Appendix A: Data

Sampling

The Chinese dataset used in this paper was collected in year 2001 and 2003 under the Investment Climate Surveys by the World Bank. These surveys cover more than 26,000 firms in 53 developing countries, and aim to understand how investment climates vary around the world and how they influence growth and poverty. The sample of firms in each country is stratified by size, sector and location.

The original Chinese samples provide annual observations for up to 3 years in the period 1998-2000 for 1548 firms distributed across 5 cities (Beijing, Chengdu, Guangzhou, Shanghai and Tianjin), and in the period 2000-2002 for 2400 firms distributed across 18 cities (Dalian, Benxi, Changchun, Ha'erbin, Hangzhou, Wenzhou, Nanchang, Zhengzhou, Wuhan, Changsha, Shenzhen, Jiangmen, Nanning, Chongqing, Guiyang, Kunming, Xi'an and Lanzhou). Overall these 3948 firms were distributed across 10 manufacturing industries (auto and auto parts, biotech products and Chinese medicine, electronic equipment, chemical products and medicine, electronic parts making, food processing, garment and leather products, household electronics, metallurgical products, and transportation equip) and 4 services industries (accounting and non-banking financial service, advertisement and marketing, business services, and information technology).

Compared with potential alternatives such as the China's Industrial Survey dataset, the dataset that we use has two important advantages. First, the sample of firms is stratified by size and therefore includes a large number of small firms. These firms are of particular interest because according to conventional wisdom, they are most likely to face an unfavorable investment climate in developing countries. However, such firms are excluded from the Industrial Survey, which only samples firms whose sales revenue exceeds 5 million Chinese Yuan. Second, the Investment Climate Survey explicitly asks about the investment expenditure on fixed assets for each firm in each year. Accurate information on investment is crucial in estimating any investment model. However, such information is not available in the Industrial Survey.

Data cleaning

Firms with number of permanent employees less than 10 or larger than 1000 are dropped to rule out extremely small or large observations. Using information from *China Statistical Yearbook*, we then deflate investment and capital stock data using province-specific price indices of investment in fixed assets and deflate profit and sales data using province-specific ex-factory price indices of industrial products. After calculating the five key variables in ratio or growth rate according to definition in Section 3.1, we trim the top and bottom 5% observations to rule out extreme values. The final sample used for estimation is a three-year panel of 3618 firms, with median number of employees 112.

Macroeconomic background over sample period

TABLE A1

Macroeconomic indicators for China over 1998 to 2002

	1998	*1999*	*2000*	*2001*	*2002*
GDP (PPP) share of world total	10.2	10.6	10.9	11.5	12.1
real GDP growth rate (secondary industry)	8.9	8.1	9.4	8.4	9.8
real GDP growth rate (tertiary industry)	8.3	7.7	8.1	8.4	7.5
inflation rate, GDP deflator	-0.8	-1.4	0.4	0.7	-0.8
fixed capital formation as share of GDP	37.1	36.7	35.1	36.3	37.9

Notes: These numbers are in percentage.

Source: World Economic Outlook, IMF; World Development Indicator, World Bank; China Statistical Yearbook; National Bureau of Statistics of China.

Appendix B: Specification tests and robustness tests

Specification tests

Table A2 reports specification tests for alternative models. For reference, the preferred full model is listed in column (1). Columns (2) to (4) show what happens when we impose respectively no quadratic adjustment costs, no irreversibility and no fixed adjustment costs. Compared with the preferred full model, a model with quadratic adjustment costs together with either irreversibility (column (4)) or fixed adjustment costs (column (3)) could fit the data reasonably well. In other words, irreversibility and fixed adjustment costs are two alternative forms of non-convex adjustment costs and are substitutes for one another if we only allow for one of them. In contrast, a model without quadratic adjustment costs (column (2)) fits the data much worse and generates a substantially higher over-identifying restriction test statistic. This is because this restricted model cannot fit the large positive serial correlation in the investment rate. Furthermore, the simulated investment rate series is too volatile, too right-skewed and over-responsive.

