ONLINE APPENDIX

Appendix A: Hypothesis Testing in Difference-in-Difference-in-Differences (DDD) Specification

The analyses analyze three differences: before versus after the time of the awarding of an innovation grant for each cohort of firms ("Post"), fraudulent versus honest firms ("Fraud"), and grant recipients versus denied grant applicants ("Win"). The regressions therefore implement a difference-in-difference-in-differences (DDD) design. In formal terms, we begin with the equation:

$$Y_{it} = \alpha + \beta_1 Post_t + \beta_2 Fraud_i + \beta_3 Win_i + \beta_4 Post_t * Fraud_i + \beta_5\, Post_t * Win_i + \beta_6\, Fraud_i * Win_i + \beta_7 Post_t * Fraud_i * Win_i + B_8 X_i + \varepsilon_{it} \qquad (1)$$

where Y_{it} measures either the resource allocation choices or innovation outcomes of firm *i* at time *t,* depending on the specification; Win_i is a binary variable set equal to 1 for all firms that receive an Innofund grant; $Post_t$ is a dummy that equals 1 for the post-grant time period; $Fraud_i$ is an indicator that equals 1 for fraudulent firms; X_i is a vector of control variables; and ε_{it} is the error term.[1]

Hypothesis I posits a difference between fraudulent and honest firms among Innofund grant winners. This is a difference-in-differences (DD) that is embedded in the DDD specification above. We calculate the value of the DD in three steps. First, we calculate the difference in recruiting behaviors (and patenting outcomes) among fraudulent grant winners between the pre-grant and the post-grant period. Second, we calculate the difference in recruiting behaviors among honest grant winners between the pre-grant and the post-grant periods. Third, we compute the difference between these two differences.

Step 1. Calculate the difference in recruiting behaviors among fraudulent grant winners between the pre-grant and the post-grant periods. We denote fraudulent winners by the two letters *fw*, and we denote before and after the time of grant with *pre* and *post*.

$$\begin{aligned} Y_{fw,post} - Y_{fw,pre} &= (\alpha + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + B_8 X_{fw} + \varepsilon_{fw,post}) \\ &\quad - (\alpha + \beta_2 + \beta_3 + \beta_6 + B_8 X_{fw} + \varepsilon_{fw,pre}) \\ &= \beta_1 + \beta_4 + \beta_5 + \beta_7 + \varepsilon_{fw,post} - \varepsilon_{fw,pre} \end{aligned}$$

Step 2. Calculate the difference in recruiting behaviors among honest grant winners between the pre-grant and the post-grant eras. We refer to honest winners by the two letters *hw*, and we use *pre* and *post* as before.

$$\begin{aligned} Y_{hw,post} - Y_{hw,pre} &= (\alpha + \beta_1 + \beta_3 + \beta_5 + B_8 X_{hw} + \varepsilon_{hw,post}) - (\alpha + \beta_3 + B_8 X_{hw} + \varepsilon_{hw,pre}) \\ &= \beta_1 + \beta_5 + \varepsilon_{hw,post} - \varepsilon_{hw,pre} \end{aligned}$$

[1] Note that for X, we have only pre-grant values that do not vary across time.

Step 3. Calculate the difference between these two differences:

$$
\begin{aligned}
(Y_{fw,post} - Y_{fw,pre}) - (Y_{hw,post} - Y_{hw,pre})& \\
&= (\beta_1 + \beta_4 + \beta_5 + \beta_7 + \varepsilon_{fw,post} - \varepsilon_{fw,pre}) \\
&- (\beta_1 + \beta_5 + + \varepsilon_{hw,post} - \varepsilon_{hw,pre}) \\
&= \beta_4 + \beta_7 + (\varepsilon_{fw,post} - \varepsilon_{fw,pre} - \varepsilon_{hw,post} + \varepsilon_{hw,pre})
\end{aligned}
$$

Because we have repeat, annual observations for each firm in the sample, it is possible to estimate the equation above with firm fixed effects. In fixed-effects specifications, all covariates and interaction terms that do not vary within-firm will be fully absorbed into the time-stationary fixed effects and therefore will drop from the estimations. After removing all time-stationary variables that are subsumed in the firm fixed effects (including the firm attributes in the X covariate vector), the final model specification can be written:

$$(Y_{fw,post} - Y_{fw,pre}) - (Y_{hw,post} - Y_{hw,pre}) = \beta_4 + \beta_7 + \varepsilon$$

To test whether fraudulent grant winners utilize Innofund grants differently than honest grant winners in recruiting, we can use Wald tests to investigate whether $\beta_4 + \beta_7 = 0$.

