Appendix: "Time to build and the real-options channel of residential investment"

Hyunseung Oh*and Chamna Yoon†

Contents

*Vanderbilt University. E-mail: hyunseung.oh@vanderbilt.edu.

†KAIST College of Business. E-mail: chamnayoon@kaist.ac.kr.

1 Micro data supplementary analysis

1.1 Definition of variables in Table 2

In this section, definitions of the variables used in the regression of Table 2 are provided. We follow the definitions provided in the Census Bureau's "Survey of Construction" webpage. For details, refer to their webpage.[1]

1. Construction category
    - Built for sale: All houses built on builder's land with the intention of selling the house and land in one transactions. Also called "speculatively-built" houses.
    - Contractor-built: All houses built for owner occupancy on the owner's land with construction under the supervision of a single general contractor.
    - Owner-built: All houses built for owner occupancy, on the owner's land, under the supervision of the owner acting as the general contractor.
    - Built for rent: All houses built on builder's land with the intention of renting the housing unit.
2. Census division states
    - New England: CT, MA, ME, NH, RI, VT
    - Middle Atlantic: NJ, NY, PA
    - East North Central: IL, IN, MI, OH, WI
    - West North Central: IA, KS, MN, MO, ND, NE, SD
    - South Atlantic: DC, DE, FL, GA, MD, NC, SC, VA, WV
    - East South Central: AL, KY, MS, TN
    - West South Central: AR, LA, OK, TX
    - Mountain: AZ, CO, ID, MT, NM, NV, UT, WY
    - Pacific: AK, CA, HI, OR, WA
3. Construction method

[1] https://www.census.gov/construction/chars/definitions/

- Modular: Finished 3-dimensional sections of the complete dwelling are built in a factory and transported to the site to be joined together on a permanent foundation.
- Panelized: A package of wall panels, roof trusses, and other components is shipped from a factory to be assembled on site. This may include all materials required to finish the house as a complete package.
- Site built: The house is built entirely on site, except that it may include some factory components as roof and floor trusses, wall panels, doorframes, etc.

4. Floor area (square foot): Completely finished floor space, including space in basements and attics with finished walls, floors, and ceilings. This does not include a garage, carport, porch, unfinished attic or utility room, or any unfinished area of the basement.

5. Metropolitan area: Whether or not the house is included in the metropolitan area.

6. Number of full bath: A full bathroom is one that has a washbasin, a toilet, and either a bathtub or shower.

7. Stories: Basement is not counted as a story even if it is finished, and 2+ story includes 1.5 stories, where living accommodations located wholly or partly within the roof frame is considered a half story.

8. Detached: Row houses, duplexes, quadruplexes, or townhouses may also defined as *attached* single-family houses if they (i) are separated by a ground-to-roof wall, (ii) have a separate heating system, (iii) have individual meters for public utilities, and (iv) have no units located above or below.

9. Deck: A deck house has a floored area without a roof, not sitting directly on the ground, typically made of wood or wood products.

10. Parking facility: (i) 1 car garage, (ii) 2 car garage, (iii) 3 or more car garage, or (iv) other

11. Foundation: (i) full or partial basement, (ii) crawl space, (iii) slab, or (iv) other (raised supports, earthen, etc)

12. Material of wall (framing material): (i) wood, (ii) steel, (iii) concrete/masonry other than insulated concrete forms, or (iv) insulated concrete forms (or SIP)

1.2 Statistical test on skewness of the distribution

Figure 4 of the paper provides the distribution of economic TTB in 2002, 2005, and 2009. Here we compute the sample equivalent of the skewness measure $\gamma = \mathbb{E}(X-\mu)^3/\sigma^3$ for each distribution. To argue the difference in skewness across the three distributions statistically, we bootstrap each distribution with replacement for 1000 times, and compute the sample skewness measure of each of the bootstrapped distribution. The skewness estimate as well as its 95 percent confidence interval for each of the three years are provided in Table A1. As observed, the 95 percent confidence intervals of the skewness measures do not overlap with each other.

1.3 Alternative measure of economic time to build

In this part, we redo the regression in Table 2 in the main text with time to build (TTB) regressed in levels instead of logs. The result is reported in table A2.

Figure A1 shows the average economic TTB based on a level regression of TTB. Economic TTB increases by 23 percent in 2008–2009, compared to its value in 2002.

The left panel of Figure A2 compares the kernel density of the alternative economic TTB measures in 2002 and 2005. Since the regression is now conducted in levels, there are scale differences in the range of economic TTB. In particular, a level TTB regression shows a larger range of residuals on the right tail. However, qualitatively, the description in the paper remains. We again observe that the overall distribution seems to have shifted to the right in between these two years.

The right panel of Figure A2 compares the kernel density of the alternative economic TTB measures in 2005 and 2009. The mass of the distribution including the mode shifting back to the left, and a fat tail appearing are also a pattern observed with the alternative economic TTB measure.

1.4 Permit to completion

Our model does not include the planning process for construction, activities such as purchasing a lot and applying for building permits. The overall construction process takes longer than the construction activity itself. Our data do include permit issuance months for housing units that require building permits. In the raw data, the time from permit to the completion of a project was about 7 months before 2002 and 10 months in 2008.

For houses requiring building permits, the survey of construction data also record the permit issue month. Typically, the period from permit to start takes about a month. However, the average permit to start period have also increased to about 2 months in 2008. Therefore, while average completion from start have increased from 6 months to 8 months in 2008, the average completion from permits have increased from 7 months to 10 months, which is a larger increase in the proportion. In Figure A3, we re-estimate our TTB measures (raw and economic) based on permit issuance month instead of start month and plot each measures as a percentage difference from 2002. Consistent with the rough observation, we find that the increase in average permit issuance to completion between 2002 and 2008 is even more in percentage than that in average start to completion in the same period. This suggests that overall construction activity including the planning stage may have slowed down in this period. Although further data is needed, we conjecture that the slowdown in construction may not be confined to TTB or time to start a permitted project, but may also extend to the time to convert a lot to a building permit, or the time to purchase a lot. Adding all these time up could be a significant slowdown in the construction industry. In that sense, our TTB measure (start to complete) is conservative in illustrating the delay in construction between 2002 and 2008.

