Bid-auction framework for microsimulation of location choice with endogenous real estate prices

Ricardo Hurtubia
Michel Bierlaire
Francisco Martínez

Urbanics
Termas de Chillán, Chile
March 28th 2012

Outline

1) Motivation
2) The bid-auction approach to location choice modeling
3) Estimation of bid-rent functions
4) Bid-auction framework for microsimulation of location choice

Motivation – Land use models

- Spatial distribution of agents and activities in a city affects:
  - Travel demand
  - Energy consumption, pollution
  - Social welfare
- Cities are complex systems:
  - Interaction of different markets
  - Many heterogeneous agents
  - Externalities
- Land use models allow to understand and forecast (?) the evolution of cities
- Location choice models are a fundamental element of land use models
- Microsimulation / agent based models are flexible and detailed, making possible to evaluate complex scenarios

Motivation – Approaches to location choice modeling

- **Choice:** agents (households and firms) select location of maximum utility as price takers
  - Most usual implemented approach in microsimulation
  - Requires prices/rents to be given (usually modeled with a hedonic price model and/or exogenous adjustments)
- **Bid-auction:** real estate goods are traded in auctions where prices and locations are determined by the best bidders
  - Usually implemented in equilibrium models (bids are adjusted so everyone is located somewhere)
  - Prices are endogenous (expected maximum bid)

Motivation – Bid-auction advantages

- Real estate goods (housing, land) are quasi-unique and usually scarce ➔ competition between agents
- Explicit explanation of the price formation process (best bid in an auction)
- Bid prices can be sensitive to scenarios of demand or supply surplus
- Estimation: no price endogeneity (spatial autocorrelation)
- But:
  - Estimates of bid function must reproduce both prices and location distribution
  - Bid-auction is not straightforward to implement in microsimulation framework
  - Detailed data is usually not available

Bid-auction approach to location choice

- B_{hi} : willingness to pay of agent *h* for location *i*.

$$B_{hi} = f(x_h, z_i, \beta)$$

x_h : characteristics of agent h (household, firm, …)
z_i : attributes of location i (housing unit, parcel of land, …)

- Probability of agent *h* being the best bidder for a location *i* (Ellickson, 1981):

$$P_{h/i} = \frac{\exp(\mu B_{hi})}{\sum_{g \in H} \exp(\mu B_{gi})}$$

H: set of bidding agents

Bid-auction approach to location choice

- Price or rent for one location:
  - Deterministic: bid of the winner of the auction
  - Stochastic: expected maximum bid

- r_i : rent/price of i = expected value of the maximum bid:

$$r_i = \frac{1}{\mu} \ln\left(\sum_{g \in H} \exp(\mu B_{gi}) \right) + C$$

H: set of bidding agents
C: unknown constant

Estimation of bid-rent functions

Estimation of bid-rent functions

- Rosen (1974): Prices as a function of location attributes (hedonic rent model)
- Ellickson (1981): stochastic bid approach, undetermined model ➔ relative prices
- Lerman & Kern (1983): bid approach + observed price is maximum bid ➔absolute prices
  - Very detailed data is required (individual transaction prices)
  - Assumption: groups of homogeneous bidding agents
  - Validation only regarding rent and marginal willingness to pay for location attributes, not agent location distribution or price forecasting

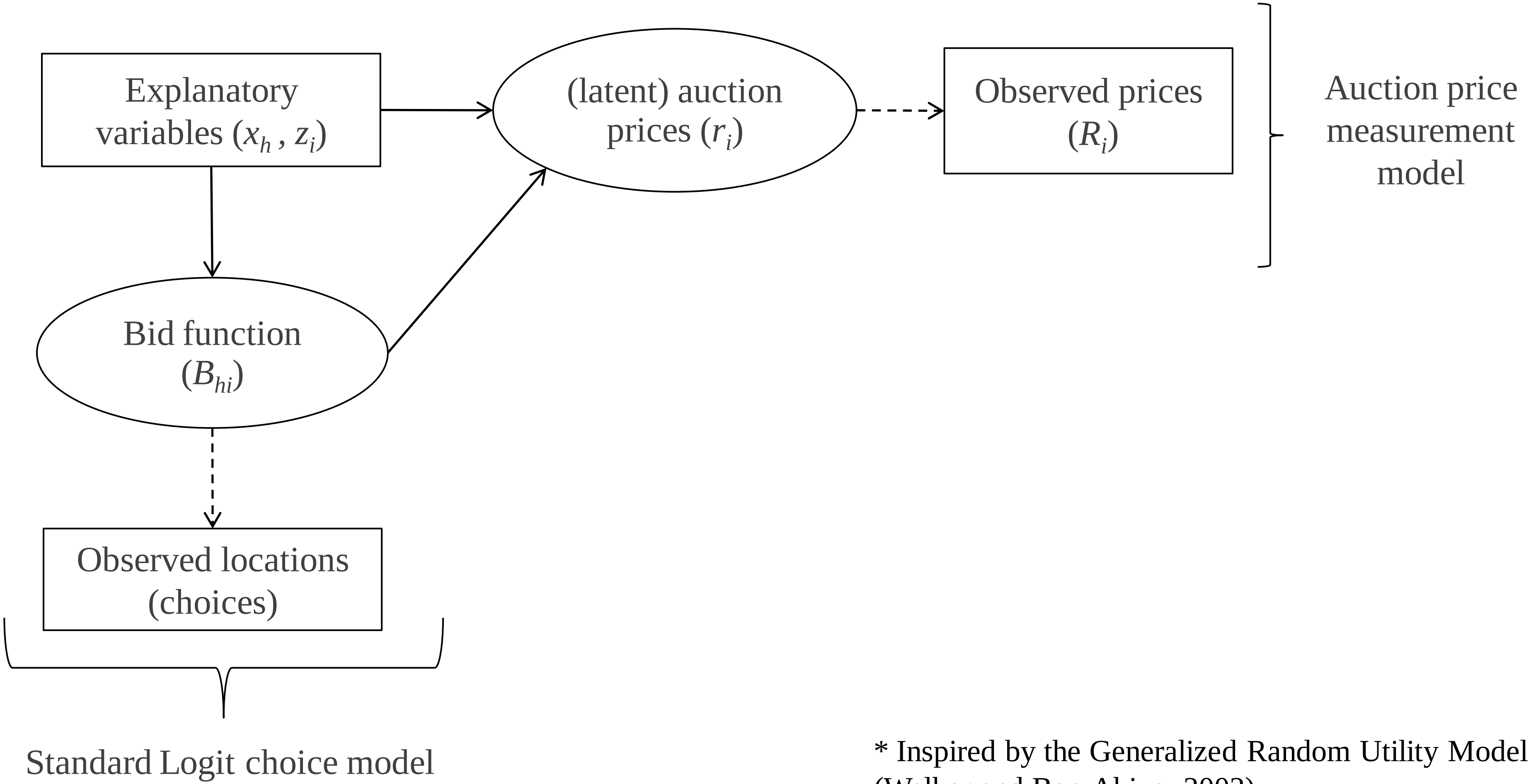
    (Gross, 1988; Gross et al 1990; Gin and Sonstelie, 1992; McMillen 1996; Chattopadhyay 1998; Muto, 2006)

Estimation of bid-rent functions

- Idea:
  - Assume structural relationship between expected outcome of the auction and observed (average) prices
  - Estimate location choice model and price model simultaneously, using observed prices as indicators

- Assumptions:
  - Auction price is a latent variable (the auction itself is a latent process)
  - All agents are potential bidders for all locations

Model with price indicator



* Inspired by the Generalized Random Utility Model (Walker and Ben-Akiva, 2002)

Model with price indicator

- Structural equation for prices:

$$r_i = \frac{1}{\mu} \ln\left(\sum_{g \in H} \exp(\mu B_{gi}) \right)$$

- Measurement equation for prices:

$$R_i = a + \gamma \cdot r_i$$

$$\sim N(0,\sigma) \Rightarrow \quad f(R_i \mid r_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{R_i - a - \gamma \cdot r_i}{2\sigma^2} \right)$$

- Likelihood:

$$L = \prod_i \left(\prod_h \left(P_{h/i} \cdot f(R_i \mid r_i) \right)^{y_{hi}} \right)$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Case study: Brussels

- Data collected for a FP7 European Union project (SustainCity)
  - Census 2001 (aggregated information by zone)
  - Household survey 1999 (~1300 observations), no detail on housing attributes
  - Average transaction prices by commune and 2 types of dwelling (house or apartment) from 1985 to 2008
  - Other geographical, land use databases

