Incorporating multiple uncertainties into projections of Chinese private car sales and stock

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

Citation
Hsieh, I-Yun Lisa, Kishimoto, Paul Natsuo and Green, William H. 2018. “Incorporating multiple uncertainties into projections of Chinese private car sales and stock.” Transportation Research Record, 2672 (47).

As Published
10.1177/0361198118791361

Publisher
SAGE Publications

Version
Author’s final manuscript

Citable link
https://hdl.handle.net/1721.1/125040

Terms of Use
Creative Commons Attribution-Noncommercial-Share Alike

Detailed Terms
http://creativecommons.org/licenses/by-nc-sa/4.0/
INCORPORATING MULTIPLE UNCERTAINTIES INTO PROJECTIONS OF CHINESE PRIVATE CAR SALES AND STOCK

I-Yun Lisa Hsieh
Department of Chemical Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue E18-509
Cambridge, MA 02139-4307
iyhsieh@mit.edu

Paul Natsuo Kishimoto
MIT Joint Program on the Science & Policy of Global Change
Massachusetts Institute of Technology
77 Massachusetts Avenue E19-411-14
Cambridge, MA 02139-4307
pnk@mit.edu

William H. Green*
(Corresponding Author)
Department of Chemical Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue E17-504H
Cambridge, MA 02139-4307
whgreen@mit.edu
ABSTRACT
China is in a fast growing stage of mobility development, and its increasing demand for private cars comes with growing energy consumption and pollutant emissions. Uncertainty in Chinese parameterization of car ownership models makes forecasting these trends a challenge. We develop an application of the Monte Carlo method, conditioned on historical data, to sample parameters for a model projecting aspects of private car diffusion, such as the mix of new and replacement sales. Our model includes changes in per-capita disposable income—both the mean and level of inequality—and a measure of car affordability.

By incorporating multiple uncertainties, we show a distribution of possible future outcomes: a low stock of 280 million (1st decile); median of 430 million; and high of 620 million vehicles (9th decile) in 2050. This illustrates the limitations of attempts to model vehicle markets at the national level, by showing how uncertainties in fundamental descriptors of growth lead to a broad range of possible outcomes. While uncertainty in projected per-capita ownership grows continually, the share of first-time purchases in sales is most uncertain in the near term and then narrows as the market saturates. Replacement purchases increasingly capture the sales market from 2025. Our results suggest that stakeholders have a narrow window of opportunity to regulate the fuel economy, pollution and other attributes of vehicles sold to first-time buyers. These may, in turn, shape consumers’ experience and expectations of car ownership, affecting their additional and replacement purchases.
INTRODUCTION
Being one of the fast-growing major economies, China has seen the largest increases in passenger vehicle sales between 2005 and 2015, overtaking the U.S. and becoming the largest auto market since 2008 (1). However, if compared to the history of U.S. motorization, per-capita car ownership in China is just passing the level in the U.S. as of 1922 (2).

The strong growth in development of the automotive industry is closely associated with the economic boom in China. On the other hand, this growth has resulted in an increase in demand for fuel, and caused more pollutant and carbon dioxide emissions associated with petroleum combustion, as well as worsened traffic congestion problems in urban regions—all of which are urgent challenges for China’s government.

Establishing the range of possible future outcomes allows stakeholders to decide how robust their strategies are to uncertainty, and to determine the risk of following status quo approaches. In addition to design of policies to limit the environmental impacts of transport, governments must choose to size infrastructure (roadways, parking facilities, etc.) to accommodate expected traffic volumes. Automakers must construct durable assets such as assembly plants and make supply chain investments to meet demand competitively in a future market of a certain size. The criteria for robustness differ according to an organization’s tolerance for risk, and type of decision or policy under consideration; but in all cases information about distributions of future outcomes is essential.

Literature review
The evolution of the auto market, especially private car market, is of interest to numerous agencies. Hamilton et al. (3) distinguishes four stages in the growth of the automobile industry: first, experiment with invention; second, luxury uses and markets; then planned price reduction and expansion of markets; and finally, price stabilization and competition in design and salesmanship. The size of China’s car market is growing robustly, and the critical feature of the future growth potential of car markets has been approached and forecasted by many researchers. Different car ownership models have been applied for various purpose and different interests, as reviewed by de Jong et al. (4). It is suggested that the preferred model type for car stock forecasting depends on the research objectives and data availability. For those countries only having limited data, like China, some of the sophisticated methodologies may not be applied. A widely used method for application to developing countries is to express car density as a logistic function of either time (5), per-capita GDP (6) or per-capita income (7), because they have the lowest data requirements, and income is widely considered to be the primary motivation behind car ownership growth.

