Supplementary Appendix:

Princelings in the Private Sector

David Szakonyi

George Washington University &

HSE University

Contents

1 Sample Construction

1.1 Individual Propiska (Registration) Data: Larix

To create the household units necessary to identify family links, I first used administrative data from the Russian propiska (i.e. registration) system. The firm Moscow Center for Economic Security acquires this registration data from the federal government, retitles it 'Larix', and sells it openly to customers. Larix databases dating back to 2000 can be found online, with new versions being released on a regular basis. I obtained a free version of Larix 10 data (from 2005) from publicly available sources on the internet; as of writing, Larix version 88.0 (updated in January 2018) was available for purchase online. The Larix 10 database used in this paper includes information on individuals' names, birthdates, places of registration, and ownership of property in the city of Moscow.

I started by cleaning the data on residency registrations. Using regular expressions and fuzzy data tools, I standardized the registration addresses into a comprehensible format common to the Russian postal system that uses four elements: street, 'dom' (house), 'korpus' (building), and 'kvartira' (apartment). For inclusion in the sample for creating family links, I required addresses to have both information on 'dom' and 'kvartira' (most in fact also contain information on buildings). In the small number of cases where the algorithm could not accurately standardize the data into the four main elements, I required an address to contain some combination of street name and a numerical listing for the house; this allows me to keep the small number of standalone homes present in Moscow. Together these two restrictions result in a loss of roughly 5% of the full address data. I also removed the approximately 1% of apartments that have more than seven individuals registered there at any one time in order to guard against mistakes or incompleteness in data entry. This left a database of 20,477,152 unique residence registrations at 7,344,794 unique addresses.

Next, I matched this residence registration data to an accompanying database within

Larix 10 that contains information on the individuals actually associated with each registration. I performed basic cleaning procedures, such as standardizing names using regular expressions, and removing individuals that were missing either their first or last names or whose names had too few characters (approximately 9% of the sample). In all, I was left with 18,008,385 individuals that could be matched to their residences. Individuals can be registered to multiple residences over the time period, hence the discrepancy between the number of unique individuals and the number of unique registrations.

I identified families by linking individuals that shared the last name and were registered at the same residence(s). After lemmatizing female last names (recoding 'Ivanova' as 'Ivanov' to create a last name 'core'), I excluded all individuals who had a unique last name ('core') that was not shared by anyone in the dataset. For these individuals, it would be impossible to link them to any other individuals, since they alone possess their last name. Of the 17,403,834 individuals remaining, 14,641,305 have entries in the cleaned housing registration database from above. I assigned each of these individuals a unique family identifier that included all other persons who both shared their last name 'core' and had registered at any of the same residences. This is the primary dataset used to measure family links.

1.2 Individual Reported Earnings Data: Pension Fund

The administrative data on reported earnings came from the Russian State Pension Fund. It covered the period from 1999 to 2004. Employers file earnings information with the Fund on an annual basis. Employers are indicated by a unique tax identification number that can be used to match organizations across both individuals and years. In addition, the database contains information on income generated from other sources (such as lotteries and gifts); I excluded these observations from the analysis. To date, the exact reasons for how and why the database appeared in the public domain in 2004, first on compact discs and then later posted to the Internet, are still unclear. In this paper, I relied on the

validation efforts of the multiple teams of academic researchers which have also utilized the data. The best of these checks involved comparing sample averages to official Moscow labor statistics collected by the Russian State Statistics Agency (Braguinsky, Mityakov, and Liscovich, 2014). The two databases are broadly comparable for each year and across all years on average. Braguinsky, Mityakov, and Liscovich (2014) also provide an exhaustive list of the resultant works, including Braguinsky and Mityakov (2015) on wage misreporting among foreign and domestic firms in Moscow, Mironov (2013) on tax enforcement, and Guriev and Rachinsky (2008) on the evolution of personal income.

To match the Pension Data to the residence registration data, I needed complete information on an individual's name (first, middle, and last) and date of birth. Recent work on the United States has shown that the combination of gender, birthday, last name, and zip code makes an individual unique at a rate of 1 in 2.7 billion (Ansolabehere and Hersh, 2017). The Pension Data has only limited information on zip code, but is available mainly for Moscow residents, which provides a geographic constraint helpful for matching to the purely Moscow registration data. In addition, the availability of first, middle (patronymic) and last names (as well as ability to infer gender from the patronymic) improved my ability to uniquely identity and match individuals across datasets.

I started with the raw Pension data file, which contains 53,578,600 observations. First, I used a series of unique identifiers to impute birthdates for individuals across years. The reason was that the full 1999-2004 Pension Data was constructed by combining individual files for each year. Several of these annual files lack data on birthdates, but contain other information such as passport numbers and tax identifiers for physical persons (FLINN) that can be used to match individuals. I developed an algorithm to iterate over combinations of full names and these unique identifiers (as well as employer tax IDs) to fill in missing birthdates between the different years. I then removed observations with abbreviated first and middle names, since they cannot be used to uniquely match individuals with the registration data. I also removed duplicate entries in the income data based on

full name, birthdate, employer, and income. In the end, I am left with 49,324,733 complete observations for 13,731,751 unique individuals that contain: full name, birthdate, year, employer tax identification number, and annual income. Each individual is assigned a unique person id based on their full name and birthdate.

Table A1 presents summary statistics for the Pension Data sample used in the analysis (by year). Overall the sample looks balanced across the years, with the exception of 2002 (the year following tax reform), where the total number individuals in the dataset somewhat drops. To account for this, I require that only officials from federal ministries that reported consistent employment numbers in all six years be included in the sample; they form the basis of the family links analysis. I also include year fixed effects in all models that help account for this over-time variation.

TABLE A1: PENSION FUND DATA: SUMMARY STATISTICS BY YEAR

Year	No. Individuals	No. Employers	Mean Age	Female (%)
1999	6,063,392	191,659	43.19	0.53
2000	6,887,893	206,930	43.33	0.52
2001	6,188,357	171,743	42.53	0.50
2002	4,996,537	150,419	42.91	0.51
2003	6,109,448	144,937	43.61	0.53
2004	6,313,468	157,789	44.95	0.52

1.3 Merging Individual-Level Datasets

The final step was merging the Propiska (Registration) Data with the Pension Fund data on reported earnings. Before getting to this point, cleaning procedures were undertaken on the two somewhat internally consistent databases. However, merging the registration and the earnings creates a number of challenges endemic to electronic databases created from paper records. Even with the standardization and cleaning, many entries in each of the databases are still marked by any number of typographical errors and abbreviations.

Of the 14,641,305 entries in the housing registration data, I was able to match 6,330,496

to their reported earnings. Officially, this match rate corresponds well to the overall employment rate in Moscow during the period. For example, representative surveys conducted by the well-regarded Levada Center found that roughly 45% of respondents aged 18 and above held some kind of paid employment during the period.[1] The remaining half of the respondents included students, pensioners, housewives, and the unemployed. In that respect, the matching algorithm performs rather well in matching housing registrations with earnings records.

We must note that far from all individuals in the Pension Data found matches in the housing registration data. This could cause problems for the analysis if certain individuals working for federal government ministries were systematically registering their families outside of Moscow (or failing to register at all). If we believed these individuals faced incentives to avoid doing so, our estimates about the labor market returns for their relatives might be biased. The Pension Data contains 40,044 federal employees from 1999-2004, as defined by the fifteen ministries and presidential administration examined in the paper. Of this number, I was able to locate housing registration records for 35,776 individuals, or a rate of 89%. Given their official status within the federal government, these individuals are thus registered in Moscow at far higher levels, improving our ability to identify members of their family (and their labor market outcomes).