- 1267997 households, 1274701 dwellings
- 157 communes
- 4975 zones
- 4 types of dwelling (with average attributes per zone)
  - Isolated house
  - Semi-isolated house
  - Joint house
  - Apartment

Case study: Brussels

Bid function specification for location (bid) choice model (Ellickson):

$$B_{hvi} = \beta_{surf} \cdot surf_{vi} \cdot \ln(N_h) + \beta_{sup} \cdot Q_i^{sup} \cdot N_h^{sup} + \beta_{house} \cdot \lambda_{vi}^{house} \cdot N_h +$$
$$\beta_{mid\_inc} \cdot I_i \cdot \gamma_h^{mid\_inc} + \beta_{high\_inc} \cdot I_i \cdot \gamma_h^{high\_inc} + \beta_{trans} \cdot Y_i^{trans} \cdot \gamma_h^{cars=0} +$$
$$\beta_{trans2} \cdot Y_i^{trans} \cdot \gamma_h^{cars>1} + \beta_{comm} \cdot Y_i^{comm} \cdot \ln(N_h) + \beta_{off} \cdot Y_i^{off} \cdot W_h + \beta_{green} \cdot Y_i^{green} \cdot W_h$$

- $surf_{vi}$ is the average surface of a residential unit in buildings type v in zone i. The building types consider 3 types of house (fully-detached, semi-detached and attached) and apartments.
- N_h is the size (number of individuals) of a household.
- W_h is number of active individuals (workers) in a household
- N_h^{sup} is number of persons in the household who achieved a university degree as their maximum education level.
- Q_i^{sup} is percentage of the population in zone i with a superior level education-degree.
- I_i is the average income of zone i (calculated from tax declarations)
- Y_i^{trans} is a measurement of the quality of public transport for zone i (accessibility)
- Y_i^{comm}, Y_i^{off}, Y_i^{green} are measurement of the presence of commerce, offices and public green areas respectively

Case study: Brussels

Table 1: Estimation results for Brussels

Parameter	Standard Logit			Logit with price indicator		
	Value	Std err	t-test	Value	Std err	t-test
β_{surf}	0.00636	0.00261	2.43	0.000311	0.000225	1.38*
β_{mid_inc}	0.0439	0.0111	3.94	-0.00317	0.00717	-0.44*
β_{high_inc}	0.0574	0.0153	3.76	0.0161	0.00998	1.61*
β_{sup}	0.403	0.108	3.73	0.728	0.0739	9.84
β_{trans0}	0.408	0.136	3	0.599	0.0849	7.06
β_{trans2}	-0.532	0.153	-3.48	-0.31	0.0791	-3.91
β_{house}	0.461	0.0615	7.5	0.0563	0.00702	8.03
β_{comm}	-1.34	0.278	-4.83	-0.0366	0.031	-1.18*
β_{green}	-0.349	0.0717	-4.86	0.136	0.0201	6.74
β_{off}	-0.295	0.0931	-3.16	0.0896	0.0413	2.17
α	-	-	-	-16.4	3.23	-5.08
γ	-	-	-	1.92	0.229	8.39
σ	-	-	-	-1.92	0.0225	-85.48
Final Log-Likelihood	-7011.03			-6387.76 (-7091.13**)		
Likelihood ratio-test	232.44			1478.97 (72.23**)		

*parameters not significant at the 95% level

** log-likelihood considering only the choice probabilities

Estimation performed with PythonBiogeme (Bierlaire and Fetiarison ,2010)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Case study: Brussels

Table 2: Estimation results for Brussels

Parameter	Standard Logit			L&K		
	Value	**Std err**	**t-test**	**Value**	**Std err**	**t-test**
β_{surf}	0.00636	0.00261	2.43	-0.00136	0.000855	-1.59*
β_{mid_inc}	0.0439	0.0111	3.94	0.0194	0.00608	3.19
β_{high_inc}	0.0574	0.0153	3.76	0.0474	0.00796	5.95
β_{sup}	0.403	0.108	3.73	0.416	0.0669	6.22
β_{trans0}	0.408	0.136	3	-1.01	0.0716	-14.1
β_{trans2}	-0.532	0.153	-3.48	-0.226	0.0887	-2.54
β_{house}	0.461	0.0615	7.5	0.0167	0.0182	0.92*
β_{comm}	-1.34	0.278	-4.83	-0.768	0.0977	-7.85
β_{green}	-0.349	0.0717	-4.86	0.286	0.0367	7.78
β_{off}	-0.295	0.0931	-3.16	-0.767	0.0533	-14.38
μ	-	-	-	1.66	0.0173	95.74
Final Log-Likelihood	-7011.03			-7569.645 (-11813.1**)		
Likelihood ratio-test	232.44			1478.97 (72.23**)		

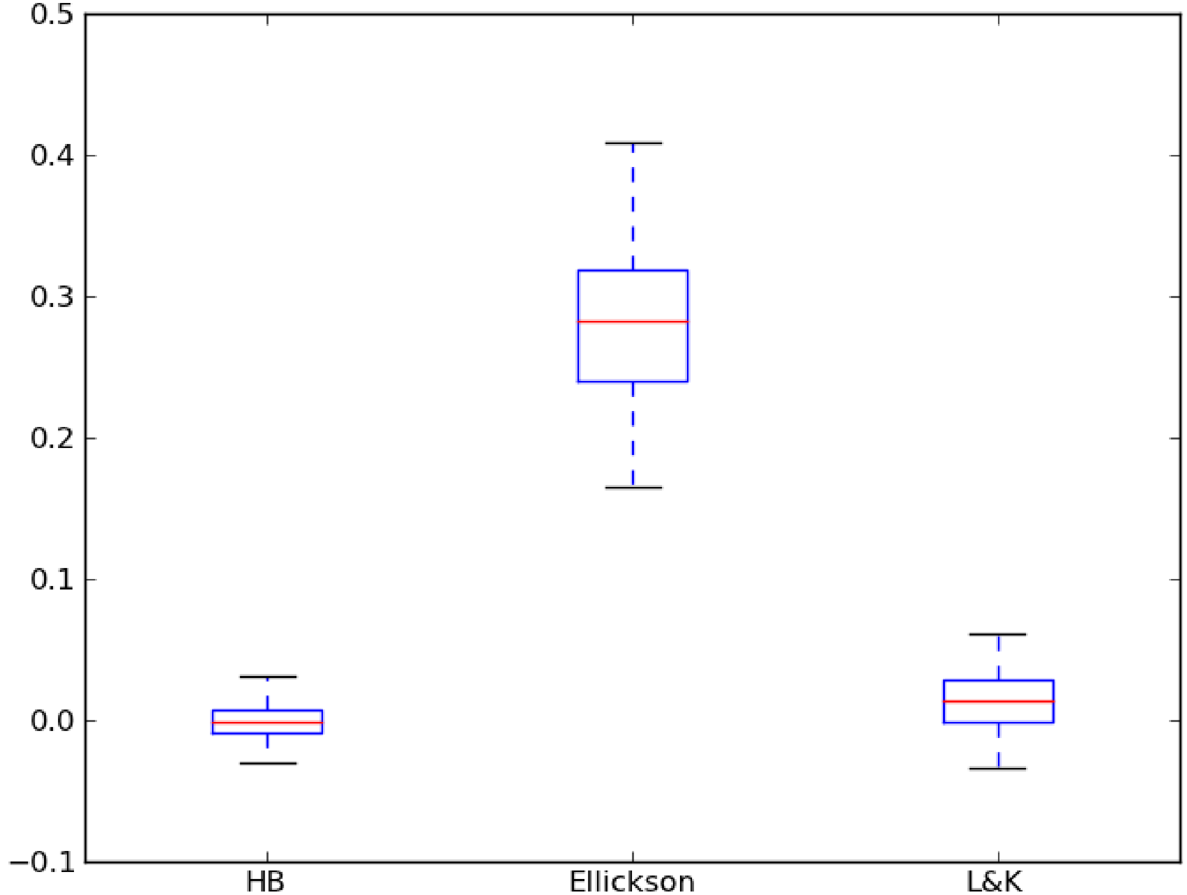
*parameters not significant at the 95% level

** log-likelihood considering only the choice probabilities

Estimation performed with PythonBiogeme (Bierlaire and Fetiarison ,2010)

