

## **Diseño de un experimento de preferencias declaradas para modelos de elección de la hora de inicio de viajes incluyendo dependencia entre los niveles de los atributos**

Julián Alberto Arellana<sup>+1,2</sup>, Andrew Daly<sup>3</sup>, Stephane Hess<sup>3</sup>, Juan de Dios Ortúzar<sup>1</sup>, Luis Ignacio Rizzi<sup>1</sup>

### **Abstract**

Departure time choice has received increased attention due to high levels of congestion in some cities and the increasing implementation of TDM strategies (e.g. road pricing/road space rationing).

In recent years, most studies related to departure time choice used SP data and are based on the scheduling model estimation. However, there is no consensus regarding the design generation process for this kind of models, ensuring both realism and simplicity in presentation to respondents.

Among the key issues in developing departure time choice experiments are (a) dependence of some attribute levels on other attribute levels within the same alternative and design, (b) design attributes are not exactly the same to that shown to respondents, and even some model attributes are obtained from attributes displayed to respondents, (c) the design should be customised based on each specific respondent's trips and therefore standard attribute levels may be inadequate in terms of experiment realism.

The aim of this paper is generate realistic SP exercises to model departure time choice, including dependency between attribute levels. The design procedure described in this paper consists of several stages that can afford various attributes dependency restrictions and generate a customised design for each respondent.

This paper describes work in progress and we will report a SP design customised for about 360 respondents of a previous RP departure time survey applied to workers in Santiago. We will present results after survey implementation focused on trips to work and also related to the joint trips before/after the work.

**Keywords:** Stated preference design, Departure time models, Dependency between attributes, scheduling model.

**Words:** 248

---

<sup>+</sup> Corresponding author. Dirección: Vicuña Mackenna 4860. Tel: 3541796. email: jarellana@gmail.com

<sup>1</sup> Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile

<sup>2</sup> Departamento de Ingeniería Civil y Ambiental, Universidad del Norte.

<sup>3</sup> Institute for Transport Studies, University of Leeds

## **1. Introducción**

La congestión en las vías y en los sistemas de transporte que a diario utilizan las personas produce incomodidades y pérdidas de tiempo irrecuperables. Los inconvenientes son más evidentes en las grandes ciudades, pues cada vez se encuentran sujetas a mayores niveles de congestión causados por el crecimiento de la población y del parque automotor.

El crecimiento de la congestión en las ciudades ha hecho que las estrategias y políticas para manejar la demanda de transporte (por ejemplo: tarificación vial, implementación de horarios de trabajo flexibles, restricciones de circulación por patente, etc) sean cada vez más populares. El propósito de las estrategias y políticas para manejar la demanda de transporte es modificar el comportamiento de viaje de las personas con el fin de redistribuir la demanda hacia algunos lugares y/o hacia períodos de tiempo menos congestionados.

Ante la presencia de congestión o cuando se implementan políticas y/o estrategias para manejar la demanda de transporte, las elecciones de modo, hora y ruta de los viajeros son las que se intentan modificar pues determinan la dimensión temporal y espacial de la demanda de los sistemas de transporte. En la literatura se reconoce que ante cambios en la condición de los viajes causados por congestión o por la implementación de políticas y/o estrategias para manejar la demanda de transporte, la segunda alternativa después del cambio de ruta, es el cambio del horario de inicio de los viajes (Ortúzar and Willumsen, 2011); siendo esta incluso una estrategia más frecuente que los cambios en el modo de transporte (Bianchi et al., 1998, Hess et al., 2007a, Hendrickson and Plank, 1984).

A pesar que la elección de la hora de inicio de los viajes es una de los componentes importantes que determinan la distribución temporal y espacial de la demanda, históricamente ha recibido menos atención que las elecciones de modo y de ruta. Sin embargo, en los últimos años la elección de la hora de inicio de los viajes ha recibido mayor atención debido a los altos niveles de congestión en algunas ciudades y al incremento en la implementación de estrategias y/o políticas de manejo de la demanda.

Modelar la elección de la hora en que los individuos realizan sus viajes durante el día es importante porque puede ayudar en la planificación para el desarrollo o la construcción de nueva infraestructura de transporte que permita acoger la demanda proyectada a diferentes horas del día, desarrollar modelos de tráfico que tengan en cuenta la reacción de las personas frente a la congestión (Tseng et al., 2011), examinar las respuestas de las personas al implementar estrategias para el manejo de la demanda, y evaluar la efectividad de las políticas de transporte implementadas en una ciudad (Bhat and Steed, 2002).