By fitting Gompertz functions to data on per-capita vehicle ownership and per-capita GDP, Wu, Zhao, and Ou (6) showed that car stock level in China has followed an S-shaped curve like most other motorizing countries, and projected that the inflection point would appear around the year 2030. Instead of relying on single economic variable, Huo and Wang (8) built up a China-specific fleet model in more detail, simulating private car ownership on a income-level basis and taking into account explicitly the influence of car price on vehicle ownership, in order to investigate the energy demand and environmental impacts of the rising number of vehicles in China.

They separated car sales into two categories: new-growth purchases (i.e., first-time car purchases and additional purchases) motivated by increased income, and replacement purchases, which households use to replace retired or scrapped vehicles. The approach in Huo and Wang (8) serves as a key example for the current paper. However, we include uncertainty in the vehicle
adoption trajectory, thereby addressing a key parameter and source of uncertainty in outcomes omitted in prior research.

We describe an approach that allows exploration of the probability distribution of the saturation level of car ownership at high purchasing power, and allows comparison of effects of uncertain parameters, including disposable income, Gini index, and car price. We show results by conducting Monte Carlo simulation, and discuss the sensitivity to major parameters. We conclude by stating some implications for policy makers and city planners.

**METHOD & DATA**

**Car stock & sales**

With the survey of Chinese urban families with different income levels owning cars, conducted by China’s Statistical Bureau, it is confirmed that it is more likely for people with higher income to own cars. Compared to income alone, per-capita income coupled with car price, called car purchasing power, was shown to have much better correlation with car ownership in China (8). Especially for emerging economies like China’s, car purchasing power is a more suitable independent variable to be used to model car ownership than income level alone, because the automotive industry is expanding as car prices decrease in the early stage of motorization.

Let $A$ represent an affordability index, or purchasing power, which is the ratio of disposable income to car price index. We arbitrarily choose the car price index in 2003 ($p_{2003}$) to be 1 and use the United Nations medium-variant population projection as our baseline (9). The total private car stock ($\hat{V}_i$) can be calculated by integrating across the entire population’s range of incomes, $x$, the product of the income distribution density, $I_i$, and the propensity, $g$, that an individual with that affordability index owns a car, as shown in (1).

$$\hat{V}_i = P_i \int_{A=0}^{\infty} I_i(A) g(A) dA$$

(1)

where

$\hat{V}_i =$ Total private car stock in year $i$

$P_i =$ Population in year $i$ [persons]

$x =$ Income of the incremental individual [RMB$_{2007}$/year]

$I_i =$ Income distribution function for year $i$

$p_i =$ Car price index in year $i$

$A = \frac{x}{p_i} =$ Car affordability index

$g =$ Deterministic relation between private car ownership (%) and affordability index

We decompose car sales, $S$, into two segments: new-growth purchases, $S^N$, associated with increases in ownership due to rising income, and replacements, $S^R$, for scrapped cars. New-growth purchases, calculated as (2), represent the spread of ownership to a larger portion of the population as incomes rise.

$$S^N_i = \hat{V}_i - \hat{V}_{i-1} = P_i \int_{A=0}^{\infty} I_i(A) g(A) dA - P_{i-1} \int_{A=0}^{\infty} I_{i-1}(A) g(A) dA$$

(2)
Even if per-capita ownership does not rise, households must buy replacements for old vehicles that are scrapped or retired. We call these replacement purchases, and calculate them as (3).

\[ S^R_i = \sum_{j=1}^{i} S_{i-j} (SR(j-1) - SR(j)) \]  

where

\[ S_{i-j} = S^N_{i-j} + S^R_{i-j} \]  

is the total number of new cars purchased in year \( i - j \) \( (4) \)

\[ SR(y) = \exp\left(-\left(\frac{y}{T}\right)^b\right) \]  

is the survival ratio [dimensionless] \( (5) \)

where \( y \) is the age of the vehicle [years], \( T \) is the average vehicle life span = 14.46 years, and \( b \) is the scrappage intensity = 4.7.

The vehicle survival rate, \( SR(y) \), describes the decreasing share of vehicles of a certain age remaining in use as the vehicle age grows. Hao et al. \((10)\) collected data from a field study and employed the Weibull distribution to simulate it, as shown in (5). Because China has no official scrappage statistics and no compulsory scrappage standards for private cars, we adopt their values for \( T \) and \( b \) (shown above) in our analysis and assume that the pattern will remain the same in the future.

The careful reader will note certain differences in equations (1) and (2) compared with those given in reference \((8)\). In (1), we use \( I_i(A)g(A) \) instead of \( I_i(x)g(A) \). This is necessary such that the integral properly represents the expectation of \( g \) over a population of individuals with car affordability index distributed as \( A \). In (2), the cited paper has something akin to \( S^N_i = P_i \int_{M-1}^{M} \int_{x}^{\infty} I'(x)g(x,p_i)dxdM \), where \( M \) is the population mean income. This again mixes functions of \( x \) and \( A \); and additionally uses a derivative of the income distribution, \( I'(x) \). In the absence of model source code, the meaning is ambiguous; a derivative of \( I(\cdot) \) with respect to \( x \) or \( A \) is negative at certain points, leading to negative projected sales in some future years, and we find that a derivative with respect to \( M \) (or \( \mu \)) yields results that do not match observations. We instead use the method of finite differences described above.