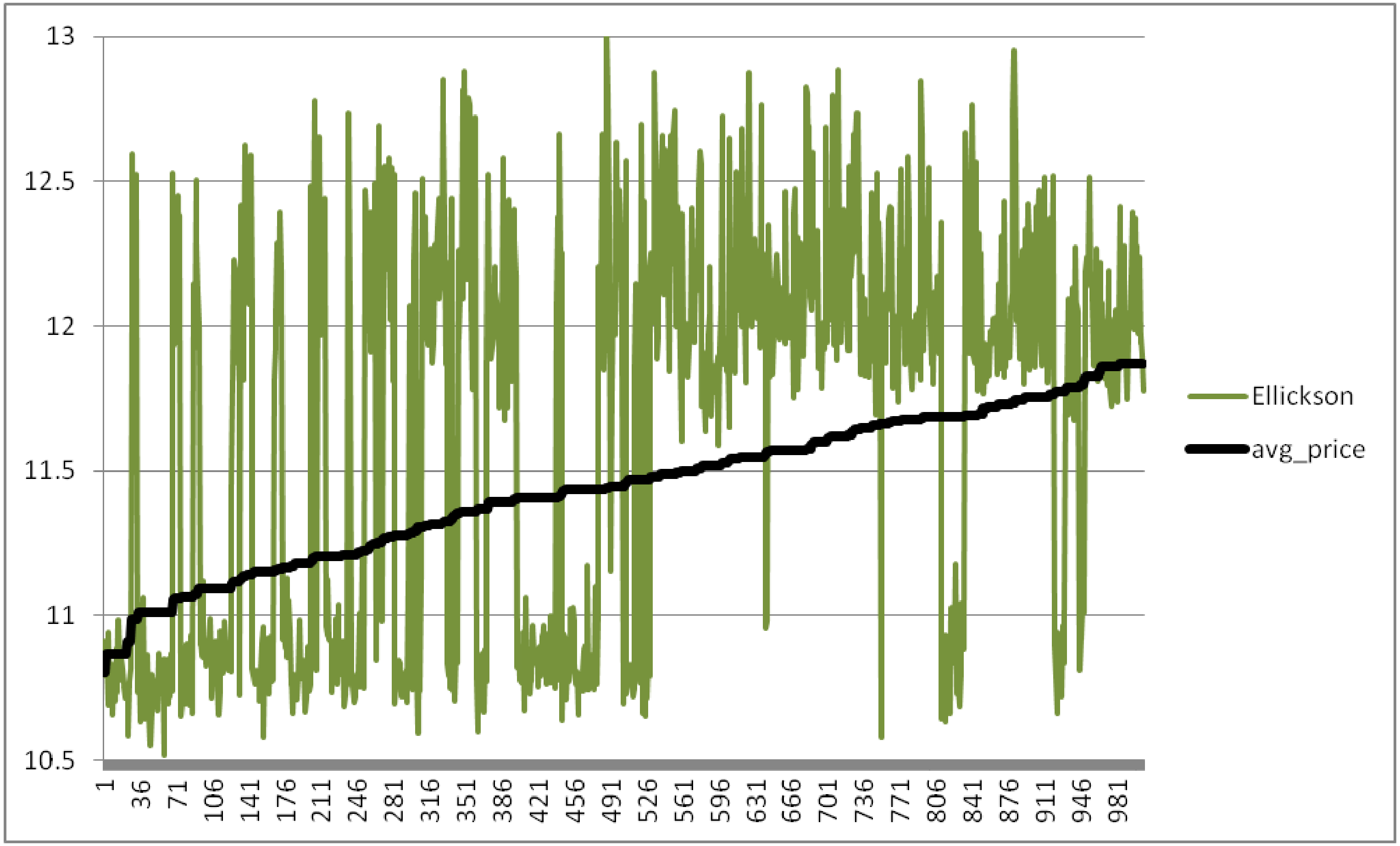
Case study: Brussels

- Prices per commune and type (% error) (over estimation dataset)



Case study: Brussels

- Prices (over estimation dataset)

