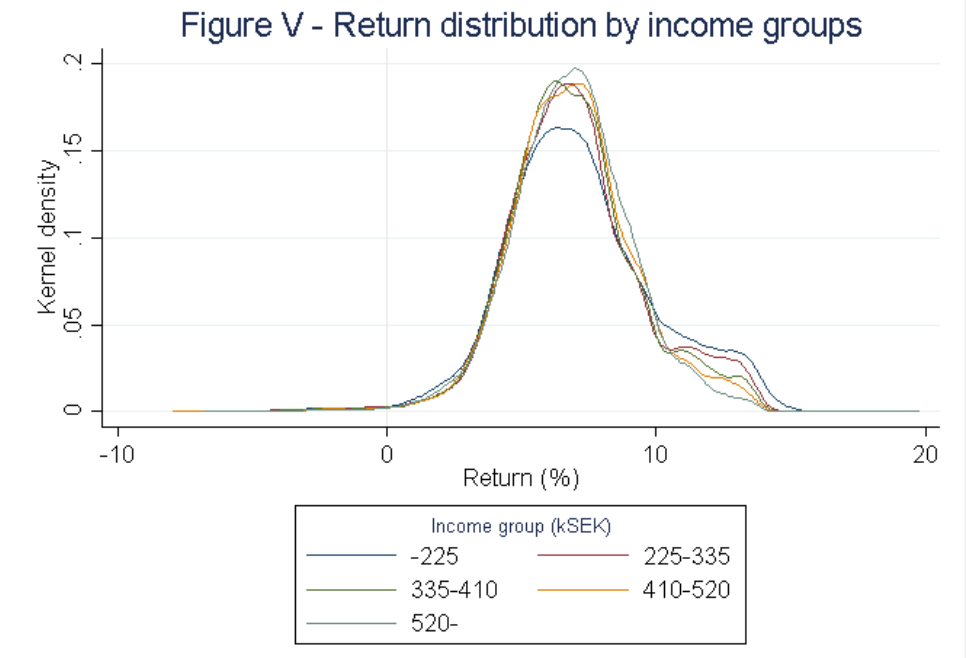

Figure V illustrates the kernel return rate distribution for different income groups.

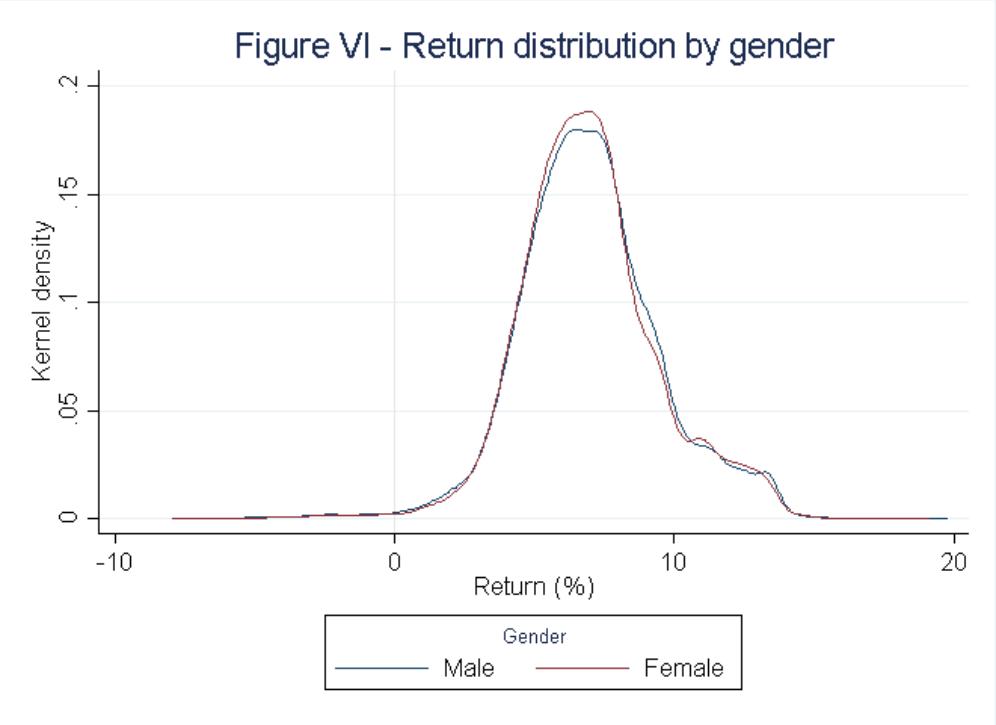

Figure VI illustrates the kernel return rate distribution for different genders.

