Time-Dependent Patterns in Freight Trip Generation

Sofia Perez-Guzman,
Ph.D. Candidate
Department of Civil and Environmental Engineering
Rensselaer Polytechnic Institute
perezs@rpi.edu

INUF 2022
Long Beach, CA, May 26th

Acknowledgments

Jose Holguin-Veras,

William H. Hart Professor

Director of the Center for Infrastructure, Transportation, and the Environment, and the VREF Center of Excellence for Sustainable Urban Freight Systems

Rensselaer Polytechnic Institute

jhv@rpi.edu

Diana Ramirez-Rios,

Lecturer

Department of Industrial and Systems Engineering

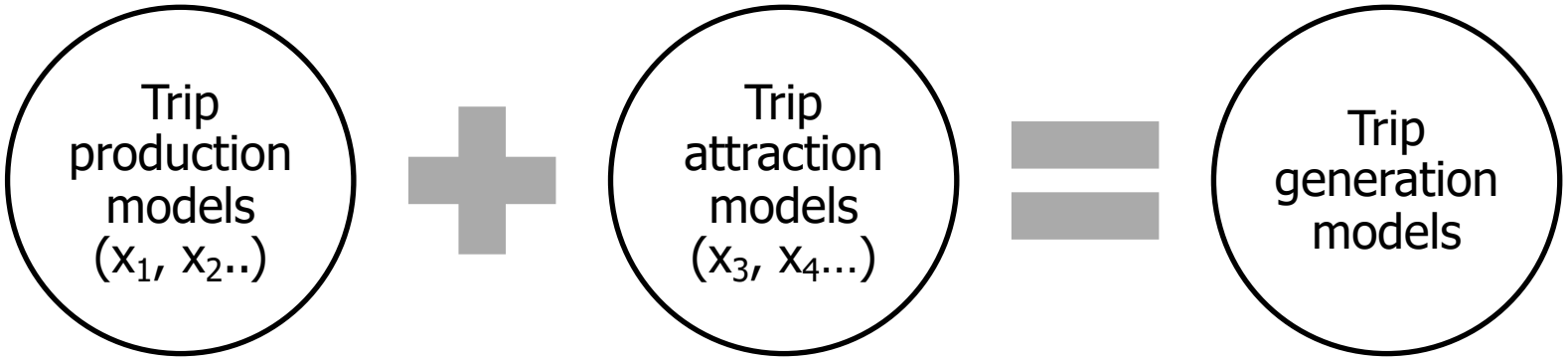
Rensselaer Polytechnic Institute

ramird3@rpi.edu

Background

The Use of Trip Generation Models

- Transportation planning, traffic impact analyses, and infrastructure design efforts rely on trip generation models to estimate the number of trips, both passenger and freight related, produced and attracted in the study area (Ortúzar and Willumsen, 2011)



***X*: Collected from surveys**

- ***Are these models transferable or stable?***
  - If so, can avoid cost and time of data collection and model estimation, and can use "imperfect" trip generation models
    - Need to know the level of errors

Stability of Demand Models

Empirical studies compare and contrast the parameters of demand models from different locations and time-periods

Kannel and Heathington, 1972; Doubleday, 1977; McCarthy, 1982; Tretvik and Widlert, 1998; Gunn, 2001; Huntsinger, 2012; Mwakalonge et al., 2012	Primary focus on geographic and temporal stability of passenger trip generation	Most overlook Freight Trip Generation (FTG), i.e., both Freight Trip Attraction (FTA) and Freight Trip Production (FTP)

Few have studied the stability of parameters of freight demand models

Oliveira-Neto et al., 2012: Geographic and temporal stability of FG	***Holguín-Veras et al., 2013:*** Geographic stability of the parameters of FTG models	***Pani et al., 2018; Pani et al., 2019; Sahu et al., 2019; Sahu and Pani, 2019:*** Geographic and temporal stability of FG at a disaggregate level

Lack of Stable Freight Demand Models

- ✓ Freight demand data collection is difficult and expensive
  - Despite availability of electronic data
- ✓ Freight surveys are hard to replace
  - The electronic data readily available usually do not include the attributes of the companies involved
- ✓ The lack of these data hampers FTG modeling efforts
  - These establishment attributes are key explanatory variables of freight demand
- ✓ Few transportation agencies collect freight data using surveys
- ✓ Less collect such data more than once
- ✓ The net result is a severe lack of FTG data to assess the stability of FTG patterns across space or time