We can also examine the difference in the effect of capital infusions on firm innovation outcomes between fraudulent and honest grant winners when we use fraudulent and honest non-winners, respectively, as the benchmark. To do this, we must consider the three-way interaction effect.

We begin with the estimating equation:

$$
\begin{aligned}
Y_{it} = \alpha + \beta_1 Post_t + \beta_2 Fraud_i + \beta_3 Win_i + \beta_4 Post_t{*}Fraud_i + \beta_5 Post_t{*}Win_i + \beta_6 Fraud_i{*}& \\
Win_i + \beta_7 Post_t{*}Fraud_i{*}Win_i + B_8 X_i + \varepsilon_{it}& \qquad (2)
\end{aligned}
$$

We can first compute values for the difference-in-differences of grant winning on fraudulent and honest firms, respectively, and then calculate the difference-in-difference-in-differences.

For instance, to test the prediction that honest firms that win an Innofund grant place more employee recruitment ads than fraudulent grant winners do (using their within-type, non-winners as benchmarks to establish the trend), we can first calculate the difference-in-differences for honest firms that win Innofund grants vs. honest firms that do not win Innofund grants. We can then calculate the difference-in-differences for fraudulent firms that win Innofund grants vs. fraudulent firms that do not win Innofund grants. These two DDs allow us to calculate the final value for the DDD to gauge the impact of state grant funding on honest firms vs. fraudulent firms (using same-type, non-winners as a benchmark, rather than within-winner comparisons). We again require three steps:

Step 1: We first calculate the DD for fraudulent firms, using the notations *fw* for fraudulent winner, *fn* for fraudulent non-winner, and *pre* and *post* as before.

$$
\begin{aligned}
DD_{fraudulent,\ win} &= (Y_{fw,post} - Y_{fw,pre}) - (Y_{fn,post} - Y_{fn,pre}) \\
&= [(\alpha + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + B_8X_{fw} + \varepsilon_{fw,post}) \\
&\qquad - (\alpha + \beta_2 + \beta_3 + \beta_6 + B_8X_{fw} + \varepsilon_{fw,pre})] \\
&\qquad - [(\alpha + \beta_1 + \beta_2 + \beta_4 + B_8X_{fn} + \varepsilon_{fn,post}) \\
&\qquad - (\alpha + \beta_2 + B_8X_{fn} + \varepsilon_{fn,pre})] \\
&= \beta_5 + \beta_7 + \varepsilon_{fraudulent,\ pre\text{-}post}
\end{aligned}
$$

Step 2: We then calculate the DD for honest firms, once again using the notations *hw* for honest winners, *hn* for honest non-winners, and *pre* and *post*.

$$
\begin{aligned}
DD_{honest,\ win} &= (Y_{hw,post} - Y_{hw,pre}) - (Y_{hn,post} - Y_{hn,pre}) \\
&= [(\alpha + \beta_1 + \beta_3 + \beta_5 + B_8X_{hw} + \varepsilon_{hw,post}) - (\alpha + \beta_3 + B_8X_{hw} + \varepsilon_{hw,pre})] \\
&\qquad - [(\alpha + \beta_1 + B_8X_{hn} + \varepsilon_{hn,post}) - (\alpha + B_8X_{hn} + \varepsilon_{hn,pre})] \\
&= \beta_5 + \varepsilon_{honest,\ pre\text{-}post}
\end{aligned}
$$

Step 3: We can now calculate the difference-in-differences-in-differences.

$$
\begin{aligned}
DDD_{fraudulent,\ honest,\ win} &= DD_{fraudulent,\ win} - DD_{honest,\ win} \\
&= (\beta_5 + \beta_7 + \varepsilon_{fraudulent,\ pre\text{-}post}) - (\beta_5 + \varepsilon_{honest,\ pre\text{-}post}) \\
&= \beta_7 + \varepsilon
\end{aligned}
$$