Durante las últimas décadas, la mayoría de estudios relacionados con la elección de la hora de inicio de los viajes se basan en la estimación del modelo de programación de horarios (SM, *scheduling model*) propuesto por Small (1982), usando datos de preferencias declaradas (PD).

El SM es considerado como el caballo de batalla dentro de la literatura de elección de la hora de inicio de viajes. Supone que los individuos eligen el periodo de inicio de sus viajes teniendo

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

en cuenta el tiempo de viaje y la utilidad que les produce llegar temprano, tarde o a la hora, respecto a una hora deseada de llegada al destino. Es decir, el individuo maximiza su utilidad si realiza el viaje más corto, iniciando su viaje a la hora que le permita llegar a su destino en la hora más cercana posible a su hora de llegada deseada.

Los estudios de elección de la hora de inicio de los viajes hacen uso de datos PD comúnmente porque los datos de preferencias reveladas (PR) que se necesitan para lograr buenas estimaciones en los parámetros del SM son difíciles de obtener (Hess et al., 2007b, Tseng et al., 2011), ya que requieren un procedimiento de recolección muy riguroso y por tanto costoso. Respecto a los datos PD, no existe consenso respecto al diseño experimental para generar las encuestas PD. En estudios previos con PD para formular SM se han utilizado desde diseños ortogonales en diferencia (Börjesson, 2008) hasta diseños factoriales fraccionales complementados con ejercicios de simulación (Small et al., 1995, Small et al., 1999, Asensio and Matas, 2008).

Los diseños experimentales utilizados en estudios previos tienden a simplificar la complejidad de la tarea de elección y utilizan los experimentos de simulación para aumentar el realismo de las situaciones de elección. Dentro de los problemas principales que dificultan el desarrollo de experimentos de elección para SM se encuentran (a) la dependencia de algunos niveles de los atributos con los niveles de otros atributos del diseño, (b) los atributos de diseño no son exactamente los mismos que se muestran a los encuestados, o algunos atributos de los modelos son obtenidos a partir de los atributos que se muestran a los encuestados (Koster and Tseng, 2010), (c) para garantizar el realismo del experimento, el diseño debe ser personalizado y basado en las características específicas de los viajes de cada encuestado.

El objetivo de este artículo es presentar una metodología que permita generar experimentos de elección eficientes, realistas y sencillos que permitan modelar la elección de la hora de inicio de los viajes basados en el SM. Este artículo pretende contribuir al estado del arte y la práctica de diseño de experimentos PD mediante la presentación de un procedimiento de diseño que aborde los 3 problemas mencionados.

El procedimiento de diseño descrito en este artículo es un diseño bayesiano eficiente en varias etapas, que permite incluir las restricciones de dependencia entre atributos, umbrales de diferencia entre los atributos de diferentes alternativas, atributos diferentes para los encuestados diferentes a los de diseño, y que incluye valores personalizados para cada encuestado generados a partir de una encuesta previa de PR para garantizar el realismo del ejercicio. Los autores no conocen en la literatura, diseños eficientes aplicados a la construcción de experimentos de elección de la hora de inicio de viajes.

Este artículo representa un trabajo en desarrollo, donde presentaremos el diseño eficiente realizado para recolectar los datos de PD que permitan estimar un SM utilizando una muestra de cerca de 360 trabajadores de Santiago. El interés principal de este proyecto es estudiar las elecciones del horario de los viajes hacia el trabajo y después del trabajo del grupo de empleados encuestados. Se pretende contribuir a la literatura con una metodología de diseño que permita generar experimentos de elección realistas, incluyendo restricciones de dependencia entre atributos. Además se pretende mostrar la aplicación de un diseño eficiente

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

para configurar experimentos de elección de horario de inicio de viajes y se espera seguir entregando evidencias de las ventajas de utilizar experimentos generados a partir de alternativas de referencia. Específicamente, cuando las alternativas de referencia provienen de una encuesta PR previa. La gracia de este tipo de experimentos es la mejora de la eficiencia del diseño y el aumento del realismo de la encuesta, tal como lo reportan Rose et al (2008), Caussade et al (2005), y Train y Wilson (2009)

El resto del artículo se organiza de la siguiente manera. La sección 2 presenta una breve discusión relacionada a los diseños de encuestas de preferencias declaradas. La sección 3 se refiere específicamente a los modelos de elección de horario de inicio de viajes y los elementos del diseño experimental que se abordaran. La sección 4 muestra la descripción del enfoque propuesto para realizar encuestas PD con las características descritas en este artículo. La sección 5 describe un caso de estudio utilizando el procedimiento de diseño presentado en este artículo. La sección 6, finalmente se concluye acerca de los resultados encontrados.