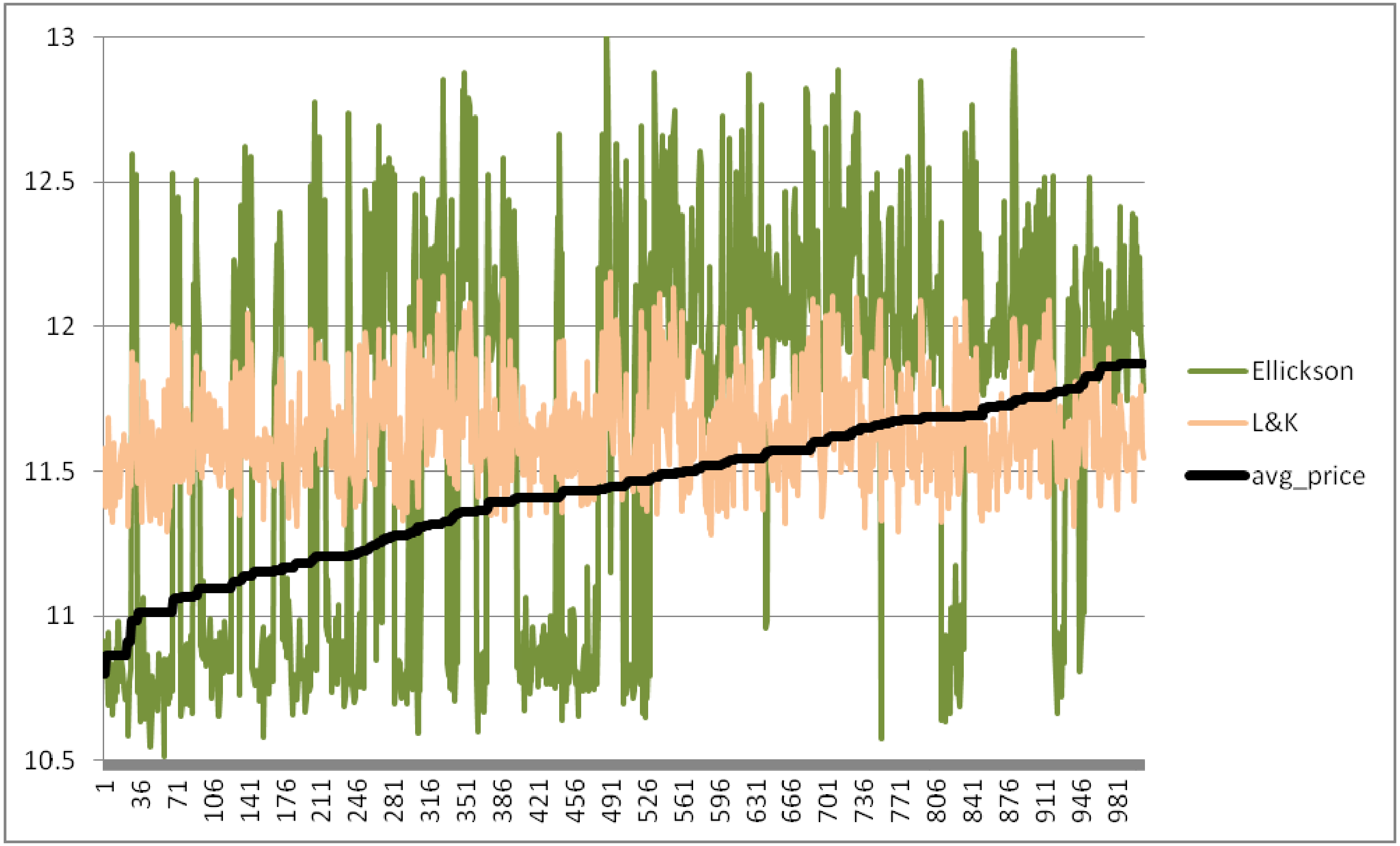


TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

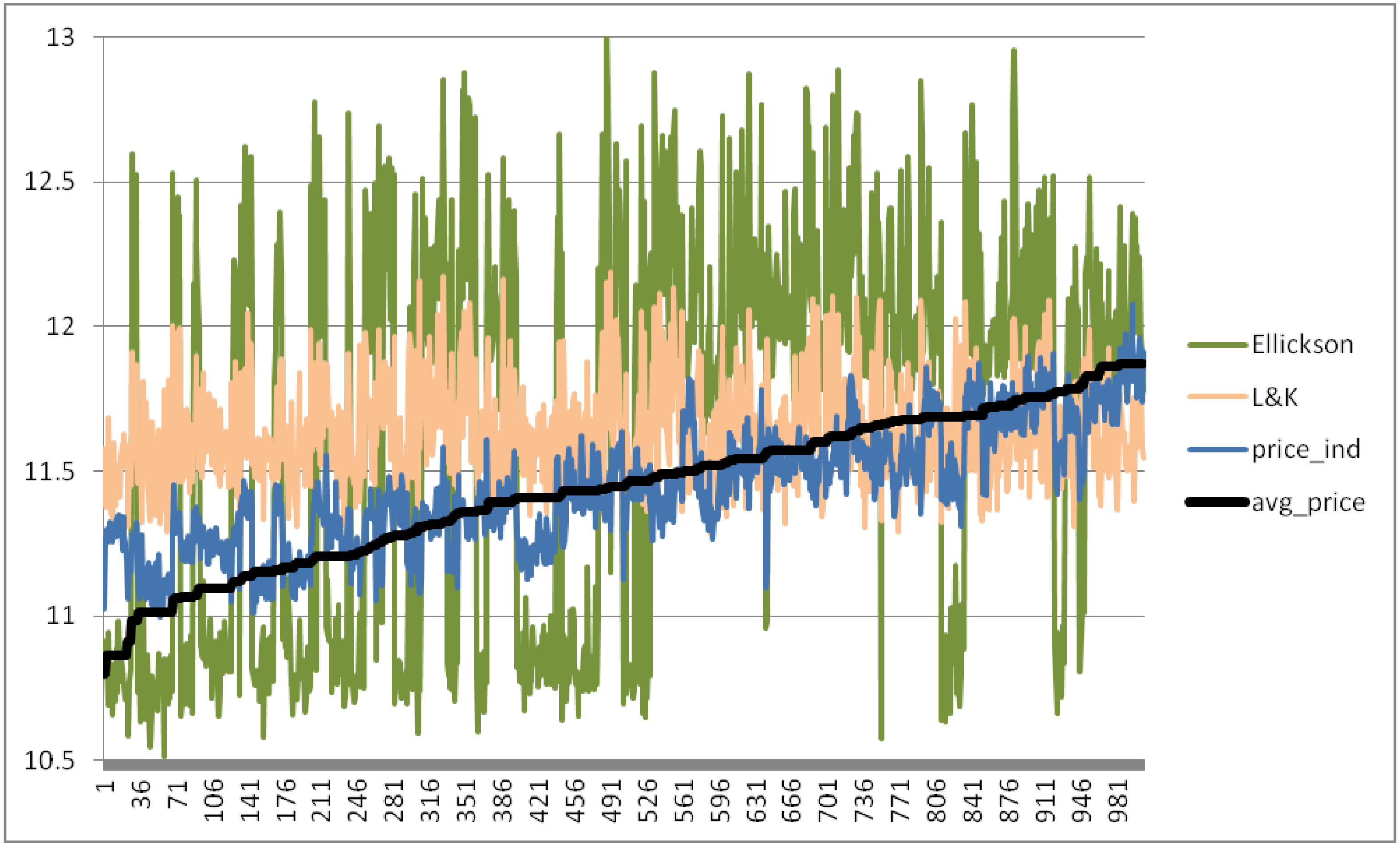
Case study: Brussels

- Prices (over estimation dataset)



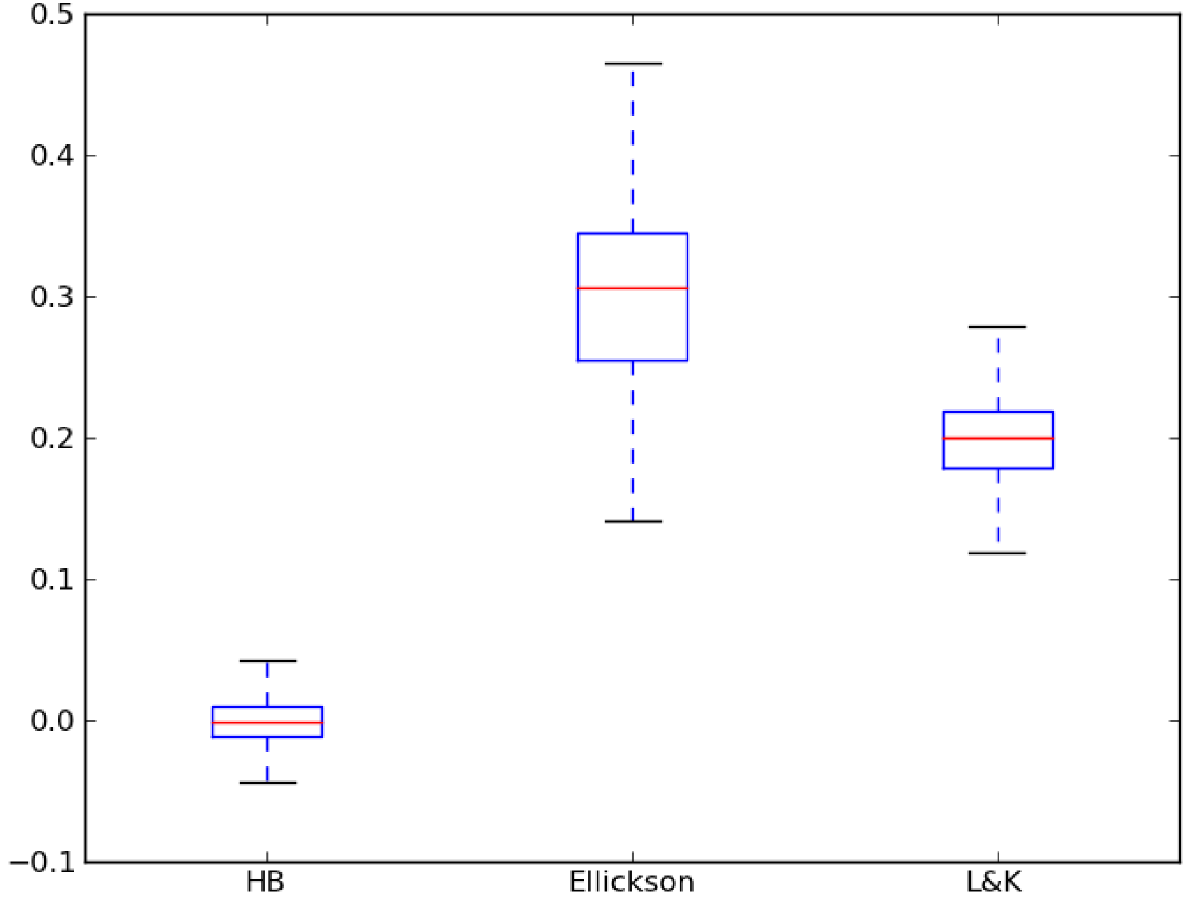
Case study: Brussels

- Prices (over estimation dataset)