In Table A2, I show results from models that examines the determinants of this missingness. The analysis is done at the individual level, with a binary outcome indicating if a person identified as working in a federal ministry or for the presidential administration was matched to his or her entry in the residence registration data. Given the limited nature of the pension data, there are few variables available to put on the right-hand side. Nevertheless, I include measures of age, income, gender, and the number of years he or

[1] This statistics comes from the first available omnibus survey Levada conducted for each year of the period under analysis. Levada Center Courier, February 19, 1999 - February 22, 1999; Levada Center Courier, December 30, 1999 - January 5, 2000; Levada Center Courier, January 10, 2001 - January 22, 2001; Levada Center Courier, January 25, 2002 - January 28, 2002; Levada Center Courier, January 24, 2003 - January 28, 2003; and Levada Center Courier, January 9, 2004 - January 12, 2004.

she worked for the federal government. I also include dummies for the specific ministry or agency. The results indicate that matching success depended primarily on how long an individual worked for the government: the longer their stay, the more likely their entry in the registration could be found. Factors such as age, income and gender are not significant predictors, which lends support to the notion that missingness is somewhat random. However, specific ministries do vary to the extent that their employees can be matched to registration records, although the point estimates are substantively small and the person fixed effects used in the models in the main text should account for this variation. This final individual-level sample must be read thus with these caveats in mind. The analysis is restricted to only individuals who have officially registered their residence in Moscow. All those employed in legal entities (either the private or public sector) but who have not registered cannot be linked to family members because of the absence of registration, and thus housing data to construct family household units.

Lastly, the number of federal employees used in the analysis sample is necessarily far lower for several reasons. First, I only include federal employees (and their family members) if they spent at least one year out of federal office during the period (without this variation, there would be no room for identification). Second, I require federal employees to make at least 30,000 rubles per year that they are working for the federal government. Next, to enter the dataset, a family (i.e. a household unit with at least one federal employee) must have at least one employed member besides the federal employee. I include this final constraint to ensure that the analysis focuses only on individuals able to enter into formal employment. Finally, I only look at close relatives (parents, children, and spouses) of relatives in order to more carefully guard against measurement errors. That leaves the final sample of 7,944 federal employees.

TABLE A2: MATCH RATE BETWEEN REGISTRATION AND PENSION FUND DATA

	Outcome: Individual Not Found in Registration Data	
	(1)	(2)
Log Age	−0.011	−0.007
	(0.019)	(0.022)
Log Income	0.005	0.001
	(0.008)	(0.009)
Female	0.012	0.017
	(0.009)	(0.012)
Number of Years Employed	−0.041	−0.039
	(0.010)	(0.010)
Ministry: Economic Development		−0.054
		(0.004)
Ministry: Education		−0.062
		(0.007)
Ministry: Culture		−0.027
		(0.012)
Ministry: Science and Technology		−0.095
		(0.007)
Ministry: Transportation		−0.063
		(0.001)
Ministry: Foreign Affairs		0.049
		(0.003)
Ministry: Emergency Situations		0.043
		(0.002)
Ministry: Taxes		−0.003
		(0.006)
Ministry: Energy		−0.051
		(0.003)
Ministry: Property Management		−0.008
		(0.007)
Ministry: Natural Resources		−0.028
		(0.008)
Ministry: Communications		−0.050
		(0.001)
Ministry: Finance		0.0003
		(0.007)
Ministry: Labor		−0.022
		(0.004)
Observations	38,302	38,302
R^2	0.047	0.063

The outcome is a binary indicator for whether I was able to match a person who had worked at any time for a federal ministry or the presidential administration with their residence registration data. The sample excludes the 4% of individuals who worked for more than one ministry. Standard errors are clustered at the ministry level.

1.4 Individual-Level Data: Additional Summary Statistics

- Table A3 gives descriptive statistics about the breakdown of employees and employment characteristics for the different federal institutions. Table A4 gives descriptive statistics for the analysis sample used in the main text. The level of analysis is the individual-year, which does not affect the statistics (means, etc.) for the variables measured at simply the individual level (such as gender) since all individuals have six years in the dataset. Finally, Table A6 gives descriptive statistics for the firms analyzed in the main text.

TABLE A3: SUMMARY STATISTICS BY FEDERAL INSTITUTION

Federal Institution	Employees	Mean Salary (ths. $)	Mean # of Relatives	Mean # of Kids	Married (%)	Female (%)
Economic Development	1,159	2.84	1.29	0.38	66.26	35.38
Education	316	2.35	1.23	0.25	76.27	56.65
Culture	57	4.94	1.23	0.39	59.65	40.35
Labor	291	2.08	1.26	0.27	76.63	49.83
Science and Technology	445	2.29	1.31	0.45	75.51	30.79
Transportation	590	2.51	1.29	0.37	70.85	36.95
Foreign Affairs	1,545	2.65	1.31	0.43	62.91	18.06
Emergency Situations	328	3.54	1.29	0.36	77.74	26.83
Taxes	401	2.94	1.19	0.19	64.84	45.39
Property Management	215	2.99	1.21	0.18	59.53	54.88
Energy	314	2.06	1.19	0.26	70.38	41.40
Natural Resources	280	2.61	1.24	0.29	77.86	38.21
Communications	181	2.21	1.30	0.37	70.17	44.20
Finance	551	3.36	1.15	0.12	69.87	65.15
Presidential Administration	1,271	4.98	1.28	0.39	72.07	26.44

Note the mean number of children will be underestimated since mother-child relationships cannot be precisely measured (patronymics cannot be used).

TABLE A4: SUMMARY STATISTICS FOR ANALYSIS SAMPLE

Statistic	N	Mean	St. Dev.	Min	Max
Individual is Female	60,384	0.472	0.499	0	1
Individual Age	60,384	42.899	14.316	16	94
Individual Age Difference from Relative in Federal Office	60,384	3.192	18.923	−56	56
Relative in Federal Office is Father	60,384	0.271	0.444	0	1
Relative in Federal Office is Son or Daughter	60,384	0.163	0.370	0	1
Relative in Federal Office is Spouse (3 yr. diff)	60,384	0.424	0.494	0	1
Relative in Federal Office is Spouse (6 yr. diff)	60,384	0.566	0.496	0	1
Relative in Federal Office is Spouse (9 yr. diff)	60,384	0.566	0.496	0	1
Relative in Federal Office	60,384	0.446	0.497	0	1
Individual Employed	60,384	0.661	0.474	0	1
Individual Salary (log)	39,884	10.744	1.374	7.601	20.392
Individual Works in Private Firm	60,384	0.269	0.443	0	1
Individual Works in SOE	60,384	0.058	0.233	0	1
Individual Works for Federal Government	60,384	0.158	0.364	0	1
Individual Works for Regional Government	60,384	0.085	0.279	0	1
Individual Works for Local Government	60,384	0.001	0.035	0	1
Individual Works in Relative's Previous Job	60,384	0.006	0.079	0	1
Individual Salary Rank	32,338	65.720	25.753	1.000	100.000

1.5 Firm-Level Data: Sample Construction and Summary Statistics

I collect the firm-level data from several publicly availably sources. First to minimize errors due to over-time changes in firm registrations, I match employers to their entries in the firm registry operated by the Russian State Statistics Agency (Rosstat). Each year, Rosstat collects information on basic firm characteristics as well as balance sheets and other financial data. Scholars have commonly accessed this data through the Professional Market and Company Analysis System (SPARK) which acquires the data from Rosstat and creates a user-friendly interface for analysts, journalists and academics to use; recent examples include Mironov (2013) investigating tax evasion and Szakonyi (2016) looking at businessperson candidacy.