This Research

Temporal patterns of FTG models

- Under-studied subject in freight demand modeling

Econometric techniques

- Identify time-dependent effects on FTG

Multi-year data from businesses in the NYC metropolitan area

- Employment and revenue models
- Employment-only models
- Fixed time effects per year
- Piece-wise linear time effects

Data Description

Years and Industry Sectors

- Cross-sectional samples → 2005, 2006, 2011 and 2014
- Businesses in *freight intensive sectors*, i.e., for which the production and consumption of supplies is essential of their economic activities

NAICS 2 Digits	Industry Sectors	Nickname
23	Construction	Constr
31-33	Manufacturing	Manuf
31	Food, Beverage, Tobacco, Textile, Apparel	Food
32	Wood, Paper, Chemical, Plastics Nonmetals	Wood
33	Metal, Machinery, Electronics, Furniture and Miscelaneous	Metal
42	Wholesale Trade	Wsale
44-45	Retail Trade	Retail
44	Motor Vehicle, Furniture, Electronics, Clothing	Motor
45	Sporting Goods, Hobby, Books, and Music Stores	Sport
48/49	Transportation and Warehousing	Transp
72	Accommodation and Food Services	Accom

Variables and Observations

- The data collected included economic and operational characteristics of the establishment, patterns of deliveries received and shipments sent out

Year	Relevant Variables	Number of observations		
		Carriers	Receivers	Total
2005	Number of shipments/ deliveries, business/ carrier type, fleet size, work hours, employment, location, revenue	192	180	372
2006	Number of shipments/ deliveries, business/ carrier type, fleet size, work hours, employment, location, revenue	139	200	339
2011	Number of shipments/ deliveries, business/ carrier type, shipment size, work hours, employment, location, revenue	-	263	263
2014	Number of shipments/ deliveries and service trips, business/ carrier type, shipment size, average payload, fleet size and type, employment, location, revenue	-	450	450

- The were cleaned and pooled, and variables were converted into equivalents per day

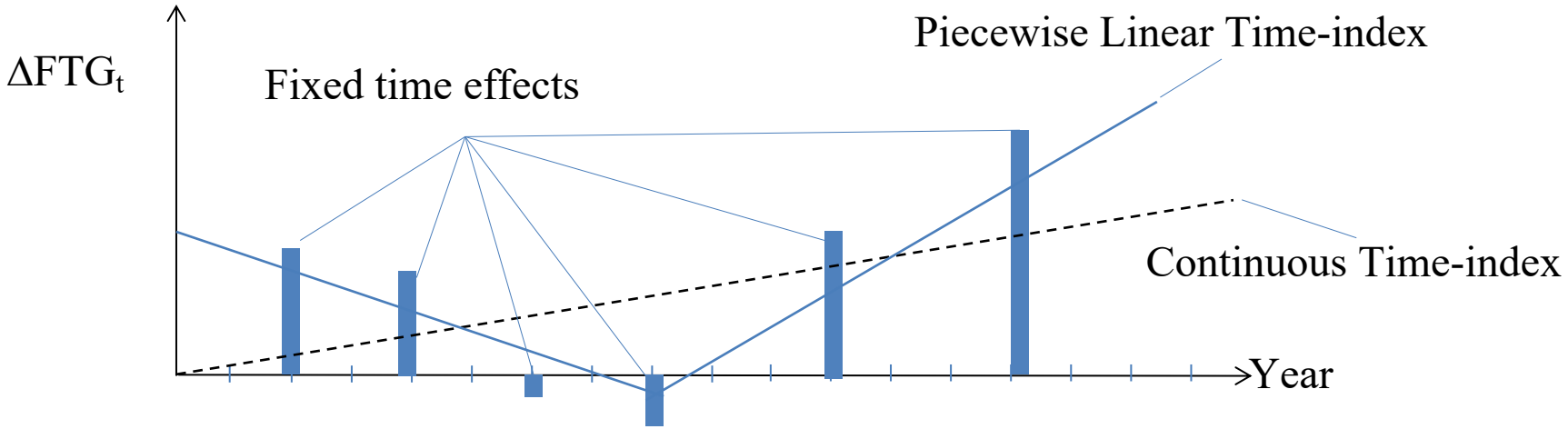
Methodology: Econometric Estimation of FTG