Column (5) presents a model which assumes a homogeneous capital share β. As one may expect, without heterogeneity in β, the model cannot match the large dispersion and the high serial correlation in log sales to capital ratio. As a result, the model over-estimates adjustment costs, heterogeneity in growth rate and level of uncertainty to match these important features of the data. Column (6) illustrates the result of imposing no heterogeneity in the growth rate μ. Such a model, first, cannot fit the positive serial correlation of the sales growth rate; and second, over-estimates the quadratic adjustment costs. Comparison between columns (5), (6) and (1) therefore highlights the importance of allowing for unobserved heterogeneities in order to get consistent estimates for the adjustment costs.

TABLE A2

Specification tests

Parameters	*col (1)* *full*	*col (2)* $b_q = 0$	*col (3)* $b_i = 0$	*col (4)* $b_f = 0$	*col (5)* $\sigma_{log\beta}=0$	*col (6)* $\sigma_\mu = 0$	*col (7)* $\sigma_{meK}=0$	*col (8)* $\sigma_{log\beta}=0$ $\sigma_\mu=0$ $\sigma_{meK}=0$	*col (9)* *free* γ	*col (10)* *free* σ
b_q	1.532	**0.000**	1.964	1.142	3.405	1.730	2.592	2.785	2.121	2.623
b_i	0.370	0.211	**0.000**	0.428	0.295	0.236	0.446	0.447	0.282	0.248
b_f	0.011	0.000	0.077	**0.000**	0.019	0.060	0.045	0.079	0.029	0.047
ε	13.953	10.164	15.459	13.591	13.509	17.657	24.980	16.627	15.099	18.376
$\mu_{log\beta}$	-2.498	-2.496	-2.448	-2.494	-2.345	-2.387	-2.201	-2.358	-2.484	-2.446
$\sigma_{log\beta}$	1.386	1.265	1.346	1.376	**0.000**	1.381	1.263	**0.000**	1.350	1.393
μ_μ	0.087	0.091	0.086	0.090	0.089	0.090	0.082	0.070	0.088	0.089
σ_μ	0.089	0.000	0.044	0.092	0.141	**0.000**	0.063	**0.000**	0.089	0.087
σ_{meK}	0.522	0.580	0.517	0.525	0.500	0.494	**0.000**	**0.000**	0.546	0.595
σ	0.569	0.285	0.585	0.555	0.639	0.655	1.160	1.160	0.580	0.629
Simulated moments										
mean(π/Y)	0.218	0.227	0.215	0.218	0.163	0.222	0.225	0.145	0.211	0.213
mean(log(Y/Khat))	0.554	0.657	0.712	0.509	0.433	0.607	0.171	-0.118	0.602	0.600
mean(I/K)	0.151	0.171	0.148	0.156	0.168	0.153	0.122	0.110	0.152	0.157
mean(ΔlogY)	0.088	0.091	0.086	0.091	0.095	0.091	0.080	0.071	0.088	0.088
sd(log(Y/Khat))	1.278	1.238	1.272	1.261	1.047	1.277	1.064	0.706	1.267	1.311
sd(I/K)	0.190	0.310	0.181	0.204	0.200	0.182	0.131	0.106	0.189	0.199
sd(ΔlogY)	0.338	0.213	0.330	0.332	0.320	0.349	0.472	0.491	0.333	0.334
skew(log(Y/Khat))	0.189	0.005	0.155	0.172	-0.249	0.166	0.641	-0.104	0.190	0.188
skew(I/K)	2.553	4.223	2.394	2.768	1.838	2.369	1.277	0.750	2.529	2.795
skew(dlogY)	0.048	0.461	0.049	0.062	0.025	0.059	0.077	0.031	0.043	0.050
scorr(log(Y/Khat))	0.843	0.817	0.846	0.839	0.786	0.857	0.921	0.804	0.827	0.813
scorr(I/K)	0.492	0.091	0.360	0.484	0.592	0.345	0.740	0.662	0.474	0.432
scorr(ΔlogY)	0.014	0.001	-0.015	0.025	0.088	-0.021	-0.021	-0.035	0.014	0.012
corr(I/K, log(Y/K))	0.407	0.275	0.402	0.406	0.756	0.414	0.549	0.831	0.412	0.418
corr(ΔlogY, log(Y/K))	0.213	0.164	0.204	0.218	0.407	0.216	0.324	0.462	0.207	0.198
corr(I/K, ΔlogY)	0.446	0.675	0.508	0.469	0.379	0.530	0.533	0.586	0.419	0.384
Prop(I/K>0.2)	0.286	0.264	0.285	0.291	0.346	0.296	0.239	0.246	0.291	0.293
Prop(I/K=0)	0.289	0.340	0.306	0.255	0.339	0.292	0.315	0.288	0.295	0.292
Prop(I/K<0)	0.002	0.008	0.003	0.002	0.000	0.002	0.004	0.000	0.002	0.001
OI	1051	4091	1374	1068	4132	1445	5136	10848	1018	969