Unfortunately, one drawback of SPARK is the limited availability of many of these indicators for the period of the early 2000s: SPARK prioritizes making current information available for analysis, rather than tracking changes over time. Therefore, I acquired a different version of the Rosstat registry available online. This registry is dated to December 2004 and contains basic information on 6,081,114 firms. The registry has complete information on firm name, tax identifiers (INN, OKPO, OGRN), a variety of official government classifiers that help identify ownership and the subnational level of operation (OKFS, OKOGU, OKOPF), sectoral membership (OKVED), and original date of registration (to calculate years in business). This allows me to more squarely match the employers in the dataset with the characteristics during the period under study. The Pension Dataset has tax identification numbers (INNs) for each employer allowing me to nearly match across organization and year between the two datasets. Of the 11,017 employers in the datasets (before banks and firms based outside of Moscow were excluded), I am able to locate 9,679, or 88% in this firm registry. I then code the ownership of employers according to state registration codes that demarcate private sector firms, state-owned firms, and those organizations owned and operated by federal, regional, or local governments.

Finally, to look at procurement outcomes, I use a database of over 540 million inter-firm

banking transactions between 1999-2004 collected by the Russian Central Bank. There are unique legal entities contained in the dataset. As described in the main text, this unique dataset comprises a near universe of registered firms and organizations, with a slight bias towards those located in Moscow (again which should not cause significant problems because the main individual-level analysis only looks individuals living in that city) (Mironov, 2013). Each entry in the dataset describes a transaction between a paying organization and a receiving organization, with information given on the date of the transaction and the amount transferred.

Table A5 gives summary statistics by year for the banking transactions data. The first column gives the total volume of transactions in each year (measured in billions of rubles), while the second denotes the number of individual transactions between payers and receivers. Next, we see that in the early years of the sample, just a fraction of the transactions contained a textual description of the goods or service being provided. Only in the last years of the sample (2003 and 2004) is there a textual description for the vast majority of the transactions.

TABLE A5: BANKING TRANSACTIONS DATA BY YEAR

Year	Volume (bil. rubles)	No. Transactions	No. Transactions w/ Description	(%)
1999	10,943.51	47,342,083	15,940,182	33.7
2000	24,122.9	60,193,580	33,063,808	54.9
2001	36,974.65	77,750,004	45,923,576	59.1
2002	44,282.29	88,400,780	68,627,241	77.6
2003	57,280.09	100,376,233	98,287,847	97.9
2004	98,001.18	130,201,636	128,514,687	98.7

This missingness somewhat complicates the task of identifying which transactions are actually the result of state procurement contracts for firms in Moscow. Moreover, many of the descriptions themselves are inscrutable and difficult to code as being payments under state procurement contracts, or other types of transactions. To identify the former and prioritize transactions for goods and services rendered to the government, I adopt a multi-step strategy. First, using the firm registration data, I code whether the 'payer' is a (1) tax agency or (2) a bank, and (3) whether a 'receiver' is a utility providing water

and/or electricity using Russian standardization and sector codes. In the first case, firms may receive VAT rebates for engaging in value-add activities such as investing in new facilities or production equipment. Removing tax agencies helps guard against those type of payments being mistaken for state contracts. In the second case, many state-owned banks hold firm deposits and facilitate cross-national transactions; therefore, I exclude all payments from banks, as indicated by their OKVED sector code and OKOGU classifiers. Finally, I follow Mironov and Zhuravskaya (2016) in excluding utilities providing electricity and water since these contracts are not awarded competitively through the state procurement system; I identify utilities using the primary OKVED sector code. Appendix Section 4.1 shows results that include the small number of employers that operate as utilities. To identify the entities that make the payments, I merge the Rosstat firm registry data with the transactions data. After excluding payments from these three types of entities, I calculate aggregate transactions to the receiver-year level. I use the same coding scheme based on state registration codes to sum the amount of transactions each employer receives from federal, regional, and local government entities.

This transactions data is then merged to the employer data using tax identification numbers. Of the 8,752 employers in the Rosstat registry after the sample restrictions are applied, 100% have at least one transaction in the Russian Central Bank Data. Overall, there are transaction data available for 90% of employer-years, a particularly high number considering many firms were not officially registered for the entire period. Summary statistics for the entire employer sample can be found in Table A6. Note that this table includes all the employers connected to individuals that have relatives working in the federal government, including government entities, private firms, SOEs, etc. The main analysis subsets to private firms, for which summary statistics are available in Table A7.

TABLE A6: SUMMARY STATISTICS (ALL EMPLOYERS)

Statistic	N	Mean	St. Dev.	Min	Max
Number of Employees Per Year	49,461	341.36	2,284.69	0	245,001
Number of Gov. Officials' Relatives	49,461	0.30	0.88	0	34
Had Federal Government Contract	49,461	0.59	0.49	0	1
Federal Government Contract (ths. $)	49,461	6,300.66	332,272.60	0	42,529,931
Had Regional Government Contract	49,461	0.44	0.50	0	1
Regional Government Contract (ths. $)	49,461	318.08	6,207.26	0	771,601
Had Local Government Contract	49,461	0.07	0.26	0	1
Local Government Contract (ths. $)	49,461	0.94	30.20	0	3,806
Privately Owned	49,461	0.62	0.49	0	1
State-Owned Enterprise	49,461	0.07	0.26	0	1
Age (years)	49,461	9.12	3.85	1	100
Construction	49,461	0.07	0.25	0	1
Light Manufacturing	49,461	0	0	0	0
Heavy Manufacturing	49,461	0	0	0	0
Other Services	49,461	0.49	0.50	0	1
Trade / Transportation	49,461	0.27	0.45	0	1
Mining / Agriculture	49,461	0.01	0.09	0	1

TABLE A7: SUMMARY STATISTICS (ONLY PRIVATE FIRMS)

Statistic	N	Mean	St. Dev.	Min	Max
Number of Employees Per Year	30,725	140.21	1,578.47	0	189,460
Number of Gov. Officials' Relatives	30,725	0.19	0.43	0	7
Had Federal Government Contract	30,725	0.50	0.50	0	1
Federal Government Contract (ths. $)	30,725	188.18	2,439.78	0	236,709.70
Had Regional Government Contract	30,725	0.35	0.48	0	1
Regional Government Contract (ths. $)	30,725	62.82	649.78	0	59,190.73
Had Local Government Contract	30,725	0.05	0.22	0	1
Local Government Contract (ths. $)	30,725	0.46	19.66	0	3,040.23
Age (years)	30,725	8.19	3.24	1	41
Construction	30,725	0.09	0.29	0	1
Light Manufacturing	30,725	0	0	0	0
Heavy Manufacturing	30,725	0	0	0	0
Other Services	30,725	0.34	0.47	0	1
Trade / Transportation	30,725	0.37	0.48	0	1
Mining / Agriculture	30,725	0.01	0.09	0	1

2 Robustness: Individual-Level

2.1 Entrances, Exits and Retirements

- Table A8 analyzes the differential effects of having relatives enter, exit or retire from political office. The subsamples are created by identifying trajectories where (1) relatives entered political office and never left until the sample ended in 2004, e.g. clean entrances, (2) relatives began the sample period in 1999 in political office but left by the end of the time series, e.g. clean exits, and (3) presumably retired. Relatives had to spend at least two years in office to enter the sample. This third category assumes that males above the age of 60 and females above the age of 55, for the most part, had to leave office according to federal law. Figures A1 and A2 plot the age when individuals left the sample for good, according to their gender. There is a sharp drop-off for males after 60 and females after 55. As discussed in the main text, individual ministries in some instances could grant exceptions to allow bureaucrats extend their time in office.