2 Details on the new house price index

The Census Bureau's "Survey of Construction" also contains sales price data for built-for-sale houses. For a completed house in each year, the survey records the sales price of that house. However, sales might not have occurred within the completion year of the house. In that case, the survey also records house prices for houses that were completed in the previous year. This implies that houses that were completed in the previous year but sold in the current year would enter the data twice. Since the data is not a panel, we cannot handle this issue directly. That is, although we know that the house was completed in the previous year, we cannot match that price to the previously completed construction data. To prevent any potential double counting in our data analysis, we drop all house prices not observed within the completion year. Using this conservative approach, the coverage of sales price data is about 87 percent of the total built-for-sale completed house sample. That is, the total observation for built-for-sale houses that started since 2000 and completed between 2002 and 2011 (with all the building characteristics) is 111,628. Sales prices within the completion year are observed for 97,300 houses.

For the observed sales price, the census constructs the single-family price index for built-for-sale houses (including lot values) by running a hedonic regression.[2] That is, for the 4 census regions (Northeast, Midwest, South, West), a regression equation is estimated for all detached houses. For attached houses, a national regression is estimated. Therefore, there are 5 different regressions run for each period.

For each regression, the dependent variable is log sales price. Based on the public data, the regressors and their categories that we use for detached houses in each of the 4 regions are listed below:

1. Size of house (log of square feet).
2. Geographic location within region (census divisions for each region).
3. Metropolitan area: control only for the Midwest.
4. Number of bedrooms: (i) less than 2, (ii) 3, (iii) 4 or more.

[2] The exact regression is listed in the Census Bureau's Construction Price Indexes webpage (https://www.census.gov/construction/cpi/). Here, we use their regression formula adjusted for the publicly available data.

5. Number of bathrooms: (i) less than 3, (ii) 3 or more.

6. Number of fireplaces: (i) 0, (ii) 1, (iii) 2 or more.

7. Type of parking facility (number of garage): (i) 0, (ii) 1 or 2, (iii) 3 or more.

8. Type of foundation: (i) no basement or finished basement, (ii) unfinished basement.

9. Presence of deck.

10. Construction method:

    - Midwest: (i) stick-built, (ii) modular/ precut/panelized,
    - Northeast, South, West: none

11. Primary exterior wall material:

    - Northeast: (i) vinyl, (i) others,
    - Midwest: (i) vinyl, (i) wood, (i) others,
    - South: (i) brick, (i) stucco, (iii) vinyl, (iv) wood and others,
    - West: (i) wood, (ii) others.

12. Central air conditioning:

    - Northeast, West: (i) central air-conditioning, (ii) no central air-conditioning,
    - Midwest, South: none.

Similarly, the regressors and their categories that we use for attached houses in the U.S. are listed below:

1. Size of house (log of square feet).

2. Geographic location: 4 regions.

3. Metropolitan area.

4. Number of bedrooms: (i) less than 3, (ii) 3 or more.

5. Number of bathrooms: (i) less than 2, (ii) 2 or more.

6. Number of fireplaces: (i) 0, (ii) 1 or more.

7. Type of parking facility (number of garage): (i) 0, (ii) 1 or more.

8. Type of foundation: (i) no basement of finished basement, (ii) unfinished basement.

9. Presence of deck: (i) deck in the Northeast, (ii) deck in the Midwest, (iii) deck in the South, (iv) deck in the West and without deck for all regions.

10. Primary exterior wall material: (i) wood in the South, (ii) wood in the Northeast and West, (iii) vinyl, (iv) stucco, (v) wood in the Midwest and others.

11. Central air conditioning: (i) no central air-conditioning, (ii) central air-conditioning.

Based on the residuals of these regressions, we compute the quality-adjusted price dispersion within each quarter and take the annual average. For the price index, the census uses the average characteristic of a 2005 built house and computes the price index using the estimated parameters.[3]

Figure A4 plots the new house price index measures reported by the census and our constructed measure based on the public data and the regression method specified in this section. We find that our measure is similar to the census in its overall pattern and almost the same before 2005. The two measures depart in 2006 and afterwards with a 1–2 percentage point difference. The source of this difference could be partially due to our usage of the public data (e.g. crude construction characteristic category used compared to the census since we also only have access to the public data, and no separate regression for special geographic areas that the census controls for), or due to differences in the regression method since we use a straightforward regression method rather than some type of a resistant regression that the census uses.

[3] Specifically, the census uses a resistant regression which incorporates Tukey's biweight, to reduce the influence of unusual characteristics.

3 Numerical solution method

We describe the key steps in the numerical method that we use to solve the builder's problem in our estimation. Given all the calibrated parameters, we solve the builder's problem for a large number of the state variables by value function iteration. We guess the remaining 21 states (P^M_{2002}, $\{p_{c,T}, \sigma_T\}^{2011}_{T=2002}$) and solve each builder's problem in each regime by interpolating the value function calculated in the first step. Based on the solution, we simulate 20,000 builders for each regime as described in the paper. The simulated moments of interest are computed subsequently, and we compare that with the empirical moments. The estimation stage searches for a combination of 21 parameters that minimizes the objective function described in Section 4 of the paper.

3.1 Value function iteration

We start by describing the value function iteration procedure.

Step 1:

- Grid points

  1. We discretize 4 continuous state variables: the idiosyncratic price (P^U_i), the macro price (P^M), bottleneck (p_c), and uncertainty (σ).
  2. The log of the idiosyncratic price follows a random walk process (with drift adjustment), where the initial price draw is from 1 whenever a new construction is considered. The process could be well approximated by a lengthy state log-centered around 1. We set the minimum and maximum values of the idiosyncratic price as 0.05 and 20, and discretize the state by setting 100 price grid points with equal log spacing.
  3. For the other three aggregate state variables, we set the ranges of the grid space wide enough to cover the values that we need to evaluate during the SMM procedure. We describe our procedure to select the ranges of the grid space below. Due to the computational constraint, we consider 10 grid points for each of the aggregate state variables.

4. There are also 2 grid points for the bottleneck indicator $B_i \in \{0, 1\}$, and 5 grid points for remaining capital to completion $K_i \in \{1, 3/4, 2/4, 1/4, 0\}$. Therefore, the model in total has 1,000,000 grid points ($100 \times 10 \times 10 \times 10 \times 2 \times 5$).