Table A1. Pre- and Post-grant Descriptive Statistics across Different Types of Firms

A. Recruitment	**Total Positions**				**R&D-related Positions**				**Non-R&D Positions**			
	Honest		**Fraudulent**		**Honest**		**Fraudulent**		**Honest**		**Fraudulent**	
	Unfunded	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**
Pre-grant era	4.814	5.587	3.408	4.247	.804	1.119	.350	.575	4.011	4.468	3.058	3.672
Post-grant era	8.259	16.707	5.755	9.986	1.063	2.354	.437	1.028	7.196	14.354	5.318	8.958
Difference	3.445^{+}	11.120$^{\bullet\bullet\bullet}$	2.347$^{\bullet}$	5.739$^{\bullet\bullet\bullet}$	.259	1.235$^{\bullet\bullet}$	.087	.453$^{\bullet}$	3.186$^{\bullet}$	9.885$^{\bullet\bullet\bullet}$	2.260$^{\bullet}$	5.286$^{\bullet\bullet\bullet}$

B. Patenting	**Total Patents**				**Invention Patents**				**Utility Patents**			
	Honest		**Fraudulent**		**Honest**		**Fraudulent**		**Honest**		**Fraudulent**	
	Unfunded	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**	**Unfunded**	**Funded**
Pre-grant era	4.188	7.081	2.757	4.810	2.509	4.040	1.350	2.314	1.679	3.040	1.408	2.497
Post-grant era	5.813	11.374	3.961	8.013	2.777	5.990	1.573	2.712	3.036	5.384	2.388	5.301
Difference	1.625	4.293^{+}	1.204	3.203$^{\bullet\bullet}$	.268	1.950^{+}	.223	.399	1.357	2.343^{+}	.981	2.804$^{\bullet\bullet}$

$^{+}$ $p < .10$; $^{\bullet}$ $p < .05$; $^{\bullet\bullet}$ $p < .01$; $^{\bullet\bullet\bullet}$ $p < .001$ for two-tailed tests.

Appendix B. Recruitment and Patenting Trends across Groups

The validity of difference-in-difference estimators hinges on the assumption that treated and control units were following the same trend in the outcome variable in the pre-period, regardless of the levels of the outcome. To assess the DD assumption, we generate figures that illustrate the trend lines for all of the outcome variables in our analyses, across the distinctions of fraudulent vs. honest firms and of grant-winning vs. non-winning firms. We hope to see that companies exhibit similar pre-trends (versus levels; the assumption is of parallel trend lines, not of identical levels).

Figure B1 illustrates the pre- and post-award trends for all recruitment and patent variables. Corresponding to hypotheses 1 and 3, panel A compares fraudulent grant winners to honest winners. Corresponding to hypotheses 2 and 4, panel B compares fraudulent grant winners vs. fraudulent non-winners.

For the patent-derived outcome variables, there is general similarity in pre-trends in the years preceding treatment. This confirms the identifying assumption. Likewise, there is visual evidence of differences between the pre- and post-grant period in the outcome variables, which is consistent with one of our central hypotheses: honest companies exhibit an increase in the rate of invention (important) patents after receiving an Innofund grant relative to the time window before the grant (panel C, image V), but the same cannot be said for fraudulent firms (panel B, image V). Conversely, image VI in the panels shows the rate of change in filing much less significant utility patents. Here, consistent with hypothesis 4, the difference favors fraudulent winning companies (panel B) rather than honest counterparts (panel C).

Eyeballing Figure B1, the pre-trends assumption appears to be more questionable for recruiting activities, especially for non-R&D-related jobs. It is possible to assess the parallel trends assumption statistically. Specifically, one can interact the treatment variable (i.e., "Win_i" for hypotheses 2 and 4, and "$Fraud_i$" for hypotheses 1 and 3) with year dummy variables and then run regressions on the recruitment and patenting outcome variables. If the pre-trends between the treatment and control groups are comparable during the pre-treatment time window, the coefficients for the interaction terms for the pre-treatment time dummies should be insignificant (Autor, 2003).