- The results in Table A8 indicate several important patterns. First, it appears that when looking at employment (Panel A), a relative's exit from federal office is roughly equally as an important than a relative's entrance: the point estimate in Column 1 is nearly the sample than that of Column 2. In Columns 3 and 4, we see that in spite of the relatively small number of observations, limiting the sample to just presumed retirements of relatives from political office returns robust coefficients. We can be more confident that the causal relationship defended in the paper holds up. The mandated retirement ages are unrelated to any unobserved shocks. Once a relative leaves federal government, the opportunities to take advantage of these family connections to secure employment disappear.

- With regards to income conditional on employment (Panel B), we see that most of the

positive effect of having a relative in federal office comes from 'entrances' rather than 'exits'. The point estimate in Column 1 are significantly larger than that of Column 2, with the latter being statistically indistinguishable from 0. Similarly, when looking at retirements (Columns 3 and 4), we see that having a relative leave office does not have a statistically significant effect on one's wages, conditional on being employed, though the size of the effects is roughly the same. This suggests that the political connections do not simply recede if an individual is already employed. Political access can survive a relative's departure.

TABLE A8: ENTRANCES, EXITS AND RETIREMENTS

	Panel A: Relative Employed			
	Clean Entrances	Clean Exits	Retirements	
	(1)	(2)	(3)	(4)
Relative in Federal Office	0.025	0.054	0.086	0.115
	(0.013)	(0.017)	(0.038)	(0.041)
Demographic Controls	No	No	No	Yes
Year FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	No
Control Group Mean	0.63	0.63	0.62	0.62
Observations	17,400	9,144	1,962	1,962
R^2	0.414	0.413	0.418	0.016
	Panel B: Total Income (log)			
	Clean Entrances	Clean Exits	Retirements	
	(1)	(2)	(3)	(4)
Relative in Federal Office	0.054	−0.063	0.111	0.043
	(0.032)	(0.048)	(0.111)	(0.145)
Demographic Controls	No	No	No	Yes
Year FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	No
Control Group Mean	1.60	4.08	3.18	3.18
Observations	11,657	6,079	1,311	1,311
R^2	0.781	0.751	0.748	0.093

The outcome variable in Panel A is a binary indicator for whether an individual was employed in a given year. The outcome variable in Panel B is log annual income for the individual in a given year, conditional on being employed. Linear probability models (OLS) are used for all models. The main predictor of interest is a binary indicator for whether the individual had a federal relative in office that year. Column 1 analyzes only clean entrances, whereby a relative entered office and never left during the sample period. Column 2 analyzes only clean exits, whereby a relative left office and never returned during the sample period. Columns 3 and 4 analyze presumed retirements where individuals left office around their federal mandated retirement age and never returned. The first three models include year fixed effects and individual fixed effects, while the fourth drops individual fixed effects. Control Group Mean displays the base employment rates and annual income (measured in thousands of dollars) when a relative was not in federal office.

FIGURE A1: MALE RETIREMENTS

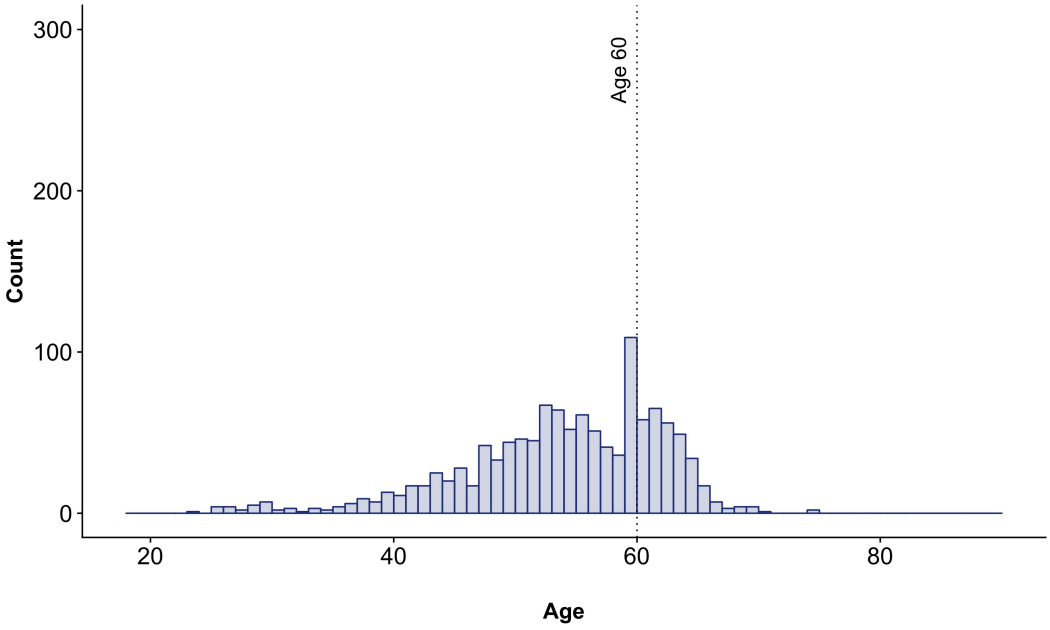
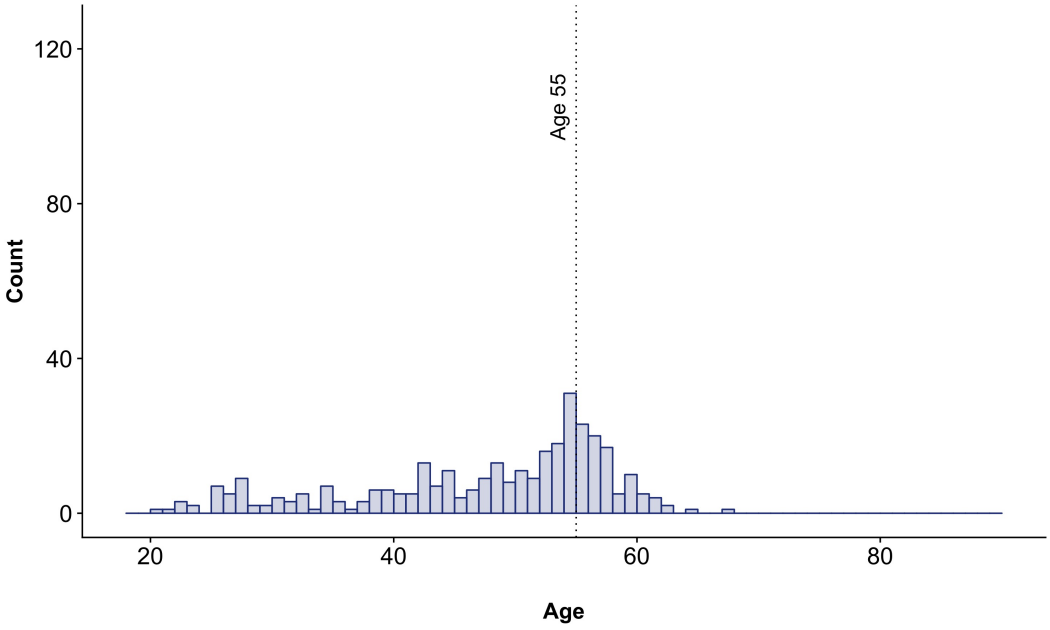


FIGURE A2: FEMALE RETIREMENTS



These two figures show the distribution of ages for relatives when they exit public office ('clean exits') permanently using the six-year sample. The two panels show different thresholds because of the different retirement ages for men and women.

2.2 Excluding Retirements from the Analysis

- Analyzing retirements helps demonstrate that the results are robust to subsetting the sample to only exogenously determined exits. However, one of the reasons that spouses appear to benefit most from nepotism could be that husbands and wife are retiring around the same time. The simultaneity of the spouses' exit from the labor market then would have little to do with political connections being severed, but rather be the result of a couple making a joint decision to retire.