- Extra grids

1. For the idiosyncratic price close to the end of each grid space, we still need to solve the value function beyond the minimum and maximum values that we set, for accuracy of the future expected value function near the end points that we set. This is because of the random walk structure of our idiosyncratic price process (Bloom, 2009). We extend the left and right ends of the state by twice the log-length from the log distance on each side with equal log distance for each grid points. We set 400 extra grid points beyond $[0.05, 20]$ where the value function is defined, with 200 grid points in $[2.7687e-7, 0.05)$ and 200 grid points in $(20, 3611743]$. The value function is linearly extrapolated at these extra grid points, based on the last two points in the grid.
2. For the three aggregate state variables, we extend the left and right ends of the state by the same log-length from the log distance on each side with equal log distance for each grid points. The value function is linearly extrapolated at these extra grid points as well.

- Transition probability

1. We approximate the transition probability of the idiosyncratic price following standard methods. For example, at each of the ith idiosyncratic price grid P_i^U, one needs to compute the transition probability on the idiosyncratic price grid vector. Since $P^U = \exp[\log P_i^U - 0.5\sigma^2 + \sigma W]$ where $W \sim N(0, 1)$, the transition probability to each price grid is computed by the sum the probability distribution function of the normal distribution between each midpoint of the adjacent price grids.
2. The transition probabilities of the three aggregate states are also approximated by the same method.

Step 2: With the above 100 grids for the price and 2 grids for the bottleneck, we solve each of the 10 value functions starting from the case when $K_i = 0$, which is the completion value of a house. The value at the completion stage is the total price of the structure. Based on these values, each value function is extrapolated for the other 400 price grids. Given these numbers, we solve the value function subsequently for $K_i = 1/4, 2/4, 3/4, 1$. Each value function is iterated until the maximum absolute distance between the previous and current value on all the grid points are below 1.0×10^{-7}.

3.2 SMM estimation

To generate the simulated data for each of the 10 regimes ($T = 2002, \cdots, 2011$), we simulate 10 economies each with 20,000 builders. Given a draw of the aggregate state variables, each builder solves its value function by linear interpolation, with its draw on the idiosyncratic price, its draw on bottleneck, and the remaining stage of construction. Each economy runs for 420 months (35 years). To prevent the bunching of investment decisions, we start the simulation by randomly assigning the 5 TTB stages to each builder. We discard the first 240 months to eliminate any effects coming from the initial conditions. For the aggregate parameters, based on a guessed initial value for P_{2002}^M we set $\{P_T^M\}_{T=2003}^{2011}$ from the real price growth rate of new construction price index divided by CPI net of shelter.

Weighting matrix. The weighting matrix is preset by the influence function technique of Bazdresch et al. (2018). In detail, we compute the asymptotic covariance matrix of the empirical moments directly used for the objective function of the SMM (mean economic TTB, economic TTB below 6m, price dispersion) for each year. Since our data observation (newly completed construction units) is not a panel, the across-year covariances are zero.

Our data is not a panel, and the empirical moments that we construct are based on each year's data observations. In constructing the weighting matrix, we correct for the sampling uncertainty that arises from differences in sample sizes across years. To simplify the discussion, let's assume there are two periods in our data, $t = 1, 2$. Let $\tilde{g}(\Lambda)$ denote the vector of our moment conditions $[m^d(x) - m^s(y; \Lambda)]$. Let $g_t(\Lambda)$ denote the subset of the moment conditions that are constructed based on period t

data. Let n_t denote the number of observations in period t. Let V_t be the asymptotic covariance matrix of $\sqrt{n_t}\ g_t(\Lambda)$. Then $\sqrt{n_1+n_2}\ \tilde{g}(\Lambda)$ can be written as follows:

$$\sqrt{n_1+n_2}\ \tilde{g}(\Lambda) = \begin{bmatrix} \frac{\sqrt{n_1+n_2}}{\sqrt{n_1}}\ \sqrt{n_1}\ g_1(\Lambda) \\ \frac{\sqrt{n_1+n_2}}{\sqrt{n_2}}\ \sqrt{n_2}\ g_2(\Lambda) \end{bmatrix}.$$

The asymptotic covariance matrix of $\sqrt{n_1+n_2}\ \tilde{g}(\Lambda)$ is given by

$$V = \begin{bmatrix} k_1\ V_1 & 0 \\ 0 & k_2\ V_2 \end{bmatrix},$$

where $(n_1+n_2)/n_t$ converges to k_t as $n_1+n_2 \to \infty$. Therefore, a consistent estimator of V is

$$\widehat{V} = \begin{bmatrix} \frac{n_1+n_2}{n_1}\ \widehat{V}_1 & 0 \\ 0 & \frac{n_1+n_2}{n_2}\ \widehat{V}_2 \end{bmatrix},$$

where $\widehat{V}_t$ is computed using influence function technique of Bazdresch et al. (2018). The covariance terms with price dispersion are computed for completed houses with sales prices within that year.

Our weighting matrix $\widehat{W}$ is the inverse of this asymptotic covariance matrix. As discussed in Bazdresch et al. (2018), one benefit of this weighting matrix is that it only uses the empirical moments and hence does not need to be calculated iteratively in each simulation step.

Inference. For parameter inference, we use large sample theory to approximate the joint sampling uncertainty in our target moments that are constructed using a 2 percent representative sample of the population (a source of sampling uncertainty of the census survey).[4] Since we are using an optimal weighting matrix in the estimation,

[4] Since we are running first step regressions to construct the economic TTB and price dispersion measures prior to the SMM estimation, one can also account for an additional source of sampling errors in our SMM parameter estimates due to the uncertainty in the estimated regression coefficients. In principle, a bootstrap method could be applied to account for this source of uncertainty. However, since the standard errors of most of the regression coefficients in Table 2 are small, the asymptotic covariance matrix for the parameter estimate is a decent approximation. Furthermore, a bootstrap method that also accounts for this source of uncertainty is infeasible due to the huge computational burden.

the covariance matrix for the parameter vector estimate $\hat{\Lambda}$ is given by

$$\frac{1}{N}\left(1+\frac{1}{J}\right)(G'\widehat{W}G)^{-1},$$

where $G = \partial m^s(y;\Lambda)/\partial\Lambda$, N is the number of total data observations, and J is the number of simulated observations per data. The term $(1 + 1/J)$ corrects for the simulation error (Bazdresch et al., 2018). For numerical values used in the paper, $N = 111,628$ (total number of completed houses) and $J = 16.3$ (16.3 times more completed houses simulated than the data).