Tables B1 and B2 show a consistent pattern of nonsignificance for the interaction terms between pre-treatment year dummies and treatment. There is only one exception: the interaction of time period T_{0-1}*Fraud for the subsample of grant-winning firms (fraudulent vs. honest companies) for the sole outcome variable, total jobs. Overall, both the pre-trend figures and the regressions show that the identifying assumption in the difference-in-difference estimator is valid for all other paired groups, namely that observational units are on similar trend lines prior to the onset of the treatment condition. Nonetheless, we should caution the reader about the total jobs regression results, although we address this point in the robustness section with a matching estimator that eliminates the problem.

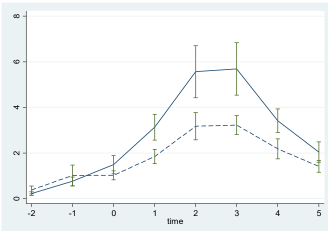
REFERENCE

Autor, D. H.
2003 "Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing." Journal of Labor Economics, 21: 1–42.

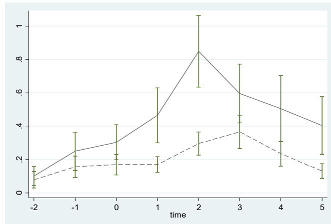
Figure B1. Recruitment and Patenting Trends across Groups

Panel A. Grant-winning Firms: Honest (Solid Line) vs. Fraudulent (Dashed Line)

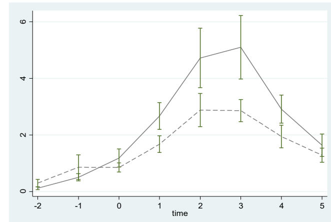
I. Total positions



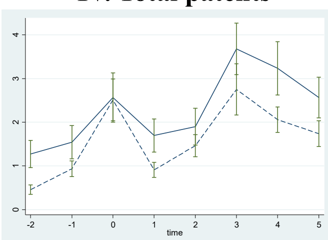
II. R&D positions



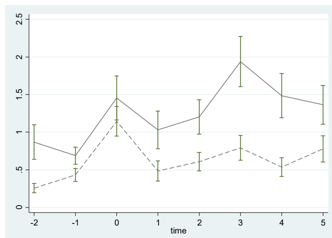
III. Non-R&D positions



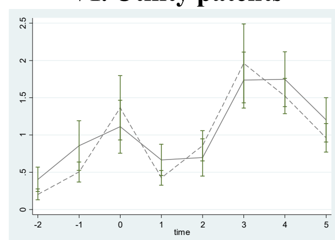
IV. Total patents



V. Inventions

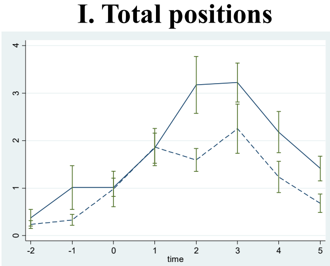


VI. Utility patents

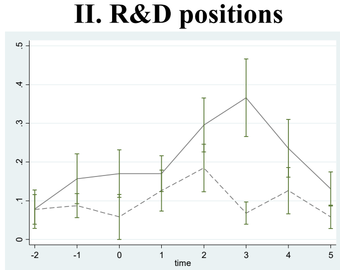


Panel B. Fraudulent Firms: Grant Winners (Solid Line) vs. Non-winners (Dashed Line)