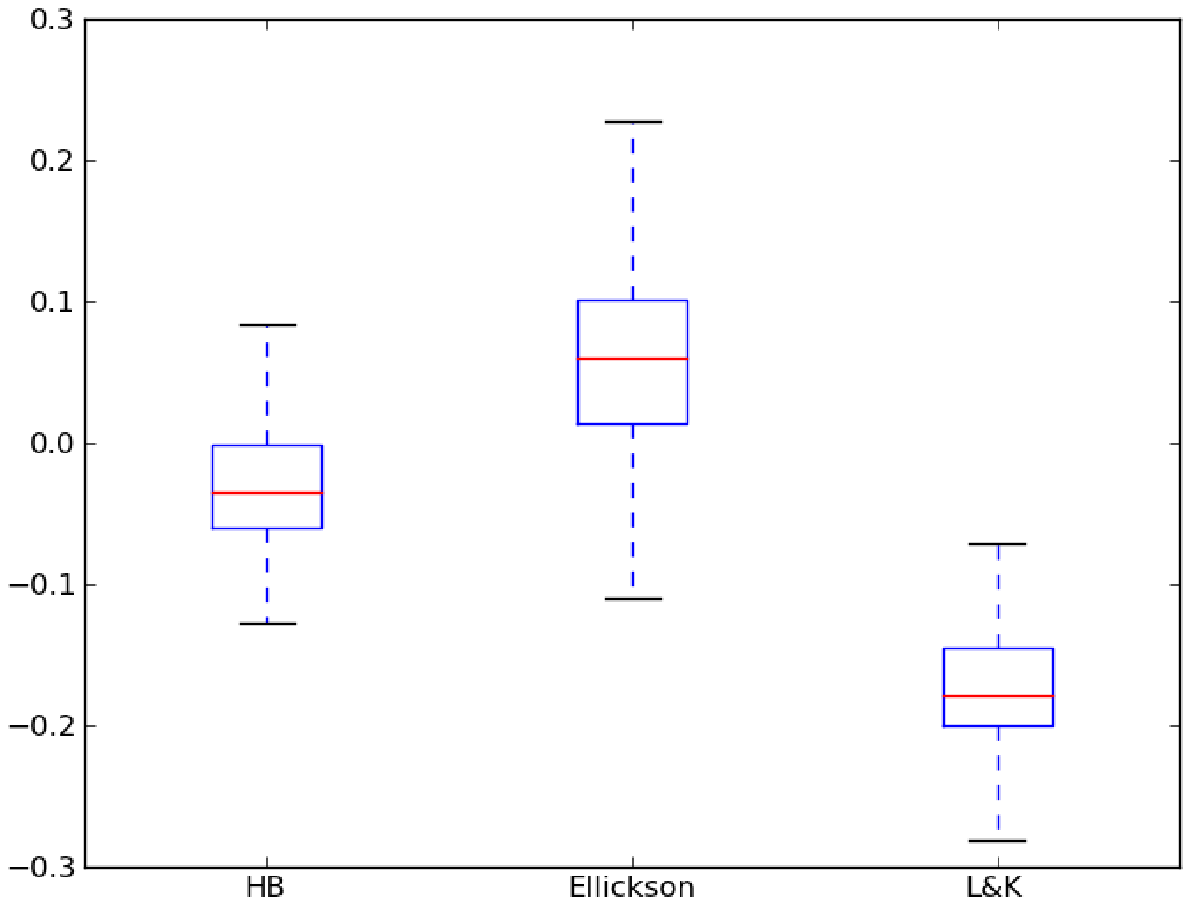
Case study: Brussels (forecasting/validation)

- Prices per commune and type (% error) (over full supply for 2001)



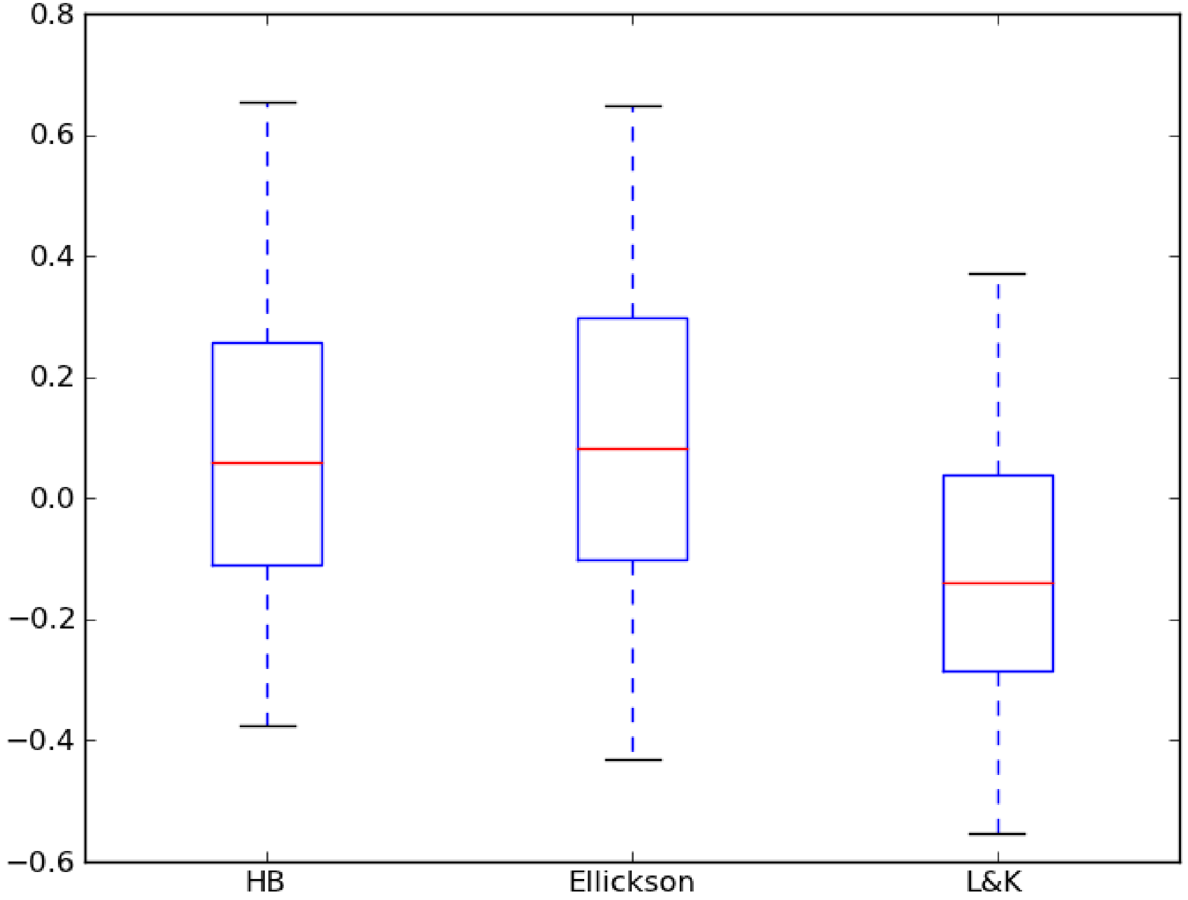
Case study: Brussels (forecasting/validation)

- Number of people per commune (% error)



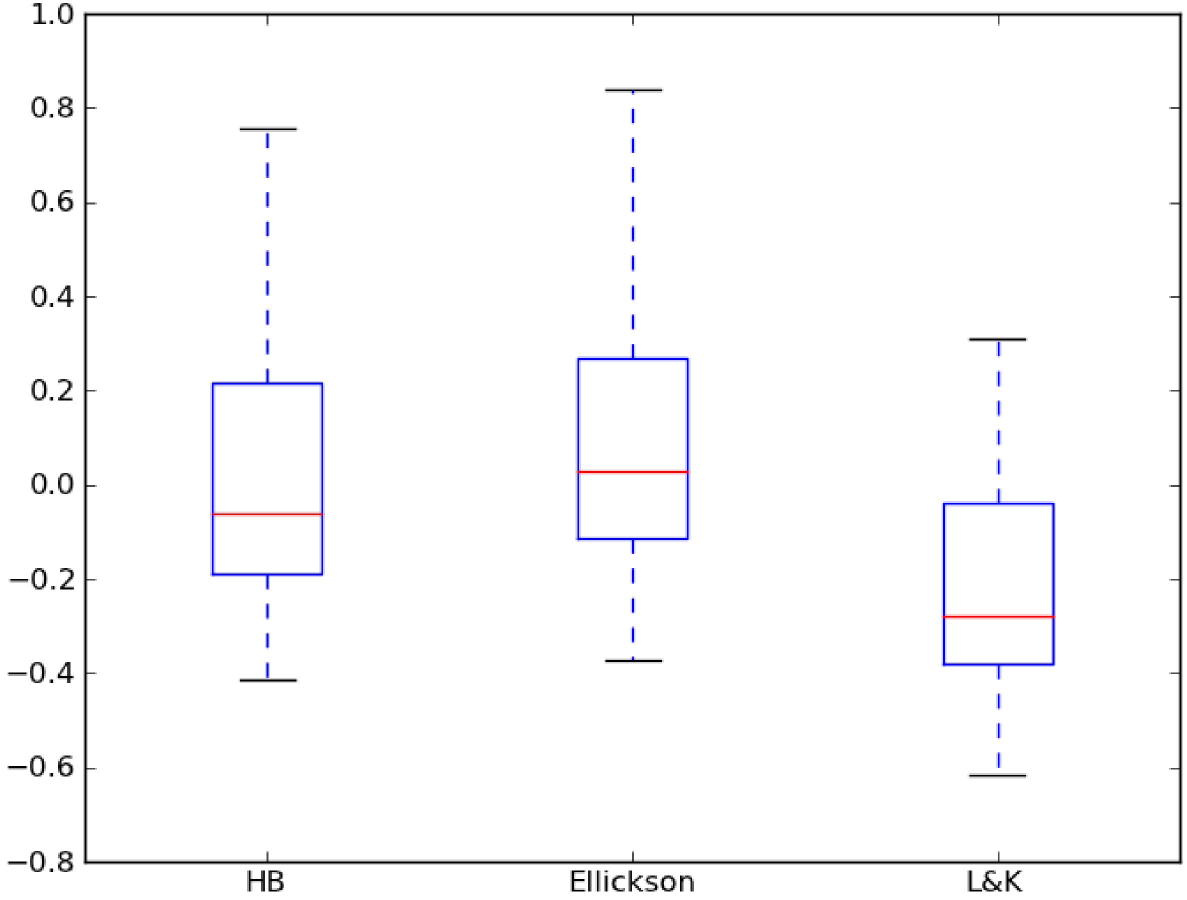
Case study: Brussels (forecasting/validation)