3.3 Selection of the grid space range of the aggregate states

The model has three aggregate state variables: P^M, p_c, σ. Without any prior knowledge, the model needs to consider a very wide range of these state variables for estimation. At the same time, the grid points should not be too coarse for accuracy of the solution. However, it is not computationally feasible to have a wide range of state variables with grid points that are close to each other. For practical purposes, we need to make a guess on the ranges of the grid spaces and verify through the estimation that the estimates are interior to the grid spaces that we considered. In this section, we describe how we came up with the guess of the ranges of the grid spaces.

Towards this, we assume that the standard deviations of the 3 aggregate state variables described in Section 4 are all equal to zero ($\sigma_{P^M} = \sigma_{p_c} = \sigma_\sigma = 0$). In this extreme case, the 30 aggregate states ($\{P^M_T, p_{c,T}, \sigma_T\}_{T=2002}^{2011}$) are parameters in each regime rather than the realization of the stochastic processes. Therefore in each of the 10 regimes ($T = 2002, \cdots, 2011$), the builder solves a value function with regime-specific values of $P^M_T, p_{c,T}, \sigma_T$.

Instead of solving for the value function upfront, we solve the value function in each of the 10 regimes given a guess of the values of the three aggregate states. The value function in each regime only has idiosyncratic states and aggregate states are treated as parameters since their standard deviations are zero. Following the same simulation and estimation procedures as above, we estimate the 21 parameters in our model with no aggregate uncertainty. For each aggregate state, we compute the minimum and maximum values that are estimated. We then set the range of the grid to be -20 percent of the minimum value and +20 percent of the maximum value. If certainty equivalence holds, the estimated aggregate values with no aggregate uncertainty should be identical to the aggregate values in our estimation with aggregate uncertainty.

4 Additional details of the estimation results

4.1 Model and data distribution plot of economic TTB

In Figure A5, we plot the estimated model distribution of economic TTB and its data equivalent for each year.[5] The frequency of houses completed within 6 months is a target moment in the estimation. Overall, the model performs well in the subsequent distribution pattern of the data. In each year, the model tends to generate more houses with lower TTB (7-8 months) and less houses with longer TTB (10 months or more) compared to the data. Nevertheless, the evolution of the empirical distribution is still well captured.

4.2 Aggregate uncertainty and sensitivity analysis

We first describe our calibration of the standard deviation of the macro price level, σ_{PM}. The macro price that we use is the price index of new single-family houses sold including lot value (census), deflated by CPI without shelter (BLS). Since the census price index is computed in quarterly frequency, we also use the quarterly CPI. We take the log of this deflated price index and take the time difference between quarters to compute the real price growth rate in each quarter. We then take the standard deviation of this measure between 1963q2 and 2011q4. Taking the variance and adjusting the frequency to monthly, we obtain this standard deviation in monthly terms which is $\sigma_{PM} = 0.00929$.

The other standard deviation parameters σ_{p_c} and σ_σ cannot be calibrated by the data since there are no apparent data counterparts. In the benchmark estimation, we assumed $\sigma_{p_c} = \sigma_\sigma = \sigma_{PM}$. In this section, we discuss the sensitivity of this assumption by estimating the model under alternative calibrations. Tables A3-A6 display the estimation results under 4 alternative assumptions: (i) $\sigma_{p_c} = 0.5\sigma_{PM}$, $\sigma_\sigma = \sigma_{PM}$; (ii) $\sigma_{p_c} = 2\sigma_{PM}$, $\sigma_\sigma = \sigma_{PM}$; (iii) $\sigma_{p_c} = \sigma_{PM}$, $\sigma_\sigma = 0.5\sigma_{PM}$; (iv) $\sigma_{p_c} = \sigma_{PM}$, $\sigma_\sigma = 2\sigma_{PM}$. We find that in this range of standard deviation parameters, the model estimates barely changes and almost identical quantitative results could be reached.

[5] The empirical economic TTB distribution is discretized at each monthly bin. For example, TTB of 8 months refers to economic TTB between 7.5 months and 8.5 months in the data.

4.3 Price dispersion and TTB

In this section, we investigate how the real-options channels affect the price dispersion of complete houses. To isolate our model channels, we also solve a counterfactual fixed TTB model in which TTB is delayed only by bottlenecks, and builders with ongoing construction must continue investment until completion. Using the estimated values for price, uncertainty, and bottleneck, we compared the implied price dispersion of the two models.

Figure A6 plots the implied price dispersion in the two models. In each graph, one aggregate state variable is allowed to vary and the other two states are fixed at their 2002 value. The first column shows the price effects on price dispersion. The rise and fall in new house prices had small effects on price dispersion for both models. In fact, the fixed TTB model generates a constant price dispersion with regards to new house price changes. The second column shows the bottleneck effects on price dispersion. Higher bottleneck probability increases price dispersion since each house is exposed to longer periods of uncertainty. However, the bottleneck effects on price dispersion do not differ across the two models. The third column shows the uncertainty effects on price dispersion. In the fixed TTB model, there is a one-to-one relationship between uncertainty and price dispersion. On the other hand, the real-options TTB model generates a smoother response of price dispersion with regards to uncertainty shifts.

The key variable that drives the difference in price dispersion across the two models is uncertainty. In the fixed TTB model, builders are forced to complete a new house regardless of the evolution of its price during construction. Therefore, there is a complete pass-through of uncertainty into the price dispersion of completed houses. In the real-options TTB model, builders slow down the completion of the house when prices turn out to be lower than initially expected. When uncertainty is high, those builders with low price observations wait longer for a better price. As a result, the realized price dispersion for complete houses is smoother than uncertainty.

References

Bazdresch, S., R. J. Kahn, and T. M. Whited (2018). Estimating and Testing Dynamic Corporate Finance Models. *Review of Financial Studies 31*(1), 322–361.

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica 77*(3), 623–685.