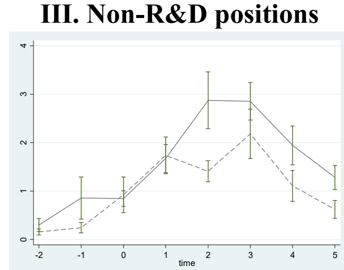
I. Total positions



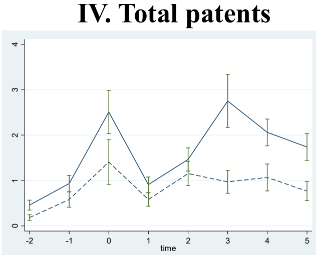
II. R&D positions



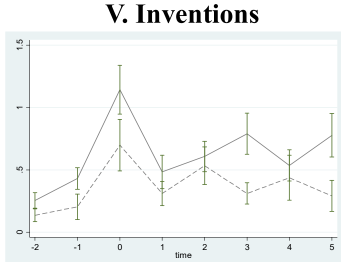
III. Non-R&D positions



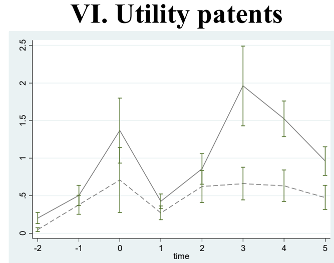
IV. Total patents



V. Inventions

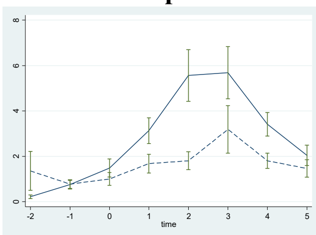


VI. Utility patents

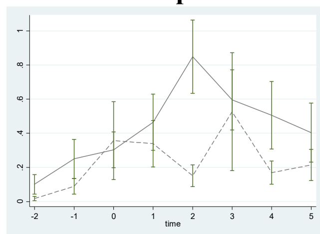


Panel C. Honest Firms: Grant Winners (Solid Line) vs. Non-winners (Dashed Line)

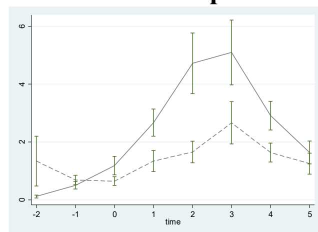
I. Total positions



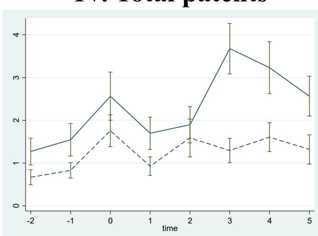
II. R&D positions



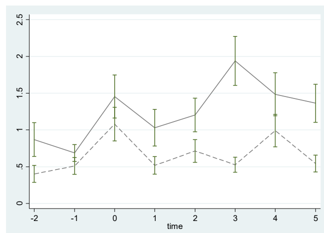
III. Non-R&D positions



IV. Total patents



V. Inventions



VI. Utility patents

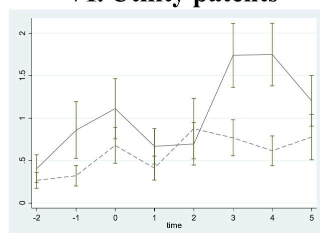


Table B1. Empirical Testing of the Parallel-trend Assumption for Recruitment (T_{0+1} as the Benchmark Period)

	Total Positions		R&D Positions		Non-R&D Positions	
	Winners: honest vs. fraudulent (H1)	Fraudulent: winners vs. non-winners (H2)	Winners: honest vs. fraudulent (H1)	Fraudulent: winners vs. non-winners (H2)	Winners: honest vs. fraudulent (H1)	Fraudulent: winners vs. non-winners (H2)
T_{0-2}*win		.165		–.043		.208
		(.661)		(.112)		(.634)
T_{0-1}*win		.705		.026		.679
		(.661)		(.112)		(.634)
T_0 *win		.058		.068		–.010
		(.661)		(.112)		(.634)
T_{0+2}*win		1.602$^{\bullet}$		.068		1.534$^{\bullet}$
		(.661)		(.112)		(.634)
T_{0+3}*win		.994		.254$^{\bullet}$		.740
		(.661)		(.112)		(.634)
T_{0+4}*win		.968		.065		.903
		(.661)		(.112)		(.634)
T_{0+5}*win		.755		.029		.726
		(.661)		(.112)		(.634)
T_{0-2}*fraud	1.446		.272		1.174	
	(.884)		(.192)		(.835)	
T_{0-1}*fraud	1.550^{+}		.202		1.349	
	(.884)		(.192)		(.835)	
T_0*fraud	.821		.162		.659	
	(.884)		(.192)		(.835)	
T_{0+2}*fraud	–1.105		–.258		–.847	
	(.884)		(.192)		(.835)	
T_{0+3}*fraud	–1.175		.065		–1.240	
	(.884)		(.192)		(.835)	
T_{0+4}*fraud	.055		.025		.030	
	(.884)		(.192)		(.835)	
T_{0+5}*fraud	.662		.021		.640	
	(.884)		(.192)		(.835)	
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2016	2048	2016	2048	2016	2048
log likelihood	–6.0e+03	–5.6e+03	–3.0e+03	–1.9e+03	–5.9e+03	–5.5e+03