- Number of people with univ degree per commune (% error)



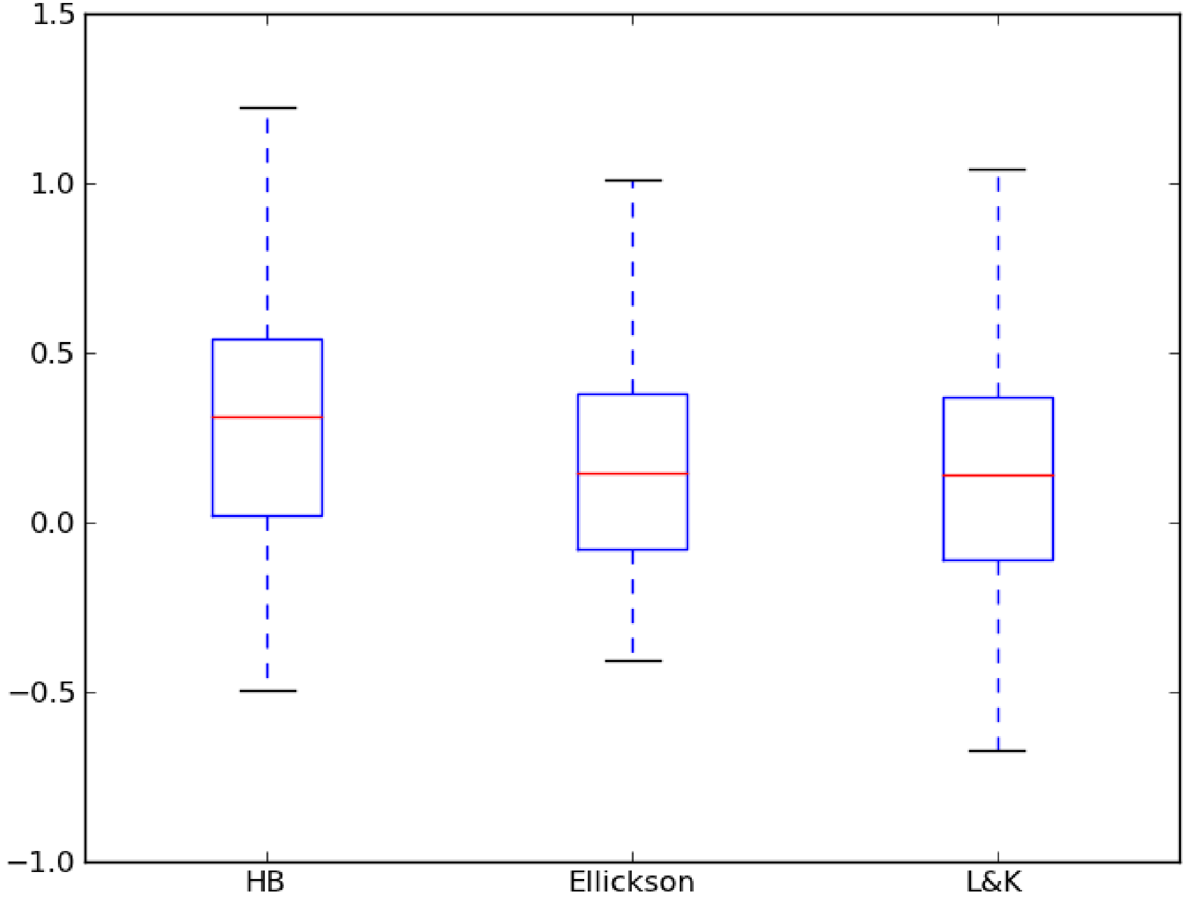
Case study: Brussels (forecasting/validation)

- Number of households with 2+ cars (% error)



Case study: Brussels (forecasting/validation)

- Number of households with 0 cars (% error)



Discussion

- The proposed estimation method finds estimates that reproduce the location distribution of agents and the average market prices of dwellings better than other methods

- Proposed method requires less detailed data ➔ more suitable for extensive land use models

- Well estimated bid functions (willingness to pay) allow to generate a good forecast of the transaction prices, without the need of hedonic price models ➔ this helps if we want to microsimulate using a bid approach

Bid-auction framework for microsimulation of location choice

Microsimulation with a bid approach

- When bids are simulated and we get:
  - Spatial distribution of agents
  - Real estate prices
- But, in order to account for competition between agents for scarce goods, we need market clearing
  - Through hedonic price models (UrbanSim)
    - Simple but not real market clearing
  - Individual auctions (ILUTE)
    - Expensive in computational terms
  - Equilibrium (MUSSA)
    - Aggregated approach

The market clearing problem

Joint probability of household *h* occupying location *i*:

$$P(i,h) = P(i \mid h)P(h) = P(h \mid i)P(i)$$

$P(h \mid i)$ Maximum bid probability

$P(i \mid h)$ Maximum surplus (utility) probability

$P(i)$ Selling probability

$P(h)$ Locating probability

Re-visiting Equilibrium

- In equilibrium models it's usually assumed that supply (*S*) equals demand (*H*)

$$P(h) = P(i) = 1 \quad \forall h, i \quad \Rightarrow H = S$$

- Possible equilibrium conditions:

$$\sum_h P(i,h) \Rightarrow \sum_h P(i \mid h) P(h) = P(i) = 1 \quad \forall i$$ **(everything is sold)**

$$\sum_i P(i,h) \Rightarrow \sum_i P(h \mid i) P(i) = P(h) = 1 \quad \forall h$$ **(everyone is located)**

Re-visiting Equilibrium

- Market clearing can be achieved by imposing one of the equilibrium conditions and finding prices/bids that produce them

$$\exists r_i : \sum_h P(i \mid h) = 1 \quad \forall i$$ **(prices clear the market)**

$$\exists b_h : \sum_i P(h \mid i) = 1 \quad \forall h$$ **(bids clear the market)**

Due to interdependence, these are usually fixed point problems

Re-visiting Equilibrium

- If we have an auction market and the best bidder rule is observed, adjusting prices or bids is equivalent in equilibrium
- When market conditions change (supply, demand, etc) utility levels of the decision makers have to be adjusted, this is reflected in the level of the prices or bids

→ idea: quasi-equilibrium

Quasi-equilibrium

- Periodical location of new and re-locating agents, given exogenous supply
- Assumption: all households looking for a location are located somewhere $P(h)=1 \quad \forall h$
  - Total supply must be greater or equal than total demand $\Rightarrow H \leq S$
  - Not all locations are necessarily used $P(i)\leq 1 \quad \forall i$

Quasi-equilibrium

- No equilibrium ➔
  - no perfect information (aggregate supply, previous prices)
  - No iterative negotiation/bidding
  - No absolute adjustment of bids/prices
- Instead, adjustment of "perception" of agents that goes in the direction of an equilibrium but does not solve it.