Establishment-Specific Variables

- ***Employment***
  - Input factors to an economic process
  - Captures the intensity of the activity at the establishment
  - Changes in this intensity translate into changes in employment level
- ***Annual revenues***
  - Indication of the market performance of the establishments
- Employment is slower to change than revenue
- Including both of them is expected to capture:
  - intensity of production
  - ups and downs of the markets

Time Variables

- Fixed factors → Yearly time-dependent effects, i.e., ’05, ’06, ’11 and ’14
- Continuous time-index → Continuous (and monotonic) time effects, i.e., elapsed since ’05
- Piecewise linear time-index → Continuous but for two periods, capturing the 2008 collapse of the financial industry



Nomenclature

f_i = Metric of FTG, freight trip attraction (FTA) or freight trip production (FTP), for establishment I

β_k = Parameters of the different independent variables k

E_i = Employment at establishment I

R_i = Revenue at establishment I

τ_{ij} = Time variables (fixed factors or piecewise linear time-index) for establishment i and year j

$E_i\tau_{ij}$ = Interaction terms between employment at establishment i and the time variables for year j

$R_i\tau_{ij}$ = Interaction terms between revenue at establishment i and the time variables for year j

Econometric Forms

- Linear:

$$f_i = \beta_0 + \beta_E E_i + \beta_R R_i + \beta_{ER} E_i R_i + \sum_j \beta_{Tj} \tau_{ij} + \sum_j \beta_{ETj} E_i \tau_{ij} + \sum_j \beta_{RTj} R_i \tau_{ij}$$

- Logarithmic:

$$f_i = \beta_0 + \beta_E \ln(E_i) + \beta_R \ln(R_i) + \beta_{ER} \ln(E_i R_i) + \sum_j \beta_{Tj} \ln(\tau_{ij}) + \sum_j \beta_{ETj} \ln(E_i \tau_{ij}) + \sum_j \beta_{RTj} \ln(R_i \tau_{ij})$$

- Exponential:

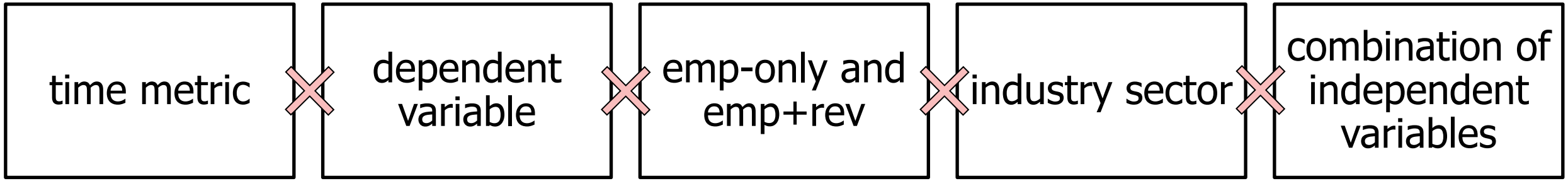
$$f_i = \exp\{\beta_0 + \beta_E E_i + \beta_R R_i + \beta_{ER} E_i R_i + \sum_j \beta_{Tj} \tau_{ij} + \sum_j \beta_{ETj} E_i \tau_{ij} + \sum_j \beta_{RTj} R_i \tau_{ij}\}$$

- Power:

$$f_i = e^{\beta_0} E_i^{\beta_E} R_i^{\beta_R} (E_i R_i)^{\beta_{ER}} \prod_j \tau_{ij}^{\beta_{Tj}} \prod_j (E_i \tau_{ij})^{\beta_{ETj}} \prod_j (R_i \tau_{ij})^{\beta_{RTj}}$$

Procedure

1. Using Ordinary Least Squares (OLS) each functional form was estimated



   - Statistical significance and conceptually validity → Best combination of independent variables: 2x2x2x11x4=352 models
2. The best functional form was selected → 2x2x2x11 = 88
3. No strong correlation between employment, revenue and time
4. Removing duplicate models → 66 final models