$^{+}$ $p < .10$; $^{\bullet}$ $p < .05$.

Table B2. Empirical Testing of the Parallel-trend Assumption for Patenting (T_{0+1} as the Benchmark Period)

	Total Patents		Inventions		Utility Patents	
	Winners: honest vs. fraudulent (H3)	**Fraudulent: winners vs. non-winners (H4)**	**Winners: honest vs. fraudulent (H3)**	**Fraudulent: winners vs. non-winners (H4)**	**Winners: honest vs. fraudulent (H3)**	**Fraudulent: winners vs. non-winners (H4)**
T_{0-2}*win		–.829		–.326		–.503
		(.544)		(.241)		(.463)
T_{0-1}*win		–.750		–.217		–.533
		(.544)		(.241)		(.463)
T_0 *win		–.776		–.272		–.504
		(.544)		(.241)		(.463)
T_{0+2}*win		–.793		–.371		–.422
		(.544)		(.241)		(.463)
T_{0+3}*win		.679		.035		.643
		(.544)		(.241)		(.463)
T_{0+4}*win		–.111		–.346		.235
		(.544)		(.241)		(.463)
T_{0+5}*win		–.130		.042		–.172
		(.544)		(.241)		(.463)
T_{0-2}*fraud	–.759		–.303		–.456	
	(.600)		(.295)		(.485)	
T_{0-1}*fraud	–.555		.055		–.610	
	(.600)		(.295)		(.485)	
T_0 *fraud	–.733		–.236		–.497	
	(.600)		(.295)		(.485)	
T_{0+2}*fraud	–.379		–.283		–.096	
	(.600)		(.295)		(.485)	
T_{0+3}*fraud	–.869		–.838$^{\bullet\bullet}$		–.031	
	(.600)		(.295)		(.485)	
T_{0+4}*fraud	–1.118^{+}		–.638$^{\bullet}$		–.480	
	(.600)		(.295)		(.485)	
T_{0+5}*fraud	–.771		–.275		–.496	
	(.600)		(.295)		(.485)	
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2016	2048	2016	2048	2016	2048
log likelihood	–5.3e+03	–5.2e+03	–3.8e+03	–3.5e+03	–4.8e+03	–4.8e+03

$^{+}$ $p < .10$; $^{\bullet}$ $p < .05$; $^{\bullet\bullet}$ $p < .01$.

Appendix C

Table C1. Coarsened Exact Matching for DDD Specifications*

	Total Jobs (1)	R&D Jobs (2)	Non-R&D Jobs (3)	Total Patents (4)	Inventions (5)	Utility Patents (6)
Post × Fraud × Win (β_7)	-5.192^{+}	$-1.600^{\bullet\bullet\bullet}$	-3.592	1.740	-1.630^{+}	$3.370^{\bullet\bullet\bullet}$
	(2.734)	(.454)	(2.622)	(1.490)	(.881)	(.943)
Post × Win (β_5)	$10.332^{\bullet\bullet\bullet}$	$1.944^{\bullet\bullet\bullet}$	$8.388^{\bullet\bullet\bullet}$	$2.833^{\bullet}$	$2.486^{\bullet\bullet\bullet}$	.347
	(2.096)	(.348)	(2.010)	(1.194)	(.706)	(.755)
Post × Fraud (β_4)	-1.243	.339	-1.582	.131	.506	$-.376$
	(1.953)	(.324)	(1.873)	(1.066)	(.630)	(.674)
Post (β_1)	2.820^{+}	.006	2.814^{+}	.458	.056	.403
	(1.497)	(.249)	(1.436)	(.854)	(.505)	(.540)
Observations (firms)	380 (190)	380 (190)	380 (190)	262 (131)	262 (131)	262 (131)
Difference in pre–post changes between honest winners and honest non-winners: (β_5)	$+^{\bullet\bullet\bullet}$	$+^{\bullet\bullet\bullet}$	$+^{\bullet\bullet\bullet}$	$+^{\bullet\bullet}$	$+^{\bullet\bullet\bullet}$	$+$
Difference in pre–post changes between fraudulent winners and fraudulent non-winners: ($\beta_5 + \beta_7$)	$+^{\bullet\bullet}$	$+$	$+^{\bullet\bullet}$	$+^{\bullet\bullet\bullet}$	$+$	$+^{\bullet\bullet\bullet}$
Difference in pre–post changes between fraudulent winners and honest winners: ($\beta_4 + \beta_7$)	$-^{\bullet\bullet\bullet}$	$-^{\bullet\bullet\bullet}$	$-^{\bullet\bullet}$	$+^{+}$	$-^{+}$	$+^{\bullet\bullet\bullet}$