- To test whether retirements are solely driving the results, Table A9 runs the same analysis but excludes all the retirement-related exits analyzed above in Table A8. The point estimates are only slightly smaller than those in the full sample are consistently significant at conventional levels. Having a relative move in and out of the federal government for reasons other than retirement also results in substantial changes to an individual's labor market prospects.

TABLE A9: EXCLUDING RETIREMENTS

	Panel A: Individual Employed				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.051	0.051	0.045	0.032	0.032
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	No	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	0.64	0.64	0.64	0.64	0.64
Observations	58,422	58,422	58,422	58,422	58,422
R^2	0.411	0.023	0.437	0.637	0.586
	Panel B: Income (log)				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.042	0.037	0.036	0.033	0.034
	(0.013)	(0.015)	(0.013)	(0.017)	(0.016)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Individual FE	Yes	No	Yes	Yes	Yes
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	4.44	4.44	4.44	4.44	4.44
Observations	38,573	38,573	38,573	38,573	38,573
R^2	0.767	0.154	0.773	0.893	0.873

The outcome variables are a binary indicator for whether an individual was employed (Panel A) and their log income, conditional on being employed (Panel B). All models are linear probability with standard errors clustered on individual and year. Column 1 includes individual fixed effects, Column 2 controls for individual age and drops fixed effects, Column 3 adds age fixed effects, and Columns 4 and 5 includes linear 'individual' time-trends and 'family' time-trends, respectively. Control Group Mean measures employment rates and income (in thousands of dollars) when a relative was not in office. The sample is restricted to all individuals who do not have relatives retiring from federal office.

2.3 Broadening the Definition of Family Members

- In the main text, I defined the analysis sample by only looking at nuclear family ties within the household. This coding decision produces much more precisely estimated coefficients and hones in on the direct family relations undergirding nepotism in Russia. Table A10 zooms out to include all family members registered in the household, even if they do not necessarily appear to be have typical nuclear family relations with the relative entering federal office: grandparents, siblings, cousins, and even marriages with larger age gaps. The empirical specifications are identical to Table 2 in the main text.

- We see that the results are robust to using this broader coding. Having a relative enter federal office can significantly affect one's employment prospects, both in terms of finding a job and earning income. However, these point estimates are considerably smaller than those shown in the main text, suggesting that federal employees may not go as far out of their way to help their more distant family members in the labor market.

TABLE A10: BROADER FAMILY DEFINITION

	Panel A: Individual Employed				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.033	0.031	0.028	0.020	0.020
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	No	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	0.64	0.64	0.64	0.64	0.64
Observations	110,088	110,088	110,088	110,088	110,088
R^2	0.410	0.014	0.437	0.639	0.550
	Panel B: Income (log)				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.017	0.014	0.014	0.009	0.015
	(0.009)	(0.011)	(0.009)	(0.012)	(0.011)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Individual FE	Yes	No	Yes	Yes	Yes
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	3.44	3.44	3.44	3.44	3.44
Observations	71,912	71,912	71,912	71,912	71,912
R^2	0.769	0.151	0.775	0.893	0.854

The outcome variables are a binary indicator for whether an individual was employed (Panel A) and their log income, conditional on being employed (Panel B). All models are linear probability with standard errors clustered on individual and year. Column 1 includes individual fixed effects, Column 2 controls for individual age and drops fixed effects, Column 3 adds age fixed effects, and Columns 4 and 5 includes linear 'individual' time-trends and 'family' time-trends, respectively. Control Group Mean measures employment rates and income (in thousands of dollars) when a relative was not in office. The sample includes all individuals with relatives in federal office, no matter what their relation was.

2.4 Varying Family Definitions

- In the main text, I define marriage as occurring if two individuals are of different genders and whose ages are within six years of one another. Table A11 presents the baseline results for the full sample in Column 1. Columns 2, 3 and 4 vary the number of years of difference between these two individuals that is used to define marriage (Column 3 corresponds to the results in the table in the main text). The results are robust to these different codings of marriage - being married to a federal employee nearly doubles one's employment opportunities. In Column 5, I add an additional category indicating siblings of the federal employee - individuals whose age is within 15 years of the employee and who share the same patronymic. This indicates both people shared the same father, but does not restrict on gender. Siblings do not see additional benefits in the labor market more than the average relative.

TABLE A11: INDIVIDUAL-LEVEL MECHANISMS: VARYING FAMILY DEFINITIONS

	Panel A: Individual Employed				
	Full Sample	Married (3 yr. gap)	Married (6 yr. gap)	Married (9 yr. gap)	Sibling
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.051	0.076	0.083	0.077	0.022
	(0.005)	(0.007)	(0.006)	(0.006)	(0.016)
Individual, Year FE	Yes	Yes	Yes	Yes	Yes
Control Group Mean	0.64	0.66	0.65	0.65	0.64
Observations	60,384	25,590	34,170	38,106	4,974
R^2	0.411	0.431	0.430	0.429	0.424
	Panel B: Total Income (log)				
	Full Sample	Married (3 yr. gap)	Married (6 yr. gap)	Married (9 yr. gap)	Sibling
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.048	0.082	0.088	0.077	−0.020
	(0.013)	(0.018)	(0.015)	(0.014)	(0.042)
Individual, Year FE	Yes	Yes	Yes	Yes	Yes
Control Group Mean	3.44	3.14	2.88	2.85	2.17
Observations	39,884	17,793	23,654	26,311	3,265
R^2	0.767	0.779	0.780	0.781	0.780

The outcome variables are a binary indicator for whether an individual was employed (Panel A) and their log income, conditional on being employed (Panel B). Column 1 is a basic model with individual fixed effects, identical to Column 1, Table 2. Column 2 measures marriage if the individual and the relative in federal office are of different genders and within 3 years of one another. Column 3 measures marriage if the individual and the relative in federal office are of different genders and within 6 years of one another. Column 4 measures marriage if the individual and the relative in federal office are of different genders and within 9 years of one another. Column 5 looks at siblings that share a patronymic and are within 15 years of age from each other. All models include year fixed effects, individual fixed effects and robust standard errors clustered on the individual and year levels. Control Group Mean measures employment rates and income (in thousands of dollars) when a relative was not in office.

2.5 Defining a Minimum Wage

- In the main text, I define employment as receiving any sort of income from an organization. The Pension Data used in the paper includes all payments from these firms, no matter the size. As a robustness check, I re-define this measure of employment to ensure that the relatives earned enough from their employer to be considered a full-time employee making at least minimum wage. This approach helps exclude any potential side payments that are not being made in exchange for real labor, but it also censors low-earning individuals.

- Table A12 presents results which only code relatives as employed if they earned any positive income, or at least 10,000-40,000 rubles per year, in ten year increments by panel. The results indicate that removing the bottom tail of the distribution does not affect the robustness of the results.

TABLE A12: INDIVIDUAL-LEVEL MECHANISMS: MINIMUM WAGE

	Panel A: Individual Employed, any positive income				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.042	0.034	0.032	0.026	0.026
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Control Group Mean	0.64	0.64	0.64	0.64	0.64
Observations	81,024	81,024	81,024	81,024	81,024
R^2	0.412	0.018	0.438	0.639	0.587
	Panel B: Individual Employed, at least 10,000 ruble wage				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.052	0.051	0.042	0.032	0.032
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Control Group Mean	0.63	0.63	0.63	0.63	0.63
Observations	68,334	68,334	68,334	68,334	68,334
R^2	0.413	0.022	0.439	0.638	0.588
	Panel C: Individual Employed, at least 20,000 ruble wage				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.051	0.051	0.043	0.032	0.032
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Control Group Mean	0.63	0.63	0.63	0.63	0.63
Observations	60,384	60,384	60,384	60,384	60,384
R^2	0.411	0.022	0.437	0.638	0.587
	Panel D: Individual Employed, at least 40,000 ruble wage				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.051	0.051	0.043	0.032	0.032
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	No	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Individual Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	0.64	0.64	0.64	0.64	0.64
Observations	60,384	60,384	60,384	60,384	60,384
R^2	0.411	0.022	0.437	0.638	0.587

The outcomes in each of the panels are a binary indicator for whether an individual was employed. Each panel uses a different cut-off for defining who was working as a full-time federal employee.