The car ownership function, \( g(x,p_i) \)

Parameter uncertainty in Gompertz functions

It is recognized that a primary driving force for the growth of private car ownership is income level \((11)\). The statistics relating to Chinese consumer price index \((12, 13)\) show that the car price in China decreased rapidly in the past decade, driven by excess capacity and competition in the rapidly growing automotive market. Based on the theory of four stages in the growth of the car industry \((3)\), car price stabilizes after the stage of the planned price reduction and expansion of markets, as shown in Figure 1(a). Compared to the early motorization stage in the U.S. \((14)\), the theory suggests that both countries would have similar price trends.

A special case of the logistic function, the Gompertz curve allows for different curvatures at extreme values of the economic factors \((15)\). We use a formulation (6) that represents the probability that an individual owns a car as a function of \( A \).
\[ g(x, p_i) = g(A) = \gamma \exp(\alpha \exp(\beta A)) \] (6)

where

- \( g = \text{Probability to own a private car} \)
- \( A = \text{Car affordability index} \)
- \( \gamma = \text{Probability of owning a private car at very high income} \)
- \( 0 < \alpha, \beta = \text{Shape parameters} \)

In order to illustrate the limitation of point estimates, we use data from the 2008, 2009, and 2010 China urban household surveys (16). (Note that prior to 2010, only Shanghai limited car ownership using a license plate quota system, but in recent years a growing number of cities have adopted such policies. To avoid obscuring the underlying relationship, we omit data from 2011 onwards.) Figure 2 shows values of \( \gamma \) ranging from 20\% to 80\% or higher produce fits that are all apparently close matches to the data. A common response in the literature is to arbitrarily choose \( \gamma \) from another country with saturated ownership, or some plausible low or high bounding values for a few scenarios.

Economic development and car ownership in China are both behind those of motorized countries, and thus as a whole the country is still below the inflection point of \( g \). Thus, the survey data tell us little about the eventual probability of possessing a car when people have higher income levels, even though the results show that people with higher car purchasing power are more likely to own cars, as expected. Due to the uncertainties arising from unclear future economic context and unpredictable consumer preference, the value of single-point estimates is limited.

**Monte Carlo simulation of the distribution of \( \gamma, \alpha, \text{and} \beta \)**

Following (17), our updated algorithm to estimate the parameters of Gompertz function by incorporating uncertainty is as follows. The unknown parameters are \( \gamma, \alpha, \text{and} \beta \) while the state variable is only \( r = \gamma \), because the latter two parameters can be further determined by linear regression at a given \( \gamma \). As the data are available for per-capita disposable income at national level, car price index, population, and private car stock numbers during 2007 to 2015, we then define the root-sum-of-square residual between back-cast private car stock \( (\hat{V}_i) \) by (1) and actual historical data \( (V_i) \) as (7).

\[
\text{RSS}(v) = \sqrt{\sum_{i=2007}^{2015} (\hat{V}_i(r) - V_i)^2} 
\] (7)

Monte Carlo simulation with 400 iterations is applied to estimate the possible outcomes of the saturation level of possibility (\( \gamma \)) forecasts. \( \gamma \) is constrained to the range between the 23\% (the observed per-capita car ownership for the highest income household in 2012 (16)) and 100\%. A discrete probability density function for \( \gamma \in (23, 100) \) is constructed by randomly sampling 50 points across the range and calculating their associated \( \text{RSS}(r) \), subtracting from \( \max(\text{RSS}) \) and
then normalizing so that the density is lowest where the residual is greatest. From the resulting distribution of $r$ for each iteration, a single sample is drawn and accepted as the $\gamma$ in that run, followed by estimating the other two parameters, $\alpha$ and $\beta$, from linear-regression model as (8), which is converted from (6) by log-linearization. As shown in (8), $\ln (-\alpha)$ and $\beta$ are linearly related and are regressed as a least-squares fit for time series data in each iteration.

$$\ln \left( \frac{\ln \gamma}{g(A)} \right) = \ln (-\alpha) + \beta A$$  \hspace{1cm} (8)

The income distribution, $I$

Various probability distribution functions have been studied in research on income disparity. By comparing observations to mathematical models, researchers can determine the distribution function with the best fit. Some scholars prefer the log-normal distribution; for example, Steyn (18) examined the inhabitants in the Orange Free State rural areas of the Union of South Africa, and showed that the logarithmic normal model is quite a good model. On the other hand, McDonald and Ransom (19) stated that the log-normal distribution performed the worst while considering the log-normal, gamma, beta, and Singh-Maddala functions as descriptive models for the distribution of family income for 1960 and 1969–1975 in the U.S.