Column (7) shows the result of imposing no measurement error, which implies $\hat{\sigma}$ =1.160 by equation (20). Not surprisingly, this restricted specification is strongly rejected, mainly because the simulated sales growth rates are too volatile at such a high level of uncertainty. To dampen the sales growth rate, the model generates much higher estimates for the adjustment costs, which in turn makes the investment rate too persistent and not dispersed enough compared with the real data.

To investigate the overall effects of the empirical innovations introduced in this paper, column (8) re-estimates the model without allowing for any heterogeneity or measurement error. Most of the simulated moments are further away from their empirical counterparts and the model is clearly rejected. The estimated capital adjustment costs from such simple model are much higher than those from the benchmark specification.

Finally, instead of imposing $\bar{\gamma}$ =0.408, one may use $\hat{\gamma} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{1+\widehat{\beta_i}(\hat{\varepsilon}-1)}$ in equation (20) where $\widehat{\beta_i}$ and $\hat{\varepsilon}$ are estimated simultaneously with σ_{meK}. Column (9) implies that even greater capital adjustment costs are obtained under this alternative value of γ. Furthermore, to examine the effect of using the restriction (20) on the estimates, column (10) reports a model which estimates σ_{meK} and σ simultaneously without using (20) at all. The alternative identification restriction is to assume that the noise-to-signal ratio is no larger than one, that is $\sigma_{meK} \leq \sigma$. Such a model estimates even larger measurement error and a higher level of uncertainty. The estimates for capital adjustment costs, the key parameters of interest, are also more substantial than those in the benchmark specification.

Robustness tests

Table A3 presents robustness checks across three different parameters. Column (1) is the benchmark model, where δ = 0.03, r = 0.14 and ρ = 0.885. Columns (2) and (3) show the results for the same model but imposing the depreciation rate to be 0.02 and 0.04 respectively. Columns (4) and (5) present the results for the same model but imposing the discount rate to be 0.13 and 0.15 respectively. Compared with the benchmark model, a model with lower depreciation rate or lower discount rate implies slightly higher quadratic adjustment costs and smaller demand elasticity, in order to match the empirical mean of log sales to capital ratio. Nevertheless, with the exception of the estimates for the mean of growth rate μ_{μ}, the estimates for other parameters are robust to the choice of depreciation rate and discount rate within the range we considered. Columns (6) and (7) show what happens when the same model is estimated but with serial correlation of 0.85 and 0.92, respectively. As expected, a model imposing higher serial correlation implies less heterogeneity in the growth rate. However, there is no significant difference between estimates reported in these two columns and the estimates in the benchmark model.