Quasi-equilibrium

- Algorithm (in each period):
  - All agents (H) observe the market: prices and supply $\left(r_i^{t-1}, z_i^{t-1}, S_i\right)$
  - All gents (simultaneously) adjust their bids, attempting to make their expected number of winning auctions equal to one:

$$\sum_{i \in S} q(h \mid i) = 1 \quad \forall h$$

q(h|i): perceived probability of being the best bidder for *i*

  - All agents bid at the same time for all locations ➔ prices and location distributions are defined
  - The assignment mechanism is an auction ➔ for each location a best bidder and a price is determined

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Quasi-equilibrium

Bid function: $B_{hi} = I_h - U_h + V_h(z_i) = V_h(z_i) - b_h$

- Perceived probability:

$$q(h|i) = \frac{\exp\left(V_h(z_i^t) - b_h^t\right)}{\sum_{g \in H} \exp(B_{gi}^t)} \approx \exp\left(V_h(z_i^t) - b_h^t - r_i^{t-1}\right)$$

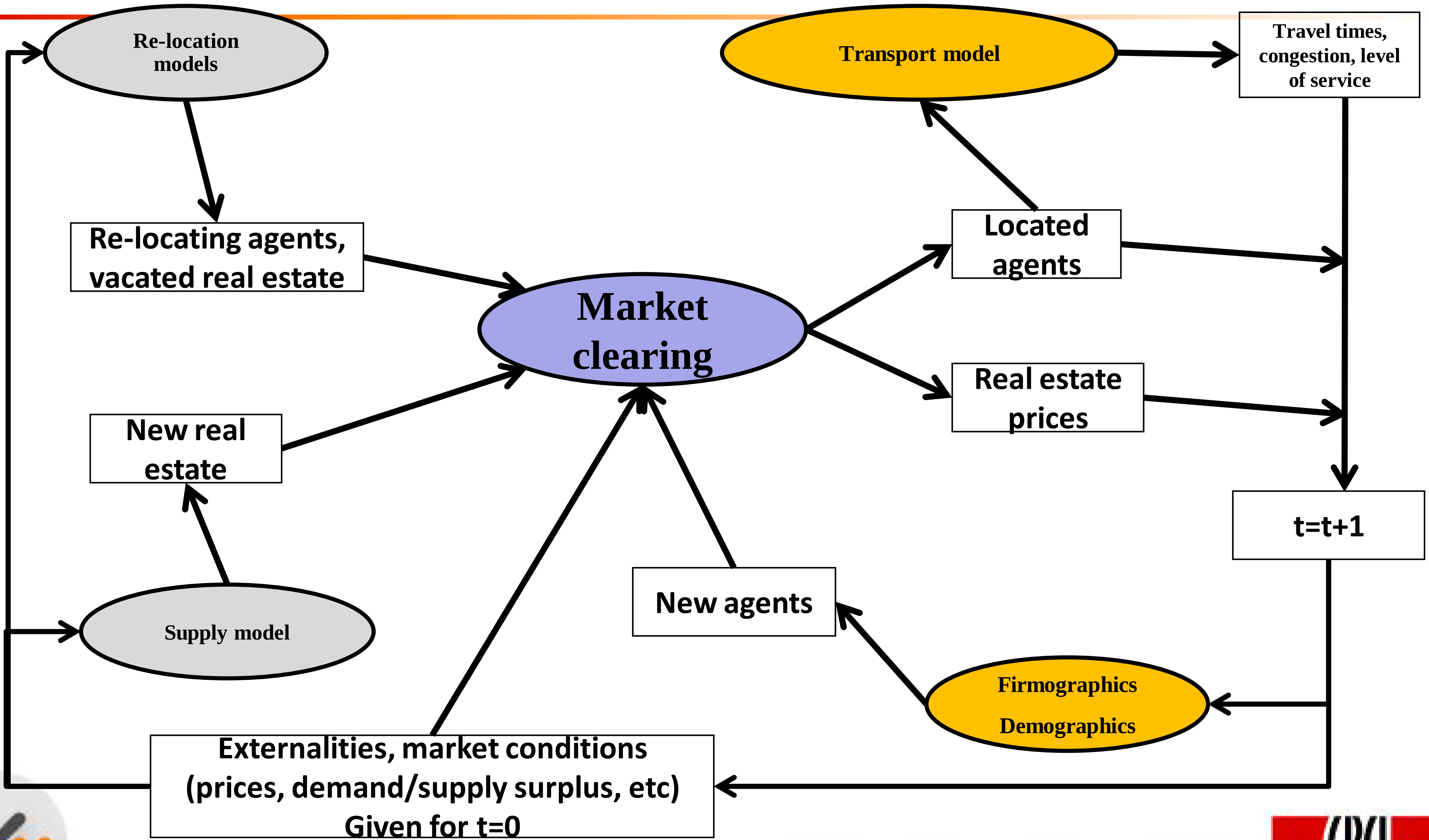
$$\sum_{i \in S} q(h|i) = 1 \quad \Rightarrow \widehat{b_h^t} = \ln\left(\sum_{i \in S} \exp\left(V_h(z_i^t) - r_i^{t-1}\right)\right)$$

Advantage: no fixed point, just evaluation of equation ➔ it is possible to apply to large populations without excessive computational cost

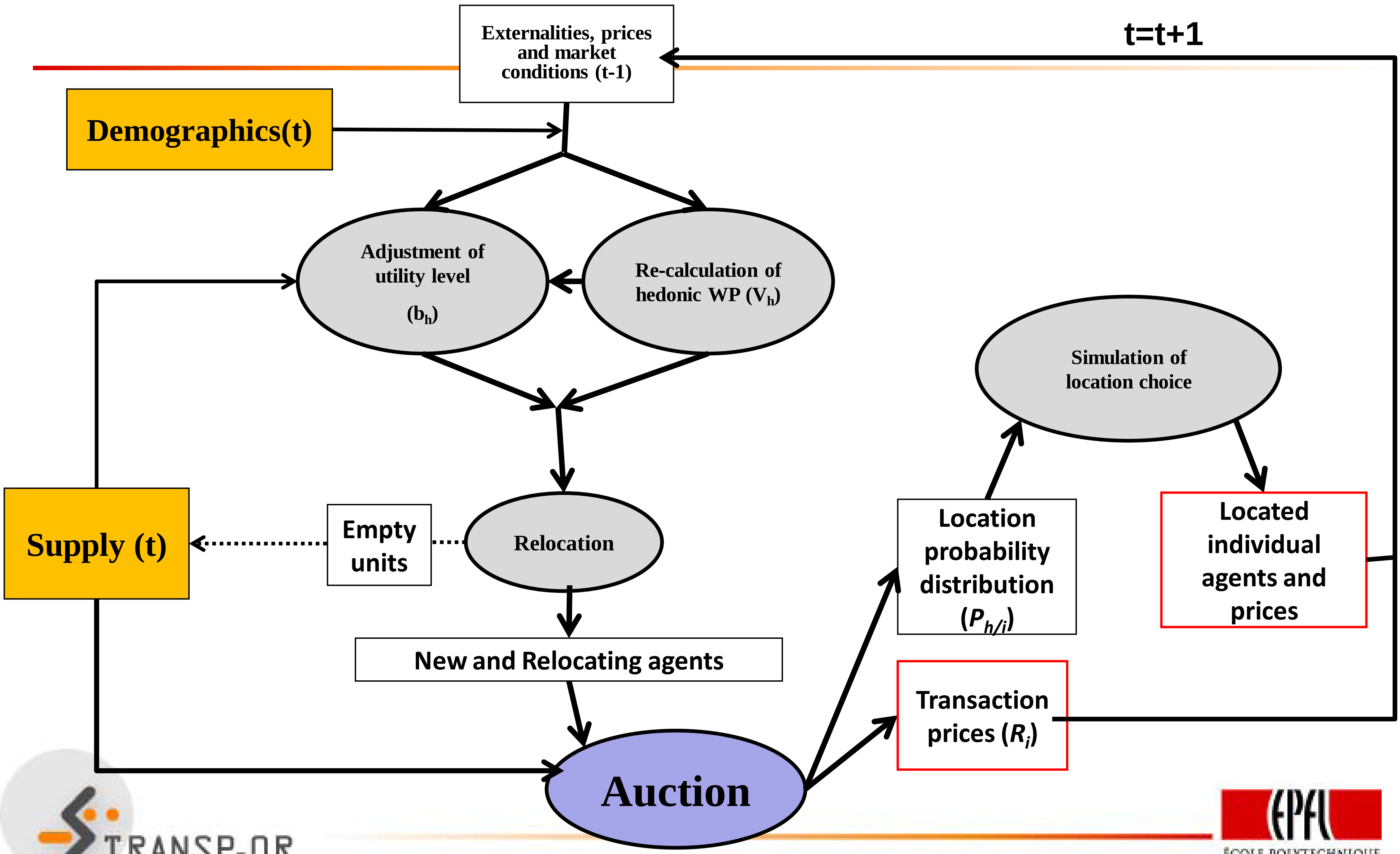
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