Data availability has been a major obstacle for research in many fields in China, including research on income inequality. Chen et al. (20) investigated 20 sets of grouped data on family income between 2005 and 2012 in China, and demonstrated that the fitting of the log-logistic distribution is better, while the log-normal distribution function underestimates the income of the high-income group due to its thinner right tail. Following (20), we approximate a continuous income distribution for China with (9).

$$f(x; a, b) = \frac{b x^{b-1}}{a^b \left(1 + \left(\frac{x}{a}\right)^b\right)^2}$$  \hspace{1cm} (9)

where

$x$ = per-capita disposable income [RMB\textsubscript{2007}/year]

$a, b$ = scale and shape parameters of the distribution

Note that if $x$ is log-logistically distributed, then $A$ has also a log-logistic distribution but with different scale parameter (10).

$$X \sim f(x; a, b) \Rightarrow A = \frac{X}{p} \sim f\left(\frac{x}{p}; \frac{a}{p}, b\right)$$  \hspace{1cm} (10)

Given historical data and projections of the mean disposable income ($M$) and Gini index ($G$), we then solve the equations, shown as (11) and (12), analytically by computer mathematical software to simulate the income distribution in a given year.

$$G = \frac{1}{b}$$  \hspace{1cm} (11)
\[ M = \frac{a \pi}{b \sin(\frac{\pi}{b})} \] (12)

For future income, we use projections from China’s 13th Five-Year Plan and the OECD, and assume that \( M \) will increase at the same rate as GDP after 2020 (21). Figure 1(b) shows the future projection of \( M \). Hu, Luo, and Yang (22) adopted the grey forecasting model, commonly applied in academic fields such as electrical engineering, mechanical engineering, and agriculture, to predict the Gini index in China until 2022. With the given difference between the official Gini index (0.465) in 2016 and Hu, Luo, and Yang’s predicted value (0.479), we first correct those from 2017–2022, and then adopt the same rate of change in subsequent years to give the future Gini index projections, shown in Figure 1(c). Figure 1(d) depicts the resulting income distributions in China for some years, \( i \), between 2007 to 2050 based on our derived projected function parameters; in computing projections, we use a distinct \( I_i(A) \) \( \forall i \).

RESULTS & ANALYSIS

Parameter space for \( \alpha, \beta, \) and \( \gamma \)

Figure 3 illustrates the parameter space for the Gompertz function at 400 samples of \( \gamma \) with the associated probability distribution from Monte Carlo simulation and regression. Note that for a randomly-selected \( \gamma \), our regression procedure for \( \alpha, \beta \) constrains these parameters to a narrow range that matches historical data, unlike (17) which treated all three as a state vector and applied Gibbs sampling, drawing each of three parameters in sequence from a joint distribution.

Private car stock

Figure 4 presents the outcomes of the private car stock and ownership in China projected by our stock model for all 400 samples, with 1st decile, mean and 9th decile shown by black lines, and the probability distribution of the outcomes in 2030 and 2050. We also show the projection results, for comparison, from Huo and Wang (8) for stock, and from TASRI (Tsinghua Automotive Strategy Research Institute, (23)) for ownership. Note that our model is forecasting private cars, rather than all vehicles as in TASRI. The model shows a range of possible future private car stock in China: 430 million (low of 280 million at 1st decile; high of 620 million at 9th decile) in 2050.

Table 1 gives a summary of the parameter values for 1st decile, median, 9th decile and the expectations with standard deviations, with the associated projected private car stock numbers and per-capita car ownership. As shown in Table 1, uncertainty is larger in estimating the parameter \( \gamma \) compared to the other two, which implies that predicting the probability of private car ownership in China is challenging, as discussed above (“Parameter uncertainty in Gompertz functions”).

Private car sales

Figure 5 draws out, for every sample, the results of private car sales in China, with 1st decile, mean and 9th decile shown in black lines, and the associated shares of replacement purchase in dark red lines, and the probability distribution of the both outcomes toward 2050. It is noted that higher car sales correspond to a lower share of replacement purchases; that is, the order of the black quantile lines is reversed for the two outcomes. The predicted distribution of the private car sales in China will be centered at 30 million from 2030 to 2050 (low of 19 million at 1st decile; high of 44 million at 9th decile). While scenario analysis, which brackets uncertain outcomes of interest, is one response to improve the value of single-point estimates, our upgraded methodology provides a more detailed characterization of the uncertainty in transport projections.
Table 2 summarizes the expected value of private car sales numbers and replacement purchases share with the associated standard deviations by our model. As stated (“Method & data”), private car sales are divided into two groups due to their purchasing motivations: new-growth purchase and replacement purchase. The split between these two groups characterizes the maturity of the auto market (8). As shown in Figure 5, current car sales in China are mainly driven by first-time car purchases—i.e., the share of replacement purchases is still low. Our results show that replacement purchases will dominate the sales market from 2025, which is a feature of maturity. However, the replacement purchase share is still lower than 100% in 2050, implying that the car market in China will not fully mature by 2050. Unlike the uncertainty in projections of private car stock/ownership and car sales, which grow continually into the future, the distribution of the replacement purchases share narrows toward the saturated (~2050) market state. We discuss the implications below (see “Discussion”).