List of Figures

Figure A1: Main text Figure 3 with alternative economic TTB

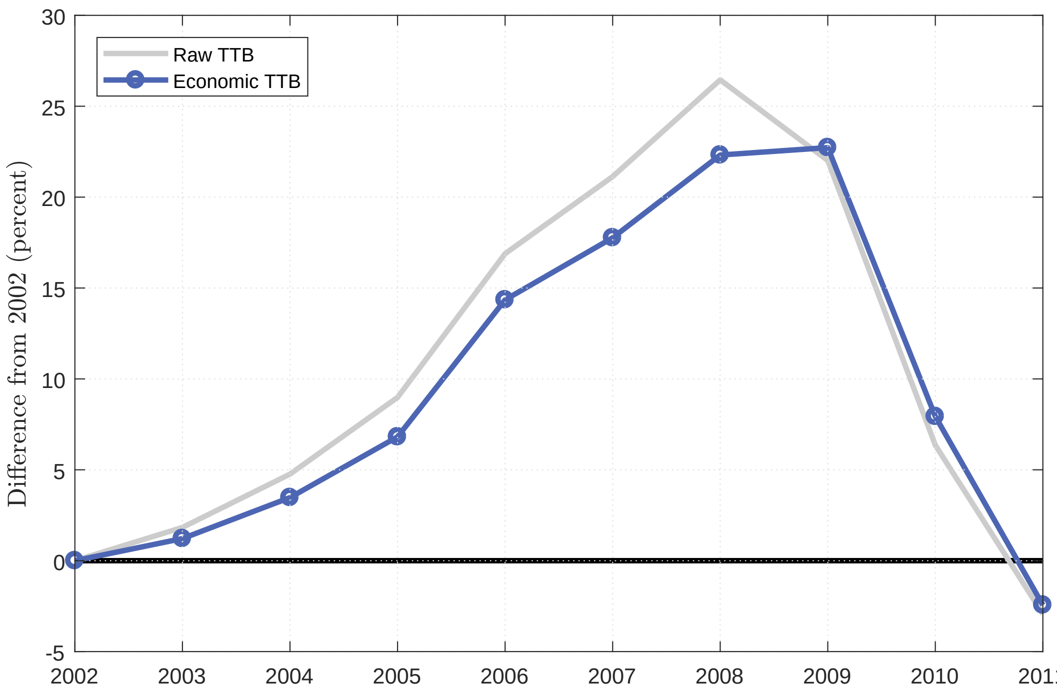
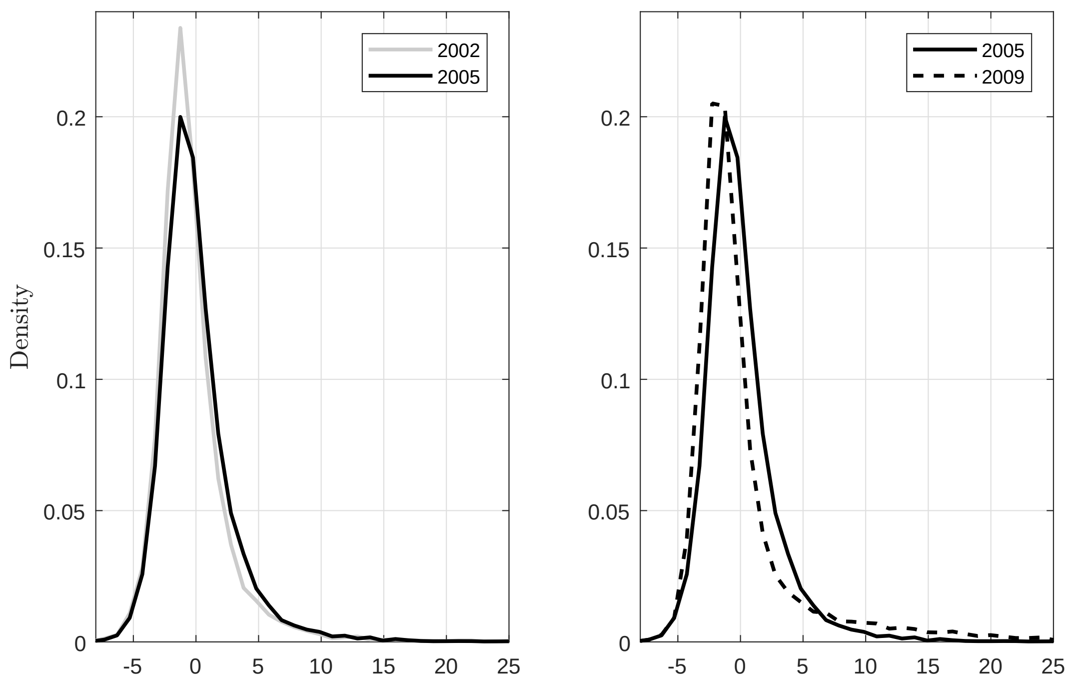
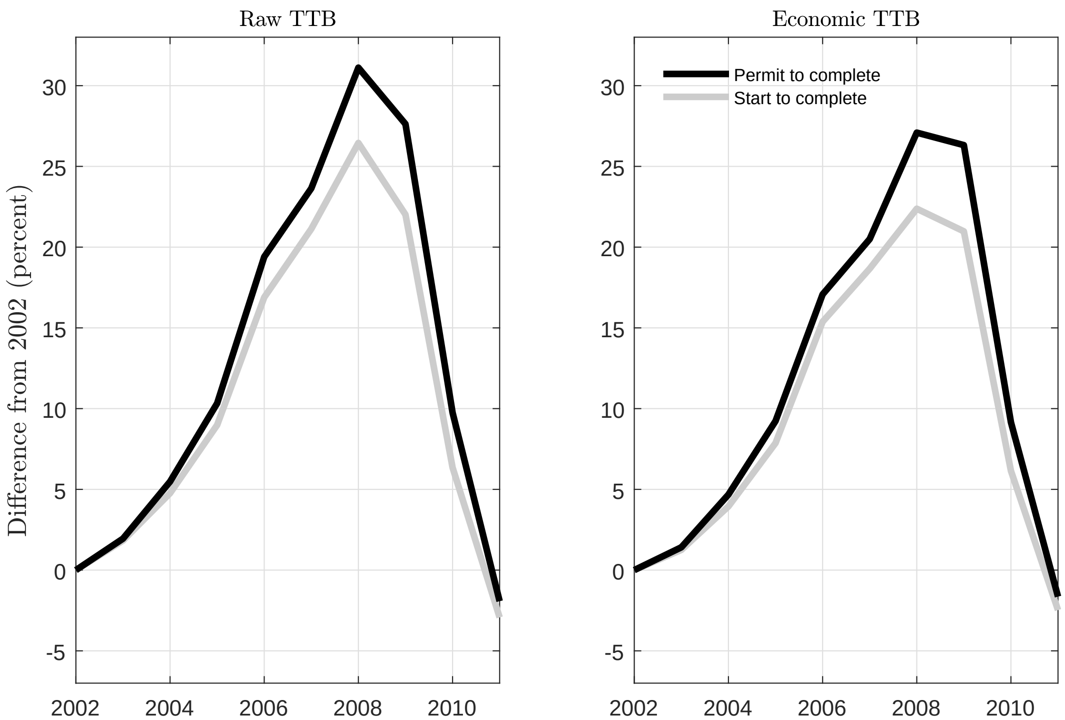