$^{+}$ $p < .10$; $^{\bullet}$ $p < .05$; $^{\bullet\bullet}$ $p < .01$; $^{\bullet\bullet\bullet}$ $p < .001$.
* Coefficient signs (rather than values) are reported for Wald tests.

Appendix D

Table D1. Firms with Political Connection vs. Firms without Political Connection

	Firms with Political Connections						Firms without Political Connections					
	Total jobs (1)	**R&D jobs (2)**	**Non-R&D jobs (3)**	**Total Patents (4)**	**Inventions (5)**	**Utility patents (6)**	**Total jobs (7)**	**R&D jobs (8)**	**Non-R&D jobs (9)**	**Total Patents (10)**	**Inventions (11)**	**Utility patents (12)**
Post × Fraud × Win (β7)	–4.181	–.069	–4.113	–4.336	–2.880$^{\bullet}$	–1.456	–3.718^{+}	–.549	–3.169	–.126	–1.241^{+}	1.115
	(3.515)	(.600)	(3.213)	(2.775)	(1.301)	(2.186)	(2.233)	(.429)	(2.111)	(1.401)	(.720)	(.987)
Post × Win (β5)	5.301^{+}	–.278	5.580$^{\bullet}$	5.040$^{\bullet}$	2.108^{+}	2.932	8.201$^{\bullet\bullet\bullet}$	1.232$^{\bullet\bullet\bullet}$	6.969$^{\bullet\bullet\bullet}$	2.408$^{\bullet}$	1.668$^{\bullet\bullet}$	.740
	(2.938)	(.501)	(2.685)	(2.319)	(1.087)	(1.828)	(1.573)	(.302)	(1.487)	(.987)	(.507)	(.695)
Post × Fraud (β4)	.221	.242	–.022	3.394	1.688	1.706	–1.316	–.253	–1.063	–1.052	–.369	–.683
	(2.792)	(.476)	(2.552)	(2.204)	(1.033)	(1.737)	(1.578)	(.303)	(1.492)	(.990)	(.509)	(.698)
Post (β1)	2.636	.091	2.545	–.727	–.545	–.182	3.533$^{\bullet\bullet\bullet}$	.277	3.255$^{\bullet\bullet}$	1.881$^{\bullet\bullet}$	.356	1.525$^{\bullet\bullet}$
	(2.262)	(.386)	(2.067)	(1.785)	(.837)	(1.407)	(1.056)	(.203)	(.998)	(.663)	(.341)	(.467)
Constant	–3.969	.094	–4.062	–1.156	.219	–1.375	–1.108	–.012	–1.096	.085	.506	–.421
	(3.866)	(.660)	(3.534)	(3.052)	(1.431)	(2.405)	(5.340)	(1.025)	(5.048)	(3.350)	(1.723)	(2.360)
Observations	204	204	204	204	204	204	730	730	730	730	730	730

$^{+}$ $p < .10$; $^{\bullet}$ $p < .05$; $^{\bullet\bullet}$ $p < .01$; $^{\bullet\bullet\bullet}$ $p < .001$.