Another way to characterize the maturity of a car market is using the ratio ($r$) of the private car sales to the change in private car stock as an indicator. Note that “one minus the inverse of the ratio” is the proportion of the new cars sold to replace the old models which have been scrapped. Therefore, the ratio should be close to 1 in an emerging car market, as the new-growth purchases dominate sales. As a check on our model projections, we compare this indicator with data from the early stage of motorization in the U.S. (24), as shown in Figure 6. According to our model, two lines (black line represents the numerator while blue dash line represents the denominator of the ratio) are close to one another in initial years but will separate gradually as the auto market in China becomes more mature, which is the same as the observed pattern in the U.S. in the 20th century.

However, care should be taken when drawing parallels between China and the United States: while some trends may be similar, the adoption of personal vehicles in China nowadays happens in a much different context compared to the one in the United States a hundred years ago. The preferences, technology options, costs, earnings, urban situation, and policy environment faced by 21st-century Chinese households differ in important ways from those of 20th-century Westerners; our projections highlight the uncertainty these distinctions create in the trajectory of motorization.

Sensitivity Analysis

The projections presented above rely on historical data, household sample surveys, and exogenous projections of the independent variables. We conduct sensitivity analysis to help assess how much the model output values are conditioned by the latter. The major parameters here are car price ($p_i$), mean per-capita disposable income ($M$), Gini index ($G_i$), and the ultimate saturation level ($\gamma$). As uncertainty in $\gamma$ has been investigated by Monte Carlo simulation, the influences of the other three parameters ($p_i$, $G_i$, and $x$) are studied, and we adopt the expectation value of $\gamma$ as a base case. For brevity, we focus only on sensitivity of this central value to assumptions; however, the method we describe makes it straightforward to study sensitivity of the spread and shape of the distributions of projected outcomes, as well.

There is disagreement in the literature over the estimation of China’s household income (25) and the precise past and present values of the Gini index (20); future forecasts contain additional uncertainty. In Figure 7, we show projections from other agencies (HSBC (26), Standard Chartered (27), and Morgan Stanley (28)) for income level and literature (Wang et al. (29)) for Gini index, and our sensitivity trajectories that result in values that are $(100 + n)\%$ of the projected
values in 2050, where \( n \in (-40, -30, ..., 30, 40) \).

Figure 8 presents how the change in car ownership as the major parameters change. As expected, car ownership increases when car purchasing power increases, due to either an increase in per-capita disposable income level or a decrease in car price index. Note that car ownership is most sensitive to car price in near-term future (2020); whereas it is nearly equally sensitive to income level along the future period.

The Gini index has a negative relationship with car ownership in near and mid-term future (2020 and 2030): as Gini index (level of income inequality) increases, the mass of income distribution shifts to the left, and thus the majority of the society becomes poorer while almost all income is held by a small fraction of very wealthy households. In this case, per-capita car ownership is lower, because a larger fraction of the population cannot afford car ownership at all. By contrast, when China’s economy and car market become developed and mature, income inequality has little effects on private car ownership. While the sensitivity of car ownership to car price shown in Figure 8 is consistent with (8), the sensitivity results with respect to per-capita disposable income and Gini index are different. Our study shows the similar sensitivity to income level toward 2050, whereas the magnitude of the changes decreases as time passes in (8); in the current results, the sensitivity of car ownership to Gini index is not always monotonically decreasing. This result is due to the different probability distribution functions chosen to simulate the income distribution in China and the different calculation mechanisms for car stock, as described above (“Car stock & sales”). For example, the log-normal distribution used in (8) somewhat underestimates the income of the highest-income households, and thus may further underestimate the sensitivity to income level in the long-term projections as more people become wealthy.

**DISCUSSION**

Our analysis simulates the possible outcomes of China’s growth in private car ownership and sales by characterizing the distribution of the adoption trajectory. Figure 4 shows that our model captures the pattern of growth, although the expected value slightly overestimates historical car ownership and stock. As mentioned (“Parameter uncertainty in Gompertz functions”), we are modeling the development of car market without many government interventions. However, besides household incomes, car price, and income inequality, policy also affects transport-related consumption. To reduce private vehicle travel and create more livable and sustainable cities, an increasing number of the largest cities in China have sharply restricted new vehicle registration by implementing quota policies (30). While Shanghai was a proactive adopter of license auctioning (starting in 1994), other cities did not impose their own quotas until 2011 or later, when the motorization was already in rapid progress. Such caps on vehicle purchases may increase the deviation of our projections from observations in the future; however, the likely distribution of our projections can aid in the design of robust transportation policy that relies on forecasts of demand.