2.6 Tax Evasion: Splitting the Sample By Time

- One alternative explanation for the positive effect on employment of having a relative enter federal office could be connected to Russia's informal sector. Individuals may be more likely to declare their income and employment, thus exiting the shadow economy, if they have family members working for the government. The positive point estimates then would reflect increased voluntary compliance with tax law, rather than any corrupt ties.

- Addressing this possibility runs into the clear difficulty of measuring tax evasion using administrative data. Validated approaches in the literature often compare official income to consumption patterns, largely using micro-level survey data. The various datasets used in this paper do not allow for this, given that the Pension Fund data in particular only tracks income, and not expenditures. As a second best option, I take advantage of an institutional change happening halfway through the analysis period that changed the incentive for voluntary compliance: the introduction of a flat tax on personal income in 2001. Gorodnichenko, Martinez-Vazquez, and Sabirianova Peter (2009) among other demonstrate using survey data that the tax reform reduced evasion and moved many people out of the shadow economy.

- We might expect then that if the nepotism results are driven by a similar dynamic, the magnitude of the point estimates should drop after tax reform was passed. More and more people, and employers, would have declared formal income during the final three years of the period because of the institutional change. Having a relative serve in the federal government would have less of an impact overall on reporting behavior.

- In Table A13, I split the sample evenly between the years 1999-2001 (pre-tax reform) and 2002-2004 (post-tax reform). As a reference point, I show the base model estimates in Column 1. The results do not indicate any drop in the size of the point

estimates after tax reform was introduced. I interpret this as evidence that nepotistic relationships bestow new labor market opportunities, and ultimately facilitate corrupt transactions between employers and the state, rather decreasing tax evasion.

TABLE A13: INDIVIDUAL-LEVEL MECHANISMS: SPLITTING THE SAMPLE BY TIME

	Panel A: Individual Employed		
	Full Sample	1999-2001	2002-2004
	(1)	(2)	(3)
Relative in Federal Office	0.051	0.036	0.034
	(0.005)	(0.008)	(0.008)
Year FE	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
Control Group Mean	0.64	0.64	0.63
Observations	60,384	30,192	30,192
R^2	0.411	0.596	0.644
	Panel B: Income (log)		
	Full Sample	1999-2001	2002-2004
	(1)	(2)	(3)
Relative in Federal Office	0.048	0.037	0.041
	(0.013)	(0.020)	(0.022)
Year FE	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
Control Group Mean	4.41	1.59	7.86
Observations	39,884	19,849	20,035
R^2	0.767	0.852	0.850

The outcome variable in Panel A is a binary indicator for whether an individual was employed in a given year. The outcome variable in Panel B is the IHS-transformed total annual income for the individual in a given year. Linear probability models (OLS) are used for all models. The main predictor of interest is a count of the number of federal employees related to the individual individual that are holding federal office in that year. All models include year fixed effects, individual fixed effects and robust standard errors clustered on the individual and year level. Control Group Mean displays the base employment rates and annual income (measured in thousands of dollars) when a federal employee was not in office.

3 Individual-level: Placebo Analysis, Banking Connections

3.1 Sample Construction

One potential concern that could undermine the difference-in-differences design is that there are still family-specific trends that the main model specifications do not absorb. In other words, there could be other network elements at play whereby having any relative enter employment in a large Moscow-based organization could affect one's individual labor market outcomes. In this section, I describe a placebo approach whereby I use the same sample construction techniques used to build the analysis sample, but instead of identifying relatives entering federal government, I identify individuals with family members entering or exiting work in one of Russia's leading banks. The presence of similar labor market effects for having a relative working in the financial sector would call into question whether the design is truly picking up unique nepotistic relationships.

To build a suitable sample of employees working in the financial sector, I first use the Pension Fund Data to identify all private banks that are based in Moscow and have employee entries for all six years in the dataset. For example, I excluded the Russian Central Bank (since that it is part of the federal government) and Sberbank (since as the successor to the Savings Bank of the USSR, the bank is responsible for multiple types of non-salary payments in the dataset). The goal is to create a sample of employers that at least through ownership connections do not have connections to the federal government.

Next, I ranked the private banks by the number of unique employees they hired over the period and created a list of the twenty largest banks by employment. Then using these tax identification numbers, I identified entries in the Pension Fund data for all individuals working for them. The same procedures were used to narrow down the sample to individuals working in full-time positions, based in Moscow, and who moved in an out of their position at the banks from 1999-2004. Then I merged the Pension Fund data with the housing registration data to identify individuals who had relatives working in the fi-

nancial step. The last step was to build an analysis sample at the individual-year level to run the placebo checks in the same manner as that used for the federal government employees.

The placebo analysis below shows results based on these 20 largest banks, while also showing a robustness check for individuals working in the top 10 banks. Several summary statistics on this sample are given in Table A14 and Table A15. Table A14 gives all the characteristics of the individual-years included in the analysis (top 20 banks), while Table A15 provides summary statistics at the bank level about the employees working there during the period.

TABLE A14: BANK EMPLOYEES: INDIVIDUAL-YEAR-LEVEL SUMMARY STATISTICS

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Individual is Female	54,336	0.374	0.484	0	0	1	1
Individual Age	54,336	42.392	13.414	16	32	52	104
Individual Age Difference from Relative in Federal Office	54,336	−5.133	17.909	−68	−23	3	47
Relative in Federal Office is Father	54,336	0.126	0.332	0	0	0	1
Relative in Federal Office is Son or Daughter	54,336	0.301	0.459	0	0	1	1
Relative in Federal Office is Spouse (3 yr. diff)	54,336	0.421	0.494	0	0	1	1
Relative in Federal Office is Spouse (6 yr. diff)	54,336	0.573	0.495	0	0	1	1
Relative in Federal Office is Spouse (9 yr. diff)	54,336	0.573	0.495	0	0	1	1
Relative in Federal Office	54,336	0.390	0.488	0	0	1	1
Individual Employed	54,336	0.673	0.469	0	0	1	1
Individual Salary (log)	36,556	10.694	1.388	7.601	9.747	11.616	19.798
Individual Works in Private Firm	35,083	0.514	0.500	0	0	1	1
Individual Works in SOE	35,083	0.105	0.306	0	0	0	1
Individual Works for Federal Government	35,083	0.197	0.398	0	0	0	1
Individual Works for Regional Government	35,083	0.127	0.332	0	0	0	1
Individual Works for Local Government	35,083	0.002	0.042	0	0	0	1

This table gives summary statistics at the individual-year level for individuals with relatives working in the financial sector in Moscow during the period. The variables describing place of employment subset to only individuals that are employed.