Figure A2: Main text Figure 4 with alternative economic TTB



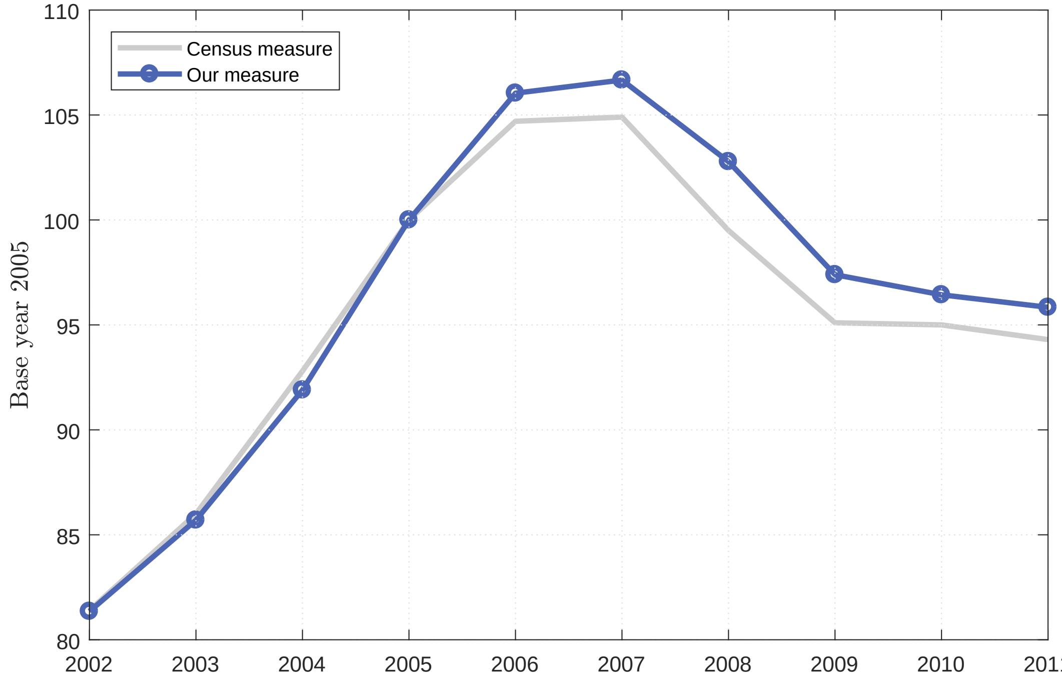
Note: Kernel density of economic TTB for total single-family houses. Left compares 2002 and 2005, and right compares 2005 and 2009.

Figure A3: Start to complete versus permit to complete



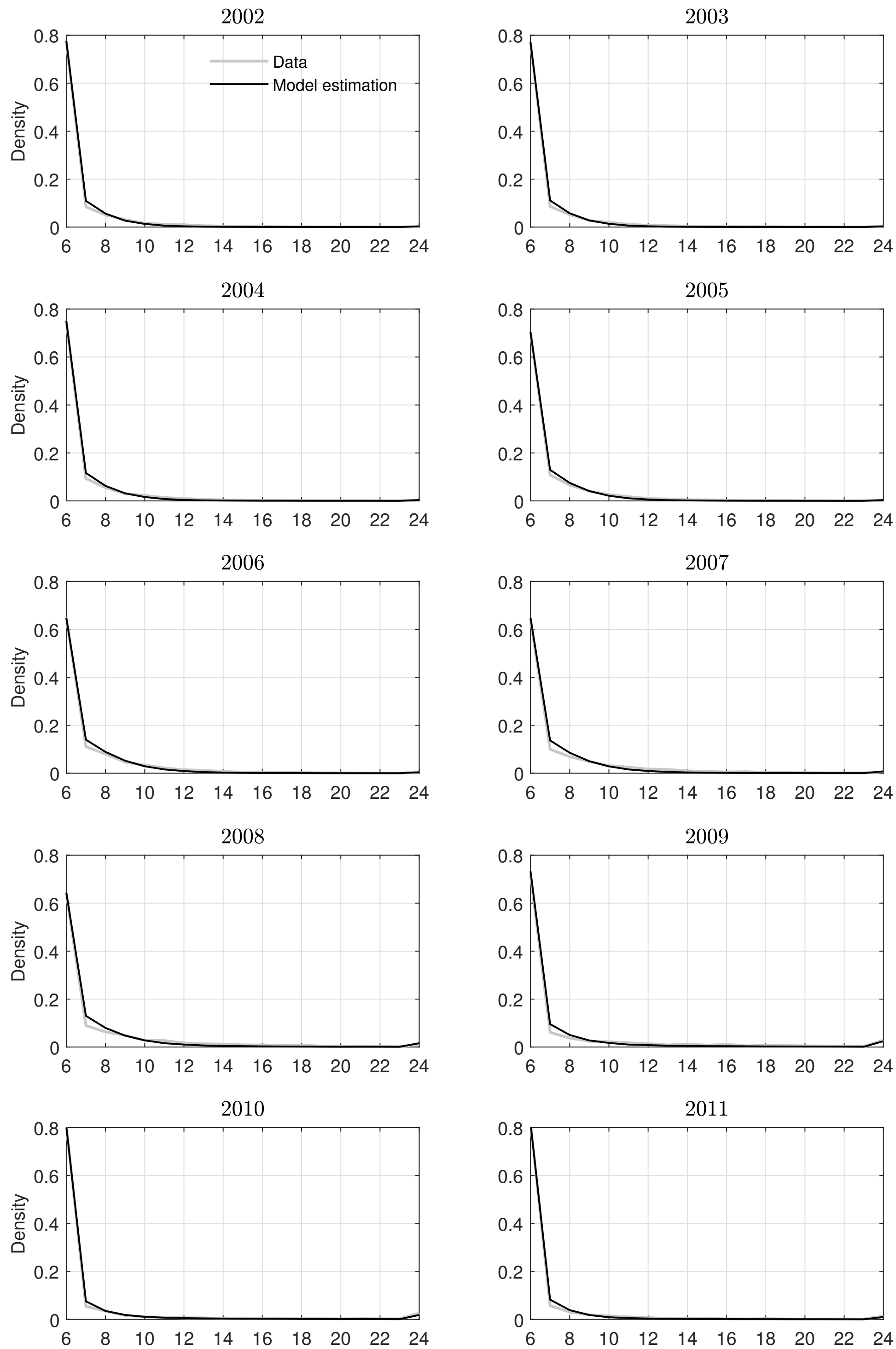
Note: *Start to complete* raw TTB and economic TTB are all the same as in Figure 3 in the main text. *Permit to complete* measures add permit to start periods to their respective *start to complete* measures.