Previous research has shown that the preferences of existing vehicle owners are, to some extent, anchored—they may prefer to get a vehicle that is similar to the one they are replacing, or already own (31); or they may even prefer to upgrade their cars in size and horsepower (32). Manufacturers both encourage and follow these preferences. We show (Figure 5) that the percentage of replacement sales can be expected to be large in the new car sales market around 2025. Therefore, in the near-term where new sales are a large share of the market, the government has significant opportunities to guide the type, size, fuel economy and pollution characteristics of vehicles sold. These will in turn condition households’ experience and expectations of vehicle ownership, and
thereby affect their future replacement and additional purchases.

According to the 13th Five-Year Plan (2016–2020), China proposed to become “a moderately prosperous society in all respects” (33). The government is aiming to double per capita income by 2020 from 2010 levels, and to improve the fairness of income distribution through expanded public services. If these objectives are achieved, Figure 8 reveals that the private car ownership rate in China will be higher in 2020 as the economy is more equal (i.e. the Gini index is lower). Proper transport demand management (TDM) strategies are required to reduce private vehicle travel, curb car congestion, and to make sustainable transport a reality—and our results show that they are even more urgent if the government is successful in pursuing its goal of reducing inequality.

Extensions
The regional distribution of China’s current vehicle market reflects the regional heterogeneity in GDP contribution across the country and the different regional urbanization level. China’s six largest civil vehicle owning provinces—Shandong, Guangdong, Jiangsu, Zhejiang, Hebei, and Henan—represented around 47% of the total vehicle fleet at the end of 2015. In particular, in terms of the contribution for incremental growth, these six provinces accounted for nearly 50% every year in the past ten years (16). This pattern shows that the coastal regions have dominated the national market.

In order to incorporate the significant regional differences in the level of economic development and vehicle ownership, and produce more precise projections with possibly narrower distributions, the national-level methodology shown in this study could be applied at the regional or provincial level. To do this, the car ownership function, $g$, would be sampled separately and survival ratio, $SR(y)$, is estimated differently for provincial-level cities and other sub-national regions, and the exogenous variables could be drawn from province-level projections.

Finally, with the car stock and car sales models at either the national or regional levels, our work allows examination of uncertainty in the future trends of life cycle energy demand and emissions from private car use in China.

Conclusion
Undoubtedly, China’s car stock will continue to grow, driven by the growing economy. However, due to the short history and the country being in a fast-growing stage of transport system expansion, national observations contain little information about the eventual saturation level. Here, we improved on existing methods by incorporating adoption uncertainties and considering multiple key influencing factors, demonstrating a novel treatment of uncertainty in transport forecasting in China. A range of future forecasts in transport demands that are conditioned on observed reality was developed, producing a characterization that can help policy makers or transport infrastructure project managers create flexible designs that are robust to eventualities. On the other hand, the use of single-point estimates could lead to a ‘flaw of averages’, by obscuring irreducible future uncertainty (34).

We found that the distribution of car stock/ownership is wider in 2050 than 2030; on the other hand, the distribution of the share of new-growth purchases is higher in the near term than in a saturated market, with replacement sales expected to dominate the car sales market after the next decade. Finally, our sensitivity analysis results suggest that the car ownership will increase further in 2020 as the society becomes less unequal; and more so if this trend goes faster than the
literature suggests. We also discussed how our method and estimates have important implications for regulators and planners, and could be a basis for regionally-disaggregated projections with similar advantages.

ACKNOWLEDGEMENT
This work was supported through the MIT Energy Initiative’s Mobility of the Future study.

AUTHOR CONTRIBUTION
Study conception and design: I.Y.L.H., P.N.K., and W.H.G.; data collection: I.Y.L.H.; analysis and interpretation of results: I.Y.L.H., P.N.K., and W.H.G; draft manuscript preparation: I.Y.L.H. and P.N.K. All authors reviewed the results and approved the final version of the manuscript.
REFERENCES

[1] OICA. 2005–2015 Sales Statistics. International Organization of Motor Vehicle Manufacturers. URL: http://www.oica.net/category/sales-statistics/ (visited on 2017-01-15).

[2] S. C. Davis, S. W. Diegel, and R. G. Boundy. Transportation Energy Data Book. Edition 33. 2014. URL: http://wedocs.unep.org//handle/20.500.11822/15991.

[3] W. Hamilton et al. “Price and Price Policy”. London, Me Graw Hill (1938), p. 27.

[4] G. de Jong, J. Fox, A. Daly, M. Pieters, and R. Smit. “Comparison of car ownership models”. Transport Reviews 24.4 (2004), pp. 379–408. DOI: 10.1080/0144164032000138733.