TABLE A15: BANK EMPLOYEES: BANK-LEVEL SUMMARY STATISTICS

Bank	Employees	Mean Salary (ths. $)	Mean # of Relatives	Mean # of Kids	Married (%)	Female (%)
MDM Bank	1,220	26.4	1.24	0.19	65.00	36.39
Sovinkom Commercial Bank	1,007	8.97	1.19	0.14	60.97	46.57
Converse Bank	923	4.47	1.22	0.17	77.14	51.46
Nikoil Bank	612	62.3	1.15	0.07	56.70	51.14
Trast Bank	498	20.8	1.22	0.12	64.06	36.35
Sobinbank	383	5.63	1.24	0.14	66.58	46.48
Tandem Bank	296	12.6	1.22	0.19	66.22	36.82
Rosbank	285	38.1	1.19	0.12	60.70	36.49
Mosnar Bank	285	14.1	1.23	0.14	67.37	41.75
Moscow Credit Bank	281	148	1.15	0.06	61.21	52.31
Bank of Moscow	274	9.09	1.30	0.28	78.10	40.15
TransCredit Bank	251	8.47	1.18	0.08	63.35	48.61
Lafko Bank	233	15.9	1.24	0.23	61.80	27.04
Absolut Bank	207	33.9	1.18	0.07	57.49	41.55
Avangard Bank	200	21.4	1.21	0.18	67.00	36.00
MTS Bank	198	25.9	1.23	0.19	78.79	46.46
Alma Bank	194	25.7	1.26	0.22	70.62	32.99
MezhTrust Bank	53	16.2	1.11	0.13	67.92	47.17
Moscow Industrial Bank	46	2.2	1.35	0.30	71.74	52.17
Dialog Optim Bank	19	1.85	1.26	0.26	68.42	42.11

This table gives summary statistics by financial institution based in Moscow that has six years of employment data in the Pension Fund data. Note that the number of employees reflect only those that enter the placebo analysis sample, or roughly 25% of the total number of employees per year. The mean number of children is also underestimated since mother-child relationships cannot be precisely measured (patronymics cannot be used).

3.2 Empirical Analysis

- I use this sample of individuals to connect identical analysis about the labor market effects of having a relative enter or exit work in a private bank based in Moscow from 1999-2004. Table A16 shows results on employment and income (log) outcomes for those with relatives working in the 20 largest banks, while Table A17 subsets the same to only those with relatives working in the 10 largest banks.

- The results indicate that there is no relationship between having a relative enter a leading financial institution and one's own labor market prospects. The point estimates for both employment (Panel A) and log income (Panel B) are not consistently above zero or statistically significant. The models specifications used are identical to those of Table 2 in the main text, giving us stronger reason to believe that there is something unique about having a familial relation to a federal government employee, rather than someone working more broadly in the private sector in Moscow during this period.

TABLE A16: FAMILY LINKS TO BANK EMPLOYEES, TOP 20 BANKS

	Panel A: Individual Employed				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	−0.007	−0.003	−0.008	−0.0001	−0.0001
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	No	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	0.67	0.67	0.67	0.67	0.67
Observations	54,336	54,336	54,336	54,336	54,336
R^2	0.411	0.014	0.433	0.630	0.588
	Panel B: Income (log)				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.010	0.027	0.009	0.009	0.007
	(0.014)	(0.017)	(0.014)	(0.017)	(0.016)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Individual FE	Yes	No	Yes	Yes	Yes
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	2.92	2.92	2.92	2.92	2.92
Observations	36,556	36,556	36,556	36,556	36,556
R^2	0.765	0.157	0.770	0.886	0.869

The outcome variables are a binary indicator for whether an individual was employed (Panel A) and their log income, conditional on being employed (Panel B). All models are linear probability. Column 1 includes individual fixed effects, Column 2 controls for individual age and drops fixed effects, Column 3 adds age fixed effects, and Columns 4 and 5 includes linear 'individual' time-trends and 'family' time-trends, respectively. Control Group Mean measures employment rates and income (in thousands of dollars) when a relative was not in office. The sample includes all individuals who have a relative enter or exit one of the 20 largest private financial institutions (by employment) in Moscow during the period.

TABLE A17: FAMILY LINKS TO BANK EMPLOYEES, TOP 10 BANKS

	Panel A: Individual Employed				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	−0.005	−0.002	−0.006	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	No	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	0.67	0.67	0.67	0.67	0.67
Observations	39,924	39,924	39,924	39,924	39,924
R^2	0.412	0.014	0.433	0.631	0.589
	Panel B: Income (log)				
	(1)	(2)	(3)	(4)	(5)
Relative in Federal Office	0.007	0.027	0.005	0.010	0.005
	(0.016)	(0.020)	(0.016)	(0.019)	(0.019)
Demographic Controls	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Age FE	No	No	Yes	No	No
Individual FE	Yes	No	Yes	Yes	Yes
Relative Time Trends	No	No	No	Yes	No
Family Time Trends	No	No	No	No	Yes
Control Group Mean	2.97	2.97	2.97	2.97	2.97
Observations	26,790	26,790	26,790	26,790	26,790
R^2	0.762	0.156	0.767	0.883	0.867

The outcome variables are a binary indicator for whether an individual was employed (Panel A) and their log income, conditional on being employed (Panel B). All models are linear probability. Column 1 includes individual fixed effects, Column 2 controls for individual age and drops fixed effects, Column 3 adds age fixed effects, and Columns 4 and 5 includes linear 'individual' time-trends and 'family' time-trends, respectively. Control Group Mean measures employment rates and income (in thousands of dollars) when a relative was not in office. The sample includes all individuals who have a relative enter or exit one of the 10 largest private financial institutions (by employment) in Moscow during the period.

4 Robustness: Firm-Level

4.1 Including Utilities

- Appendix Section 1.5 described the process of cleaning the transaction data to hone in on only those payments made for goods and services. In the main text, I excluded all firms operating as utilities since the payments they receive from government entities for water and electricity are not competitively allocated. Therefore, we might not expect them to depend significantly on utilities having political connections to officials. Table A18 shows analysis that includes the small number of utilities (150). The point estimates are robust to their inclusion.

TABLE A18: FIRM-LEVEL MECHANISMS: EXCLUDING UTILITIES

	Panel A: Received Government Contract			
	(1)	(2)	(3)	(4)
No. of Individuals with Relatives in Federal Office	0.018	0.017	−0.006	0.079
	(0.004)	(0.007)	(0.007)	(0.024)
No. of Individuals with Relatives in Federal Office * Firm Age				−0.007
				(0.003)
Organization, Year FE	Yes	Yes	Yes	Yes
Organization Time Trends	No	Yes	No	No
Employment Controls	Yes	Yes	Yes	Yes
Control Group Mean	0.47	0.47	0.84	0.50
Observations	30,799	30,799	3,679	30,799
R^2	0.602	0.741	0.683	0.602
	Panel B: Amount of Government Contracts			
	(1)	(2)	(3)	(4)
No. of Individuals with Relatives in Federal Office	0.113	0.122	0.068	0.401
	(0.033)	(0.049)	(0.033)	(0.250)
No. of Individuals with Relatives in Federal Office * Firm Age				−0.031
				(0.024)
Organization, Year FE	Yes	Yes	Yes	Yes
Organization Time Trends	No	Yes	No	No
Employment Controls	Yes	Yes	Yes	Yes
Control Group Mean	163.62	163.62	1391.82	125.39
Observations	15,317	15,317	3,218	15,317
R^2	0.718	0.863	0.851	0.718

The outcome variables are a binary indicator for whether an organization received contracts from the federal government (Panel A) and the log amount of contracts received, conditional on a contract being given. Column 1 is a reduced-form model that only includes private firms. Column 2 adds firm-specific time trends. The sample is subset to only state-owned enterprises in Column 3. Column 4 subsets again to only private firms while adding an interaction for age. All models control for the total number of employees an organization has in each year. Standard errors are clustered on the firm (organization) and year level. Both outcome variable exclude payments from tax agencies and banks, but include employers (firms) who operate as utilities.