Figure A4: New house price index



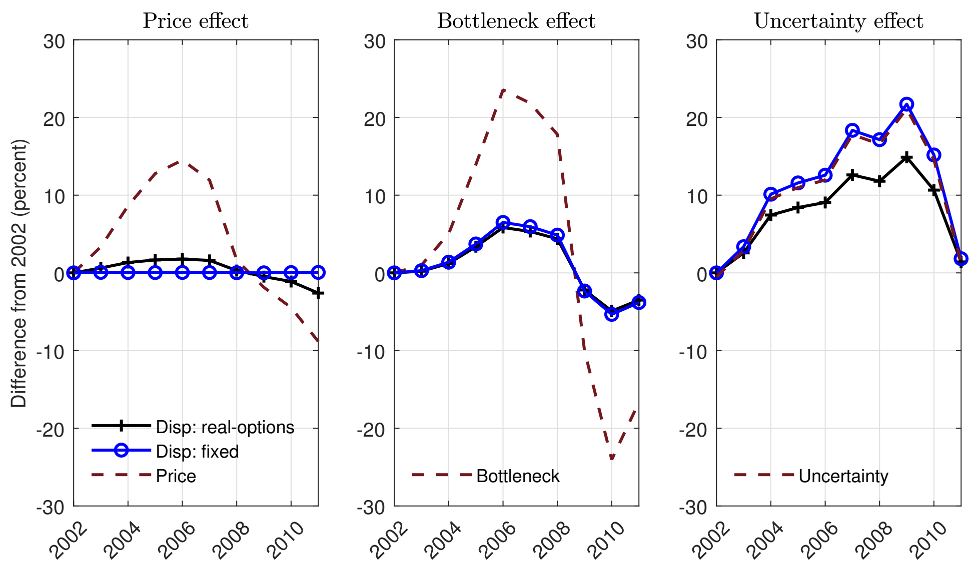
Note: Census measure indicates the census reported measure of the annual new house price index. Our measure indicates the new house price index computed based on the public data and our specified hedonic regression model. Both measures are normalized to 100 at their respective 2005 values.

Figure A5: Distribution of economic TTB (model and data)



Note: The data refers to the economic TTB data for built-for-sale houses as in the draft. The x-axis is TTB in months. The left end refers to houses completed within 6 months, and the right end refers to houses that took 24 months or more to complete.

Figure A6: Price dispersion and TTB



Note: Price dispersion (in standard deviation) is plotted by varying each aggregate state in each column holding fixed the other two aggregate states at their 2002 level. Real-options TTB indicates our model and fixed TTB indicates a model in which TTB is delayed only by bottlenecks.

List of Tables

Table A1: Sample skewness of economic TTB

Year	Skewness	95% confidence interval
2002	0.412	[0.346, 0.478]
2005	0.247	[0.181, 0.313]
2009	1.018	[0.948, 1.088]

Note: Skewness is the sample equivalent measure of $\mathbb{E}(X - \mu)^3/\sigma^3$, where $\mathbb{E}$ is the expectations operator, X is the observation, μ is the population mean, and σ is the population standard deviation. The confidence interval is computed by bootstrapping each distribution with replacement 1,000 times.

Table A2: Regression on TTB (TTB in levels)

	Frequency		
New England	0.019	2.131	(0.1520)
Middle Atlantic	0.052	2.146	(0.1016)
East North Central	0.095	1.329	(0.0751)
West North Central	0.067	1.153	(0.0874)
South Atlantic	0.306	0.285	(0.0538)
East South Central	0.054	0.807	(0.0727)
West South Central	0.142	-	
Mountain	0.119	0.305	(0.0701)
Pacific	0.146	1.148	(0.0737)
Modular	0.010	-1.095	(0.1670)
Panelized	0.023	-1.093	(0.0757)
Site built	0.967	-	
Square feet ($\times 100$)		0.059	(0.0024)
Constant		5.343	(0.1524)
Other controls:			
Metropolitan area		yes	
Number of full bath		yes	
Number of story		yes	
Detached		yes	
Deck		yes	
Parking facility		yes	
Foundation		yes	
Material of wall		yes	
Observations		111,628	

Note: Robust standard errors in parentheses.