[5] T. Zachariadis, Z. Samaras, and K.-H. Zierock. “Dynamic modeling of vehicle populations: an engineering approach for emissions calculations”. Technological Forecasting and Social Change 50.2 (1995), pp. 135–149. DOI: 10.1016/0040-1625(95)00057-H.

[6] T. Wu, H. Zhao, and X. Ou. “Vehicle ownership analysis based on GDP per capita in China: 1963–2050”. Sustainability 6.8 (2014), pp. 4877–4899. DOI: 10.3390/su6084877.

[7] J. M. Dargay, D. Gately, and M. Sommer. “Vehicle ownership and income growth, worldwide: 1960–2030”. The Energy Journal (2007), pp. 143–170. URL: http://www.jstor.org/stable/4132125.

[8] H. Huo and M. Wang. “Modeling future vehicle sales and stock in China”. Energy Policy 43 (2012), pp. 17–29. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2011.09.063.

[9] United Nations. Department of Economic. World population prospects: The 2015 revision. 2015. URL: https://www.populationpyramid.net/china/2017/ (visited on 2017-04-30).

[10] H. Hao, H. Wang, M. Ouyang, and F. Cheng. “Vehicle survival patterns in China”. Science China Technological Sciences 54.3 (2011), pp. 625–629. ISSN: 1674-7321. DOI: 10.1007/s11431-010-4256-1.

[11] J. M. Dargay. “The effect of income on car ownership: evidence of asymmetry”. Transportation Research Part A: Policy and Practice 35.9 (2001), pp. 807–821. DOI: 10.1016/S0965-8564(00)00018-5.

[12] National Development and Reform Commission. Consumer Price Index. URL: http://en.ndrc.gov.cn/ (visited on 2017-05-10).

[13] China Industry Information Network website. URL: http://www.chyxx.com (visited on 2017-04-10).

[14] S. C. Davis, S. W. Diegel, and R. G. Boundy. Transportation Energy Data Book. Edition 35. 2016. URL: http://cta.ornl.gov/data/index.shtml.

[15] J. Dargay and D. Gately. “Income’s effect on car and vehicle ownership, worldwide: 1960–2015”. Transportation Research Part A: Policy and Practice 33.2 (1999), pp. 101–138. DOI: 10.1016/S0965-8564(98)00026-3.

[16] National Bureau of Statistics of China. China Statistics Yearbook 2016. URL: http://data.stats.gov.cn/english/easyquery.htm?cn=C01 (visited on 2017-04-22).
[17] P. N. Kishimoto. “Applying Advanced Uncertainty Methods to Improve Forecasts of Vehicle Ownership and Passenger Travel in China”. 31st USAEE/IAEE North American Conference. (Austin, TX, US, Nov. 4, 2012). 2012.

[18] H. S. Steyn. “A model for the distribution of incomes”. *South African Journal of Economics* 27.2 (1959), pp. 149–156. DOI: 10.1111/j.1813-6982.1959.tb01820.x.

[19] J. B. McDonald and M. R. Ransom. “Functional forms, estimation techniques and the distribution of income”. *Econometrica: Journal of the Econometric Society* 47.6 (1979), pp. 1513–1525. DOI: 10.2307/1914015.

[20] J. Chen, F. Fang, W. Hou, F. Li, M. Pu, and M. Song. “Chinese Gini coefficient from 2005 to 2012, based on 20 grouped income data sets of urban and rural residents”. *Journal of Applied Mathematics* 2015 (2015). DOI: 10.1155/2015/939020.

[21] OECD. “Long-term baseline projections, No. 95 (Edition 2014)”. 2017 (1). DOI: 10.1787/data-00690-en. (Visited on 2017-05-02).

[22] J. Hu, Y. Luo, and X. Yang. “Forecast and Policy Suggestions of Gini Coefficient in China in the Future Decade - Analysis Based on Grey Forecasting Model”. *West Forum on Economy and Management* 26.1 (2015). (In Chinese).

[23] Beijing Times. *Experts: 2030 China’s vehicles per 1000 people will maintain the amount of 300*. Sept. 2016. URL: http://epaper.jinghua.cn/html/2016-09/22/content_336471.htm (visited on 2017-07-26). (In Chinese).

[24] United States Department of Transportation. *National Transportation Statistics*. URL: https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/index.html (visited on 2016-12-10).

[25] X. Wang and W. T. Woo. “The size and distribution of hidden household income in China”. *Asian Economic Papers* 10.1 (2011), pp. 1–26. DOI: 10.1162/ASEP_a_00064.

[26] K. Ward. *The world in 2050: Quantifying the shift in the global economy*. HSBC Global Research, Jan. 4, 2011.

[27] C. Daily. *Per capita income to rise by up to 50% by 2030*. Oct. 2011. URL: http://www.chinadaily.com.cn/bizchina/2012-10/11/content_15810209.htm (visited on 2017-07-05).