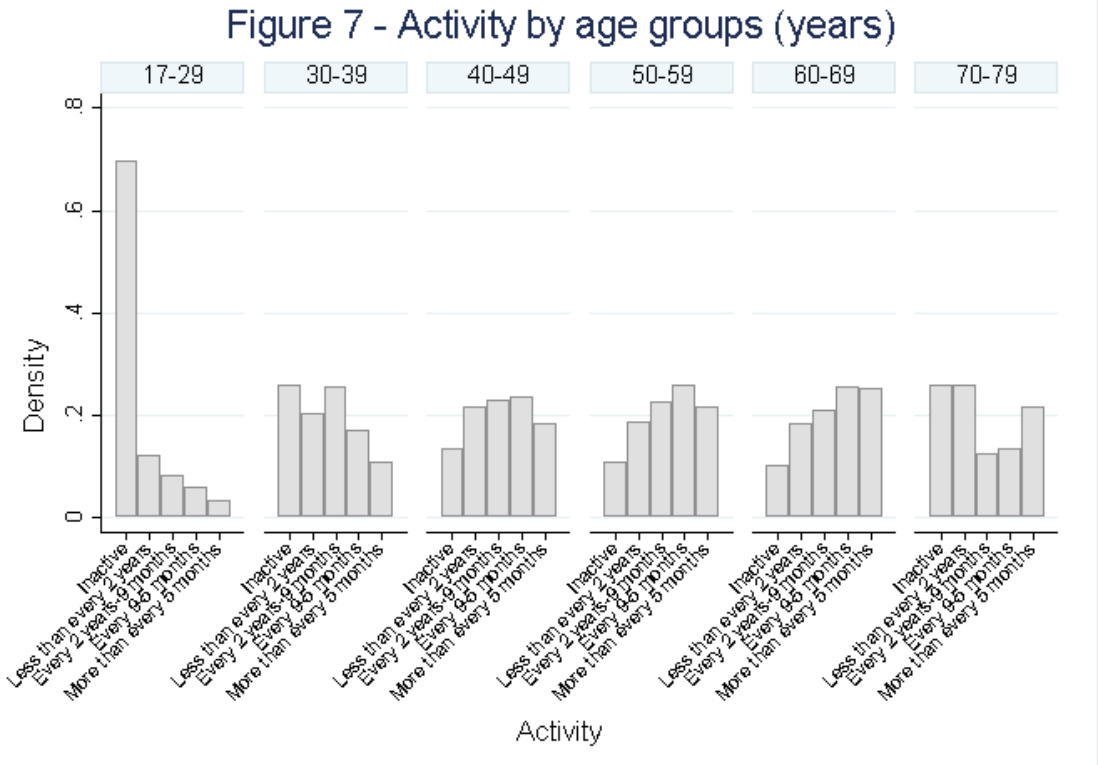

Figure IV illustrates the activity distribution for different age groups.