Table A3: Model parameter estimates and standard errors; case (i)

Parameter	Year	Estimate	Standard error	95% confidence interval
P_T^M	2002	2.085	(0.00748)	[2.070, 2.099]
$p_{c,T}$	2002	0.338	(0.00258)	[0.333, 0.343]
	2003	0.340	(0.00271)	[0.335, 0.346]
	2004	0.355	(0.00257)	[0.350, 0.360]
	2005	0.384	(0.00248)	[0.379, 0.389]
	2006	0.418	(0.00250)	[0.413, 0.423]
	2007	0.411	(0.00284)	[0.405, 0.416]
	2008	0.397	(0.00343)	[0.390, 0.404]
	2009	0.309	(0.00529)	[0.299, 0.320]
	2010	0.256	(0.00559)	[0.245, 0.267]
	2011	0.280	(0.00509)	[0.270, 0.290]
σ_T	2002	0.395	(0.00285)	[0.392, 0.401]
(annual)	2003	0.411	(0.00327)	[0.405, 0.417]
	2004	0.435	(0.00274)	[0.433, 0.441]
	2005	0.442	(0.00317)	[0.438, 0.448]
	2006	0.445	(0.00402)	[0.440, 0.453]
	2007	0.472	(0.00364)	[0.464, 0.479]
	2008	0.463	(0.00359)	[0.458, 0.471]
	2009	0.479	(0.00383)	[0.478, 0.487]
	2010	0.455	(0.00417)	[0.450, 0.463]
	2011	0.403	(0.00373)	[0.398, 0.410]

Note: We set $\sigma_{p_c} = 0.5\sigma_{P^M}$ and $\sigma_\sigma = \sigma_{P^M}$.

Table A4: Model parameter estimates and standard errors; case (ii)

Parameter	Year	Estimate	Standard error	95% confidence interval
P_T^M	2002	2.122	(0.00755)	[2.107, 2.137]
$p_{c,T}$	2002	0.338	(0.00271)	[0.333, 0.344]
	2003	0.340	(0.00268)	[0.335, 0.346]
	2004	0.354	(0.00253)	[0.349, 0.359]
	2005	0.384	(0.00251)	[0.379, 0.389]
	2006	0.418	(0.00258)	[0.413, 0.423]
	2007	0.411	(0.00279)	[0.405, 0.416]
	2008	0.397	(0.00346)	[0.390, 0.404]
	2009	0.307	(0.00559)	[0.296, 0.318]
	2010	0.256	(0.00574)	[0.245, 0.267]
	2011	0.284	(0.00520)	[0.274, 0.294]
σ_T	2002	0.394	(0.00316)	[0.388, 0.400]
(annual)	2003	0.411	(0.00330)	[0.404, 0.417]
	2004	0.437	(0.00264)	[0.432, 0.442]
	2005	0.443	(0.00352)	[0.436, 0.450]
	2006	0.447	(0.00416)	[0.438, 0.455]
	2007	0.473	(0.00347)	[0.466, 0.480]
	2008	0.465	(0.00330)	[0.458, 0.471]
	2009	0.482	(0.00400)	[0.474, 0.490]
	2010	0.458	(0.00386)	[0.451, 0.466]
	2011	0.404	(0.00372)	[0.397, 0.412]

Note: We set $\sigma_{p_c} = 2\sigma_{P^M}$ and $\sigma_\sigma = \sigma_{P^M}$.

Table A5: Model parameter estimates and standard errors; case (iii)

Parameter	Year	Estimate	Standard error	95% confidence interval
P_T^M	2002	2.085	(0.00743)	[2.070, 2.099]
$p_{c,T}$	2002	0.338	(0.00254)	[0.333, 0.343]
	2003	0.340	(0.00271)	[0.335, 0.346]
	2004	0.355	(0.00257)	[0.350, 0.360]
	2005	0.384	(0.00248)	[0.379, 0.389]
	2006	0.418	(0.00255)	[0.413, 0.423]
	2007	0.411	(0.00282)	[0.405, 0.416]
	2008	0.397	(0.00343)	[0.390, 0.404]
	2009	0.309	(0.00529)	[0.299, 0.320]
	2010	0.256	(0.00559)	[0.245, 0.267]
	2011	0.280	(0.00505)	[0.271, 0.290]
σ_T	2002	0.395	(0.00287)	[0.390, 0.401]
(annual)	2003	0.411	(0.00327)	[0.404, 0.417]
	2004	0.435	(0.00273)	[0.430, 0.441]
	2005	0.442	(0.00317)	[0.436, 0.448]
	2006	0.445	(0.00396)	[0.438, 0.453]
	2007	0.472	(0.00363)	[0.465, 0.479]
	2008	0.464	(0.00358)	[0.457, 0.471]
	2009	0.479	(0.00382)	[0.472, 0.487]
	2010	0.455	(0.00416)	[0.447, 0.463]
	2011	0.403	(0.00372)	[0.396, 0.410]

Note: We set $\sigma_{p_c} = \sigma_{P^M}$ and $\sigma_\sigma = 0.5\sigma_{P^M}$.

Table A6: Model parameter estimates and standard errors; case (iv)

Parameter	Year	Estimate	Standard error	95% confidence interval
P_T^M	2002	2.085	(0.00748)	[2.070, 2.099]
$p_{c,T}$	2002	0.338	(0.00254)	[0.333, 0.343]
	2003	0.340	(0.00271)	[0.335, 0.346]
	2004	0.355	(0.00257)	[0.350, 0.360]
	2005	0.384	(0.00248)	[0.379, 0.389]
	2006	0.418	(0.00249)	[0.413, 0.423]
	2007	0.411	(0.00284)	[0.405, 0.416]
	2008	0.397	(0.00343)	[0.390, 0.404]
	2009	0.309	(0.00529)	[0.299, 0.320]
	2010	0.256	(0.00559)	[0.245, 0.267]
	2011	0.280	(0.00505)	[0.271, 0.290]
σ_T	2002	0.395	(0.00288)	[0.390, 0.401]
(annual)	2003	0.411	(0.00327)	[0.404, 0.417]
	2004	0.435	(0.00274)	[0.430, 0.441]
	2005	0.442	(0.00317)	[0.436, 0.448]
	2006	0.445	(0.00400)	[0.437, 0.453]
	2007	0.472	(0.00364)	[0.465, 0.479]
	2008	0.464	(0.00359)	[0.457, 0.471]
	2009	0.479	(0.00383)	[0.472, 0.487]
	2010	0.455	(0.00417)	[0.447, 0.463]
	2011	0.403	(0.00372)	[0.396, 0.410]

Note: We set $\sigma_{p_c} = \sigma_{P^M}$ and $\sigma_\sigma = 2\sigma_{P^M}$.