[28] E. Curran. *China Will Avoid a Bank Crisis, Reach High Income Status: Morgan Stanley*. Feb. 2017. URL: https://www.bloomberg.com/news/articles/2017-02-14/morgan-stanley-says-china-to-avoid-bank-shock-reach-high-income (visited on 2017-07-05).

[29] X. Wang, K. Z. Chen, S. Robinson, and Z. Huang. “Will China’s demographic transition exacerbate its income inequality?–CGE modeling with top-down microsimulation”. *Journal of the Asia Pacific Economy* 22.2 (2017), pp. 227–252. DOI: 10.1080/13547860.2016.1263043.

[30] L. Xue. *Four Lessons from Beijing and Shanghai Show How China’s Cities Can Curb Car Congestion*. WRI Ross Center for Sustainable Cities. Apr. 10, 2015. URL: http://www.wri.org/blog/2015/04/4-lessons-beijing-and-shanghai-show-how-china%E2%80%99s-cities-can-curb-car-congestion (visited on 2015-12-18).
[31] M. Sivak and B. Schoettle. *What do current owners of hybrids and non-hybrids think about hybrids?* Publication UMTRI-2014-25. University of Michigan, Ann Arbor, Transportation Research Institute, 2014.

[32] C. R. Knittel. “Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector”. *American Economic Review* 101.7 (Dec. 2011), pp. 3368–3399. DOI: 10.1257/aer.101.7.3368.

[33] J. Ross. *China’s Five Year Plan to achieve a ‘moderately prosperous society’*. Oct. 2015. URL: http://www.china.org.cn/opinion/2015-10/30/content_36935303.htm (visited on 2017-07-25).

[34] R. de Neufville and S. Scholtes. *Flexibility in engineering design*. MIT Press, Cambridge MA, 2011. ISBN: 9780262299145.
FIGURES & TABLES

| Parameter                        | 1st decile | Median | 9th decile | $E[\cdot] \pm \sigma$ |
|----------------------------------|------------|--------|------------|-----------------------|
| $\gamma$ (%)                     | 28.206     | 48.090 | 74.538     | 49.94 ± 17.29         |
| $\alpha$                         | -5.307     | -5.491 | -5.796     | -5.400 ± 0.184        |
| $\beta$                          | $3.18 \times 10^{-5}$ | $2.38 \times 10^{-5}$ | $2.00 \times 10^{-5}$ | $-(3.32 \pm 0.44) \times 10^{-5}$ |
| Stock $\hat{V}_{2030}$ (million) | 235.62     | 342.75 | 470.14     | 423.84 ± 86.78        |
| Stock $\hat{V}_{2050}$ (million) | 282.26     | 434.21 | 620.32     | 504.93 ± 125.37       |
| Vehicle per 100 capita$_{2030}$  | 16.634     | 24.213 | 33.213     | 29.942 ± 6.131        |
| Vehicle per 100 capita$_{2050}$  | 20.938     | 32.210 | 46.016     | 37.456 ± 9.300        |
**TABLE 2 Private car sales projections from car sales model**

| Year | $E^{} \pm \sigma$ | Car Sales (million) | Replacement purchases share (%) |
|------|--------------------|---------------------|--------------------------------|
| 2020 | 21.27±5.39         | 24.16±2.56          |
| 2025 | 24.17±6.24         | 52.49±3.17          |
| 2030 | 29.16±7.78         | 70.38±2.70          |
| 2040 | 30.40±8.68         | 83.97±2.18          |
| 2050 | 30.80±9.13         | 94.95±1.54          |
FIGURE 1 Major parameters and functions in car stock model: (a) Car price index in China (historical data 2003-2016 and future projection to 2050) and the historical data in U.S. (1913-1962); (b) Per-capita disposable income (2007 RMB), 2007–2050; (c) Gini index, 2007–2050; and (d) Income distribution with the log-logistic assumption, 2007–2050
FIGURE 2 Private car ownership and car affordability index in China, $g(x, p_i)$, with assumed $\gamma$ between 20% and 80%
FIGURE 3 Parameter space for $\alpha$, $\beta$, $\gamma$ with the probability density distribution of $\gamma$
FIGURE 4 Possible outcomes from Monte Carlo simulation: (a) 2010–2050 private car ownership with the probability distribution in 2030 and 2050; and (b) 2010–2050 private car stock with the probability distribution in 2030 and 2050.
FIGURE 5 Possible outcomes of the private car sales and replacement purchase proportion out of total car sales in China with their probability distributions.
FIGURE 6 Comparison of our model results for China 2010–2050 with the motorization-stage in U.S. 1915–2000.
FIGURE 7 Trajectory of the major parameters varied in the sensitivity analysis.
FIGURE 8 Sensitivity of private car ownership with respect to major parameters (solid line) and the corresponding ownership level in 2050 (dotted line).