4.2 Robustness Tests

- The main outcome variable to investigate why private firms hire the family members of public officials measured the level of contracts from the federal government that these companies could hope to earn. An additional empirical implication holds that hiring family members of federal officials might not translate into financial gains when dealing with regional and local governments. After all, federal officials may not have the access or opportunities to influence decision-making over contracts at these lower levels, however prominent they may be. Looking at different levels of government contracting helps ensure that federal connections are actually translating into real benefits.

- The first two columns in Table A19 present the results. The outcome variables are binary indicators for a company receiving any money from regional or local government institutions, respectively; the sample is limited to private companies. Having a relative hold federal office appears to help firms secure contracts from regional governments, but not local governments.

- This table also presents two robustness checks for the main results. First, Column 3 subsets the data to only years where the private company appeared in the wage data to ensure consistency across the analyses. The main result of having more relatives still holds. Secondly, Column 4 presents the actual empirical results used to create Figure 2 in the main test.

TABLE A19: CONTRACTS: ROBUSTNESS TESTS

	Regional Gov.	Local Gov.	Federal Gov.	
	(1)	(2)	(3)	(4)
No. of Individuals with Relatives in Federal Office	0.008	0.001	0.015	0.015
	(0.005)	(0.003)	(0.005)	(0.006)
No. of Individuals with Relatives in Federal Office * Construction				0.015
				(0.031)
No. of Individuals with Relatives in Federal Office * Heavy Manufacturing				−0.031
				(0.014)
No. of Individuals with Relatives in Federal Office * Light Manufacturing				−0.007
				(0.015)
No. of Individuals with Relatives in Federal Office * Mining / Agriculture				−0.020
				(0.076)
No. of Individuals with Relatives in Federal Office * Trade/Transportation				0.019
				(0.010)
Organization, Year FE	Yes	Yes	Yes	Yes
Employment Controls	Yes	Yes	Yes	Yes
Control Group Mean	0.33	0.05	0.47	0.47
Observations	30,725	30,725	23,213	30,725
R^2	0.635	0.505	0.643	0.602

	Regional Gov.	Local Gov.	Federal Gov.	
	(1)	(2)	(3)	(4)
No. of Individuals with Relatives in Federal Office	0.122	−0.025	0.074	0.083
	(0.060)	(0.151)	(0.037)	(0.069)
No. of Individuals with Relatives in Federal Office * Construction				0.145
				(0.284)
No. of Individuals with Relatives in Federal Office * Heavy Manufacturing				0.077
				(0.131)
No. of Individuals with Relatives in Federal Office * Light Manufacturing				−0.098
				(0.119)
No. of Individuals with Relatives in Federal Office * Mining / Agriculture				−0.053
				(0.713)
No. of Individuals with Relatives in Federal Office * Trade/Transportation				0.070
				(0.085)
Organization, Year FE	Yes	Yes	Yes	Yes
Employment Controls	Yes	Yes	Yes	Yes
Control Group Mean	50.68	0.26	163.83	163.83
Observations	10,687	1,517	13,795	15,280
R^2	0.766	0.789	0.740	0.718

The outcome variables for Columns 1 and 2 are binary indicators for whether a private company received any money from regional and local governments, respectively. Column 3 looks only at private companies were wage data was available for the years they entered the analysis. Column 4 interacts the main predictor of interest with a categorical variable indicating sectoral membership. The reference category for Column 4 is 'light manufacturing'. All models only include private companies, as well as firm and year effects.

4.3 All Individuals

- In the main text, I use the analysis sample to calculate the number of individuals with relatives working in the federal government that are employed in each organization. In other words, this calculation is done using information on those individuals who are connected to federal employees that move in and out of government during the period, the key driver behind the identification strategy. This approach excludes all federal employees that worked for the government the entire period, since we cannot employ fixed effects in the model specifications.

- Table A20 brings the individuals connected to these 'always employed' back into the analysis. The predictor thus includes all individuals with any relatives working in federal government, whether or not they were employed continuously or came in and out. The specifications are identical to those in Table 5, but the sample is larger because there are more employers entering the sample, given the larger number of politically connected individuals. The results are robust to this different sample construction. Firms that hire individuals with more family members in government are more likely to win federal contracts, and conditional on winning access, earn larger volumes in transfers. Younger firms are again more likely to benefit from this strategy.

TABLE A20: CONTRACTS: ALL INDIVIDUALS

	Panel A: Received Government Contract			
	(1)	(2)	(3)	(4)
No. of Individuals with Relatives in Federal Office	0.018	0.015	−0.004	0.097
	(0.004)	(0.006)	(0.007)	(0.021)
No. of Individuals with Relatives in Federal Office * Firm Age				−0.009
				(0.002)
Organization, Year FE	Yes	Yes	Yes	Yes
Organization Time Trends	No	Yes	No	No
Employment Controls	Yes	Yes	Yes	Yes
Control Group Mean	0.46	0.46	0.83	0.50
Observations	34,325	34,325	3,862	34,325
R^2	0.602	0.741	0.681	0.602
	Panel B: Amount of Government Contracts			
	(1)	(2)	(3)	(4)
No. of Individuals with Relatives in Federal Office	0.108	0.100	0.075	0.438
	(0.032)	(0.045)	(0.030)	(0.209)
No. of Individuals with Relatives in Federal Office * Firm Age				−0.035
				(0.021)
Organization, Year FE	Yes	Yes	Yes	Yes
Organization Time Trends	No	Yes	No	No
Employment Controls	Yes	Yes	Yes	Yes
Control Group Mean	158.84	158.84	1196.02	120.03
Observations	16,940	16,940	3,378	16,940
R^2	0.719	0.864	0.843	0.720

The outcome variables are a binary indicator for whether an organization received contracts from the federal government (Panel A) and the log amount of contracts received, conditional on a contract being given. Column 1 is a reduced-form model that only includes private firms. Column 2 adds firm-specific time trends. The sample is subset to only state-owned enterprises in Column 3. Column 4 subsets again to only private firms while adding an interaction for age. All models control for the total number of employees an organization has in each year. The sample includes all individuals with relatives working in federal government, regardless of whether they entered the analysis sample.

References

Ansolabehere, Stephen, and Eitan D Hersh. 2017. "Adgn: An Algorithm For Record Linkage Using Address, Date Of Birth, Gender, And Name." *Statistics and Public Policy* 4 (1): 1–10.

Braguinsky, Serguey, and Sergey Mityakov. 2015. "Foreign Corporations and the Culture of Transparency: Evidence from Russian Administrative Data." *Journal of Financial Economics* 117 (1): 139–164.

Braguinsky, Serguey, Sergey Mityakov, and Andrey Liscovich. 2014. "Direct Estimation of Hidden Earnings: Evidence from Russian Administrative Data." *Journal of Law and Economics* 57 (2): 281–319.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter. 2009. "Myth And Reality Of Flat Tax Reform: Micro Estimates Of Tax Evasion Response And Welfare Effects In Russia." *Journal of Political Economy* 117 (3): 504–554.

Guriev, Sergei, and Andrey Rachinsky. 2008. "The Evolution of Personal Wealth in the Former Soviet Union and Central and Eastern Europe." In *Personal Wealth from a Global Perspective*, ed. James B. Davies. Oxford: Oxford University Press.

Mironov, Maxim. 2013. "Taxes, Theft, and Firm Performance." *The Journal of Finance* 68 (4): 1441–1472.

Mironov, Maxim, and Ekaterina Zhuravskaya. 2016. "Corruption in Procurement and the Political Cycle in Tunneling: Evidence from Financial Transactions Data." *American Economic Journal: Economic Policy* 8 (2): 287–321.

Szakonyi, David. 2016. "Businesspeople in Elected Office: Identifying Private Benefits from Firm-Level Returns." *American Political Science Review* 11 (2): 322–338.