General results

Breakdown of Final Models by Type

Type of model	Functional Form				Total
	Linear (Lin)	Logarithmic (Log)	Exponential (Exp)	Power (Pow)	
(1) Employment Only					
Freight Trip Attraction	6	1	2	11	20
Freight Trip Production	3	0	6	12	21
Totals (1)	9	1	8	23	41
Percent FTA of Total (1)	66.7%	-	25.0%	47.8%	48.8%
(2) Employment and Revenue					
Freight Trip Attraction	10	0	4	2	16
Freight Trip Production	2	0	4	3	9
Totals (2)	12	0	8	5	25
Percent FTA of Total (2)	83.3%	0.0%	50.0%	40.0%	64.0%
Totals (1) and (2)	**21**	**1**	**16**	**28**	**66**
Percent FTA of Total	**76.2%**	**100.0%**	**37.5%**	**46.4%**	**54.5%**
Percent Total by Type of Model	**31.8%**	**1.5%**	**24.2%**	**42.4%**	

Predominant functional forms → 74.2%

Time-Dependent Patterns and Counts of Time-Dependent Effects

- ✓ 86.4% had statistically significant time-dependent effects
  - ✓ For 77.3%, the time-dependent effects were significant at 95%
  - ✓ In 5 cases (7.6%) at 90%, and in one case (1.5%) at 85%

FTG patterns are predominantly time dependent

Counts of Time-Dependent Effects by Type of Model and Time Period

Variables	FTA models					FTP models					Totals
	Lin	Log	Exp	Pow	Total	Lin	Log	Exp	Pow	Total	
Pre-2011	9	1	3	11	24	2	0	6	9	17	41
Post-2011	14	1	8	8	31	1	0	15	13	29	60
Totals	23	2	11	19	55	3	0	21	22	46	101
Percent	22.8%	2.0%	10.9%	18.8%	54.5%	3.0%	0.0%	20.8%	21.8%	45.5%	

- ✓ Most of the time effects in power (40.6%) and exponential (31.7%) forms
- ✓ 54.6% of the time effects were in FTA models

Counts of Time-Dependent Effects

Counts of Time-Dependent Effects by Sign and Time Period

Variables	Employment-Only Models						Employment and Revenue Models						Totals
	FTA models			FTP models			FTA models			FTP models			
	(+)	(-)	Total	(+)	(-)	Total	(+)	(-)	Total	(+)	(-)	Total	
Pre-2011	9	4	13	11	1	12	11	0	11	5	0	5	41
Post-2011	14	2	16	17	1	18	11	4	15	9	2	11	60
Totals	23	6	29	28	2	30	22	4	26	14	2	16	101
Percent	22.8%	5.9%	28.7%	27.7%	2.0%	29.7%	21.8%	4.0%	25.7%	13.9%	2.0%	15.8%	

- ✓ Adding revenue as explanatory variable does not completely eliminate the time-dependent effects, it only reduces their number:
  - ✓ 41.5% effects in Emp+Rev vs 58.4% in Emp-only

Other time-related factors—not captured by revenue—affect the FTG temporal patterns, e.g., e-commerce