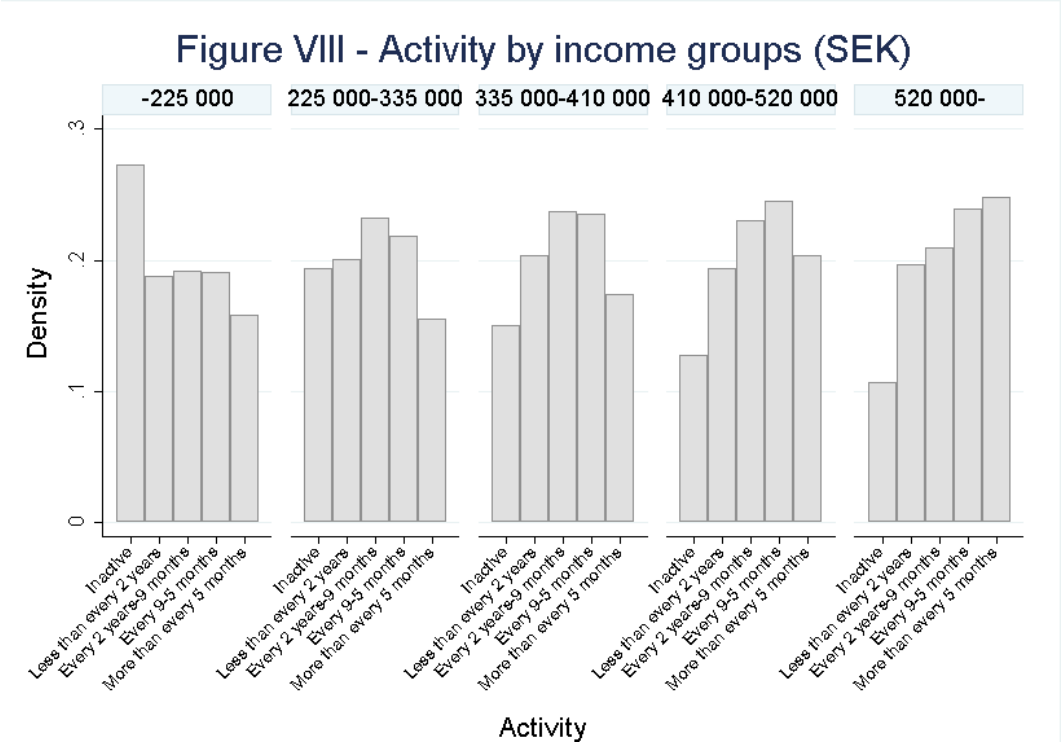

Figure VIII illustrates the activity distribution for different income groups.

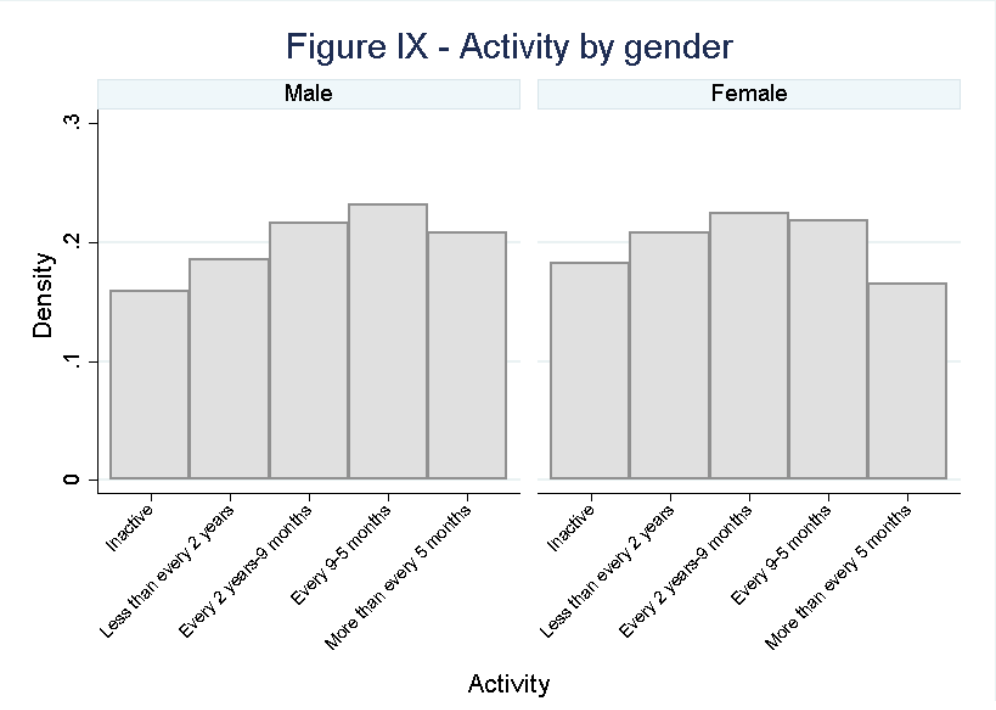

Figure IX illustrates the activity distribution for different genders.