General framework



Market clearing

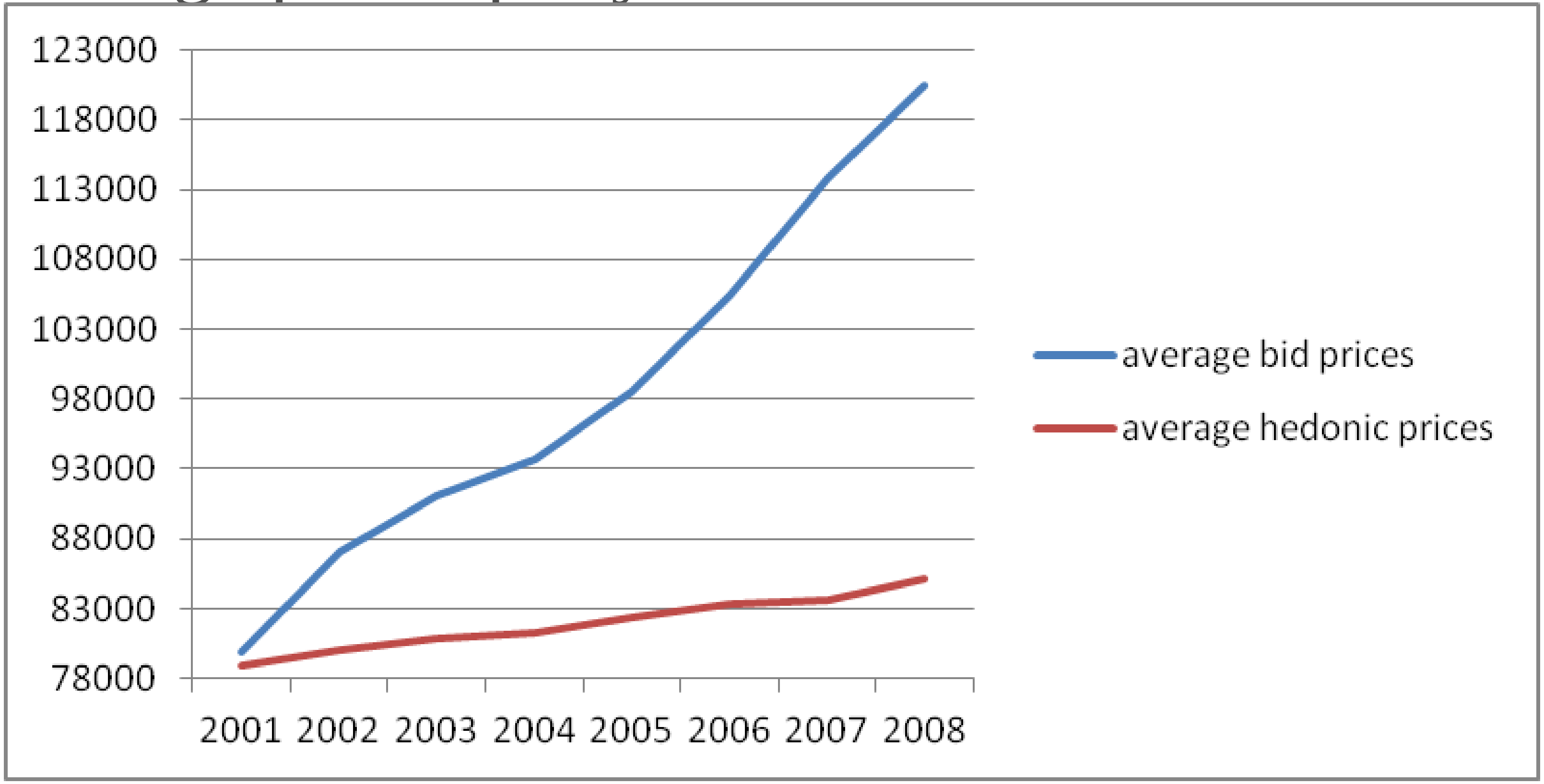


TRANSP-OR

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Some preliminary results

- Average prices per year

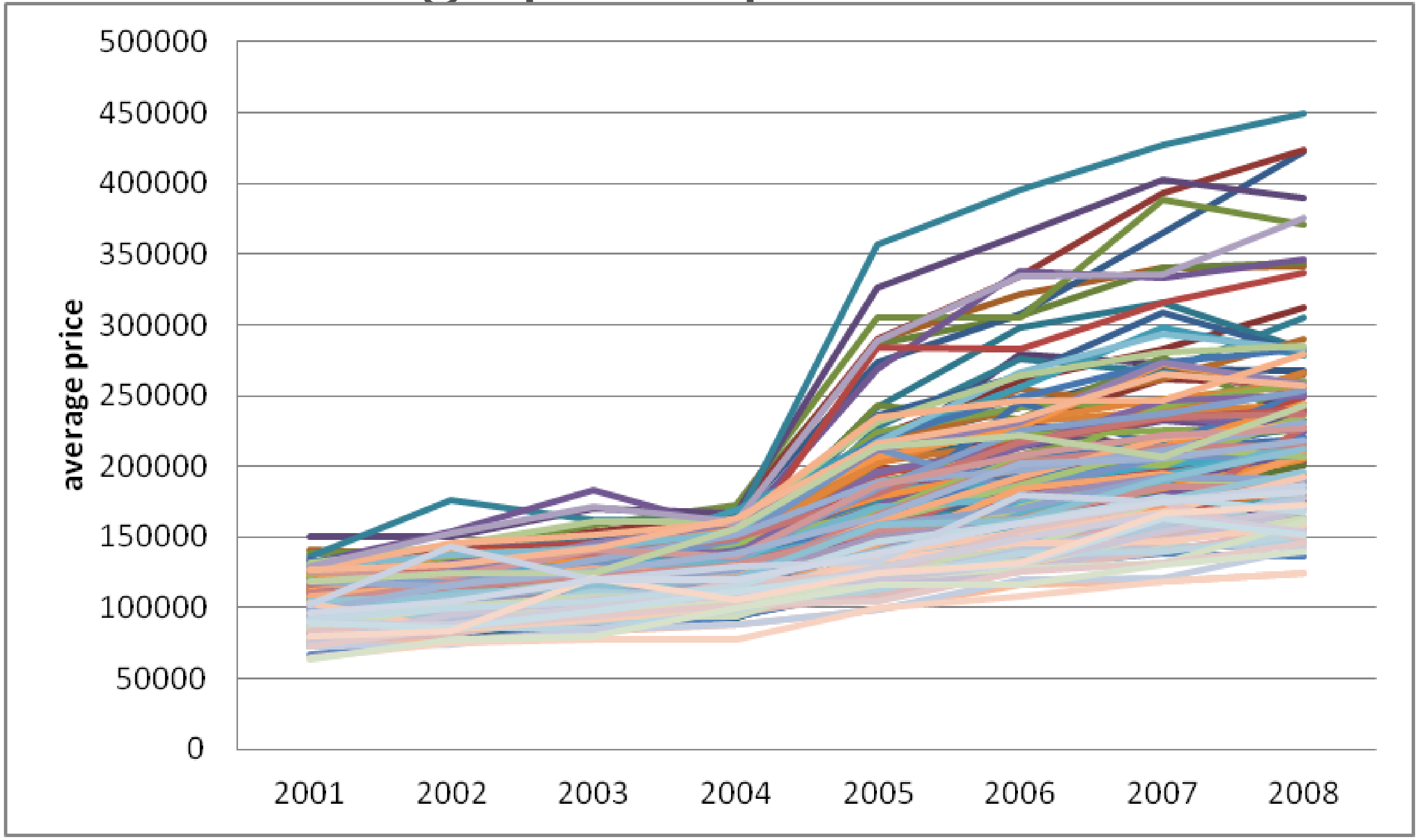


  - Average price growth: BID: 50%, HEDONIC: 7%

TRANSP-OR

EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Observed average prices per commune



- Average price growth :108%

TRANSP-OR

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Advantages

- Agents have an individual behavior but they relate to a “higher level” market mechanism through the utility level adjustment and the simultaneous auction.
- Quasi-equilibrium :
  - Demand is not cleared: utility adjustment does NOT assure allocation
  - Supply is not cleared
  - System tends to equilibrium but does not clear
- Adjustment of utility levels instead of prices allow to
  - Explain price formation (no need for hedonic price models)
  - Detect all agents utility levels, including those not active in the market, triggering future re-location

Thank you

Main assumptions of the general framework

- Auction market
- Agents adjust their utility level (individually in each period)
  - to ensure location (ex-ante expectations)
  - given market conditions: previous period rents, current supply
- Time lag:
  - In production of real estate goods:
  - In perception of attributes of locations (non-instantaneous)
- Simultaneous (macro level) bid of all agents for all locations
  - Location (best bidder) distributions and expected rents (R_i).
  - No iterative transactions.
  - Computationally simpler than transaction-specific price clearing
- Microsimulation:
  - Actual allocation following macro distributions (simulation of auctions)
  - Rents at micro level (r_i)