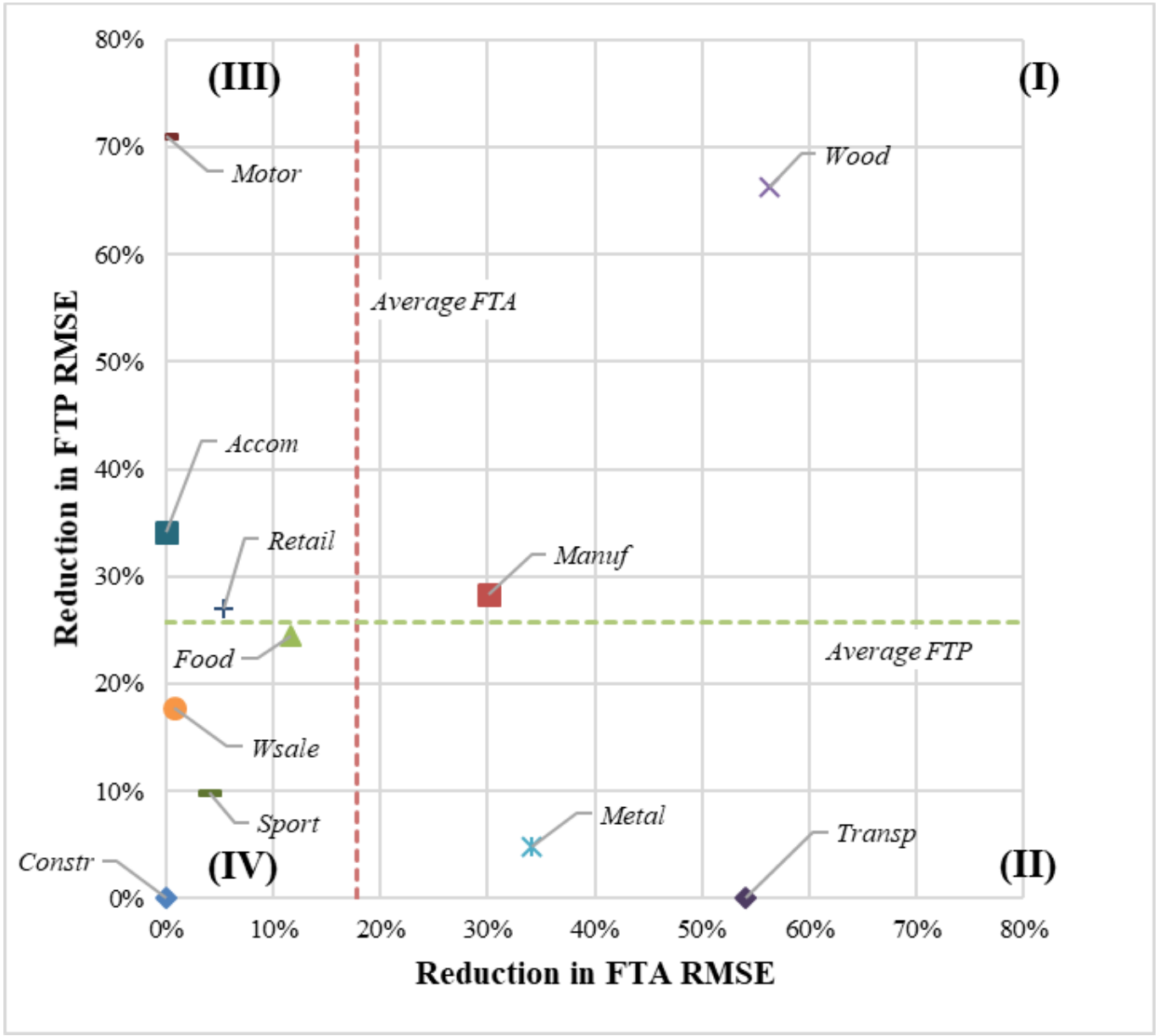
- ✓ More positive effects in the post-2011 period than in the pre-2011 period

Time-Dependent Effects by Sector

Model	FTA				FTP			
	Employment Only		Employment + Revenue		Employment Only		Employment + Revenue	
Sector	TD Effects?	# of Effects	TD Effects?	# of Effects	TD Effects?	# of Effects	TD Effects?	# of Effects
Constr	No	-	No	-	Yes	2	No model	-
Manuf	Yes	4	Yes	4	Yes	4	Yes	3
Food	Yes	3	Yes	5	Yes	4	Yes	3
Wood	Yes	3	Yes	2	Yes	4	Yes	4
Metal	Yes	3	No model	-	Yes	2	No model	-
Wsale	Yes	4	Yes	3	Yes	4	Yes	4
Retail	Yes	3	Yes	6	Yes	3	No model	-
Motor	Yes	2	Yes	4	Yes	4	No model	-
Sport	Yes	4	No	-	Yes	2	No model	-
Transp	Yes	3	No model	-	No	-	No	-
Accom	No	-	Yes	2	Yes	1	Yes	2

- ✓ Significant throughout most industry sectors, except construction
- ✓ More time effects in manufacturing, food and retail
- ✓ More time effects in FTA than in FTP

Estimation Errors (RMSE): Static vs Time-Dependent Models



The time-dependent (Emp-only) models perform significantly better than the static models, with a few exceptions