TABLE A3

Specification tests

Parameters	col (1) benchmark	col (2) δ = 0.02	col (3) δ = 0.04	col (4) r= 0.13	col (5) r = 0.15	col (6) ρ = 0.85	col (7) ρ = 0.92
b_q	1.532	1.801	1.408	1.840	1.398	1.436	1.728
b_i	0.370	0.317	0.400	0.313	0.401	0.349	0.384
b_f	0.011	0.026	0.009	0.021	0.009	0.007	0.020
ε	13.953	13.183	15.794	13.199	14.244	13.431	13.142
$\mu_{log\beta}$	-2.498	-2.498	-2.490	-2.497	-2.499	-2.499	-2.499
$\sigma_{log\beta}$	1.386	1.371	1.394	1.370	1.387	1.365	1.386
μ_μ	0.087	0.094	0.080	0.085	0.084	0.086	0.089
σ_μ	0.089	0.083	0.090	0.086	0.085	0.090	0.087
σ_{meK}	0.522	0.523	0.514	0.523	0.521	0.531	0.521
σ	0.569	0.564	0.594	0.563	0.573	0.568	0.540
Simulated moments							
mean(π/Y)	0.218	0.220	0.213	0.220	0.217	0.218	0.222
mean(log(Y/Khat))	0.554	0.540	0.587	0.535	0.597	0.575	0.548
mean(I/K)	0.151	0.146	0.156	0.147	0.148	0.150	0.153
mean(ΔlogY)	0.088	0.094	0.081	0.085	0.085	0.087	0.089
sd(log(Y/Khat))	1.278	1.279	1.273	1.276	1.276	1.272	1.275
sd(I/K)	0.190	0.180	0.197	0.181	0.190	0.187	0.194
sd(ΔlogY)	0.338	0.338	0.341	0.337	0.338	0.341	0.327
skew(log(Y/Khat))	0.189	0.190	0.191	0.193	0.185	0.169	0.201
skew(I/K)	2.553	2.425	2.604	2.436	2.680	2.539	2.598
skew(dlogY)	0.048	0.040	0.058	0.040	0.054	0.045	0.051
scorr(log(Y/Khat))	0.843	0.841	0.846	0.840	0.842	0.835	0.845
scorr(I/K)	0.492	0.482	0.497	0.490	0.489	0.480	0.502
scorr(ΔlogY)	0.014	0.005	0.019	0.008	0.012	-0.007	0.040
corr(I/K, log(Y/K))	0.407	0.400	0.412	0.404	0.411	0.398	0.418
corr(ΔlogY, log(Y/K))	0.213	0.205	0.219	0.207	0.216	0.206	0.219
corr(I/K, ΔlogY)	0.446	0.436	0.459	0.431	0.455	0.433	0.451
Prop(I/K>0.2)	0.286	0.278	0.294	0.281	0.277	0.285	0.290
Prop(I/K=0)	0.289	0.305	0.280	0.296	0.288	0.285	0.292
Prop(I/K<0)	0.002	0.001	0.002	0.002	0.002	0.002	0.002
OI	1051	1052	1094	1046	1072	1028	1125

Appendix C: The effects of other factors

As illustrated by equation (9) and Figures 1a to 1c, optimal investment behavior is determined by five factors in this model: the Jorgensonian user cost of capital (J), production technology (β), demand schedule (ε), stochastic process characterizing investment opportunities (Z), and different forms of adjustment costs (b_q, b_i, b_f). The aggregate output loss in this paper is derived by reducing capital adjustment costs while keeping all other factors constant. Thus the effects simulated in Sections 6.1 and 6.2 are by nature comparative static analyses. It is therefore interesting to discuss how these effects might be, in a general equilibrium framework where user cost of capital, production technology, demand schedule and investment opportunities are all potentially endogenous in the long run.

First, when aggregate capital stock increases with the elimination or reduction of capital adjustment costs, the interest rate r will endogenously fall due to the general equilibrium effect. The will induce even more investment expenditure at the firm level until a lower *MRPK* is equalized with a lower Jorgensonian user cost of capital in the long run. Second, when the economy becomes more abundant in capital, the relative factor prices between production factors will change. If firms can endogenously choose production technology, it will be optimal for them to adopt a more capital intensive technology. This implies a higher value of β. Third, with lower capital adjustment costs the net profit of each firm will increase. In an environment with free entry, higher profit of the incumbents will bring more new comers and lead to a more competitive market structure. This implies a higher value of ε. According to equations (3), $1 - \gamma$ increases with both β and ε. Since the gain of aggregate output is an increasing function of $1 - \gamma$, endogenous technology and market structure will imply even larger gain from reducing investment frictions. Finally, if capital investment has positive externality or spillover effects, the growth rate of Z will endogenously increase with more aggregate capital stock. This implies a higher value of μ. Given that the aggregate output loss is also an increasing function of μ, the endogenous growth theory will predict even more aggregate output loss due to the presence of capital adjustment costs.

To sum, in a more general model where all other factors that affect investment decisions are allowed to change with an increase in aggregate capital stock, the aggregate output gain might be even greater than what has been derived from the comparative static analyses. Therefore, the effects reported in this paper could be taken as the lower bound of the true effects from a more general model. Estimating such a model is beyond the scope of this paper and will be an interesting task for future research.