Final Remarks

Conclusions

- Most of the time-dependent effects took the form of interaction terms with employment
  - Changes in the marginal generation of FTA and FTP at the establishment-level
- The interaction term between employment and revenue replaced some time-dependent effects found in the employment-only models
  - Not all the time effects can be captured by revenue
- Some industry sectors have stronger time-dependent effects
- Time-dependent effects outperformed the static models in all cases where a time-dependent model was found, but the improvement is not necessarily symmetrical for all sectors

Major challenge to freight transportation modeling → temporal stability of parameters is not likely to hold

Thank you!

Descriptive Statistics

- High heterogeneity across and within industry sectors

Number of deliveries received per day											
Sector	**Constr**	**Manuf**	**Food**	**Wood**	**Metal**	**Wsale**	**Retail**	**Motor**	**Sport**	**Transp**	**Accom**
Mean	4.33	5.85	2.92	8.38	6.13	6.08	6.41	6.79	5.67	0.72	3.87
Standard deviation	5.90	11.70	2.41	18.24	9.65	11.17	15.43	16.29	13.63	4.88	3.99
Minimum value	0.20	0.18	0.40	0.18	0.40	0.20	0.20	0.20	0.20	0.00	0.36
Maximum value	30	100	15	100	50	85	150	150	100	50	25
Number of observations	45	160	45	44	71	152	236	157	79	179	99

Number of shipments sent out per day											
Sector	**Constr**	**Manuf**	**Food**	**Wood**	**Metal**	**Wsale**	**Retail**	**Motor**	**Sport**	**Transp**	**Accom**
Mean	1.65	8.81	4.58	12.23	8.51	3.54	2.94	2.53	3.71	3.10	0.62
Standard deviation	3.48	14.15	7.27	15.68	15.67	12.42	11.24	10.67	12.30	4.14	2.57
Minimum value	0.00	0.18	0.55	0.18	0.36	0.00	0.00	0.00	0.00	0.18	0.00
Maximum value	20	90	25	50	90	150	100	100	100	35	20
Number of observations	44	95	25	34	36	261	245	161	84	179	99

Employment											
Sector	**Constr**	**Manuf**	**Food**	**Wood**	**Metal**	**Wsale**	**Retail**	**Motor**	**Sport**	**Transp**	**Accom**
Mean	33.44	44.34	36.64	52.55	44.62	38.33	22.16	18.93	28.26	19.08	27.16
Standard deviation	46.57	59.85	46.34	70.62	61.05	45.99	39.11	28.71	53.22	29.66	31.82
Minimum value	5.00	1.00	2.45	2.45	1.00	1.00	1.00	1.00	1.00	1.00	4.00
Maximum value	250	350	200	300	350	350	300	202	300	210	180
Number of observations	45	179	56	50	73	262	246	161	85	186	100

Revenue (million USD per year)											
Sector	**Constr**	**Manuf**	**Food**	**Wood**	**Metal**	**Wsale**	**Retail**	**Motor**	**Sport**	**Transp**	**Accom**
Mean	18.86	35.34	21.36	61.53	28.51	27.65	21.48	19.45	25.34	5.52	2.38
Standard deviation	41.69	140.20	46.74	246.39	74.87	138.27	93.94	79.74	117.33	20.83	7.53
Minimum value	0.280	0.003	0.003	0.175	0.010	0.002	0.001	0.002	0.001	0.049	0.140
Maximum value	244	1495	210	1495	349	1931	756	650	756	191	58
Number of observations	44	130	42	36	52	212	122	80	42	163	84

Nature of the Time-Dependent Effects

Time Variables	Interaction with Establishment Attributes						Total	Percent
	FTA			FTP				
	None	Employment	Revenues	None	Employment	Revenues		
Fixed factors	6	13	8	13	11	4	55	54.5%
PW Time-Index	8	15	5	8	9	1	46	45.5%
Totals	14	28	13	21	20	5	101	100%
Percent	13.9%	27.7%	12.9%	20.8%	19.8%	5.0%	100%	

- ✓ Without interaction, fixed factors are a more dominant time variable
- ✓ More of effects (47.5%) in the interaction of time and employment