## **2. Diseño de encuestas de preferencias declaradas**

En la actualidad, la mayoría de las aplicaciones de los modelos de elección discreta se hacen utilizando datos PD. Los datos PD presentan ciertas ventajas sobre los PR en cuanto al costo de recolección, el número de observaciones por encuestado, el tratamiento de la correlación entre atributos, y la información que conoce el modelador en cuanto a la elección y a los atributos de la misma. Sin embargo, los datos PR representan el comportamiento real de las personas y no un comportamiento hipotético como en el caso de las PD, que puede estar afectado por ciertos sesgos (afirmación, de política, etc).

En el área del transporte utilizar experimentos de elección es una práctica común para obtener datos a un costo menor que el de las PR, que permitan obtener la influencia independiente de ciertas variables sobre la elección estudiada, estimar disposiciones al pago, modelar la demanda actual o incorporar alternativas hipotéticas de transporte, entre otros. Los experimentos de elección como los define Louviere et al (2000) es un tipo de encuestas de preferencias donde se presentan múltiples situaciones hipotéticas de elección a los encuestados descritas por una información de contexto y un conjunto de alternativas y atributos que son escogidas/controladas en su totalidad por el modelador. Cabe destacar que cuando nos refiramos a datos PD en este artículo, nos estamos refiriendo a datos obtenidos a partir de experimentos de elección.

La construcción de las situaciones hipotéticas de elección que enfrentaran los encuestados, es decir la definición del contexto, las alternativas, los atributos y los niveles de los atributos a presentar es lo que se conoce como diseño de un experimento de elección. El diseño del experimento no sólo es la matriz de valores que indica donde va cada elemento del diseño dentro del experimento PD sino que va más allá, de hecho, es un proceso que demanda tiempo y se encuentra lejos de ser un proceso simple (Ortúzar and Willumsen, 2011). El diseño de un experimento de elección juega un rol crucial en los resultados que se obtienen al aplicar

## Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

una encuesta PD, pues tiene que ver con que presentamos, cómo lo presentamos y en qué cantidad lo hacemos.

El objetivo de un diseño PD es generar encuestas que maximicen la información con el menor número de observaciones. Un aspecto clave en el éxito de los diseños de experimentos PD es el realismo y la complejidad del mismo. Estas características de realismo y complejidad se deben balancear para obtener resultados adecuados y además garantizar que las personas entiendan correctamente la situación de elección que se les presenta. Para una mayor discusión en el tema, ver Ortúzar y Willumsen (2011).

Existen dos alternativas marcadas para generar diseños de experimentos de elección. La primera alternativa es la que ofrecen los modelos de diseño de experimentos tradicionales, donde se encuentran entre otros los diseños factoriales y los ortogonales, que no realizan supuestos algunos relacionados a los parámetros a estimar en los modelos de elección (Louviere et al., 2000, Street and Burgess, 2004). La otra corriente de pensamiento, intenta aprovechar información a-priori de los parámetros para generar diseños que utilicen menor cantidad de datos (Rose and Bliemer, 2008).

### 2.1 Diseños que no realizan supuestos a priori de los parámetros del modelo

Este tipo de diseños fueron los primeros en ser usados para generar experimentos de elección y aún siguen contando con gran popularidad. A su favor cuentan con la facilidad para ser generados pues vienen implementados dentro de paquetes estadísticos muy utilizados (por ejemplo, SPSS), existen páginas web donde se pueden encontrar (por ejemplo, *Neil Sloane: Home page*), o también pueden generarse a partir de los diseños publicados por Kocur et al (1982).

Los diseños anteriores se basan en la teoría del diseño experimental para modelos de regresión lineal, donde la varianza de los parámetros estimados viene dada por la ecuación 1.

$$Var(\hat{\beta}_i) = \frac{\sigma^2}{\sum x_i^2(1-r^2)} \quad (1)$$

Donde  $\sigma^2$  representa la varianza del modelo,  $x_i$  representa los niveles de los atributos, y  $r^2$  la correlación entre atributos.

El propósito de los diseños experimentales es obtener datos que generen estimaciones de los parámetros con los mínimos errores estándar posibles. Los errores estándar se obtienen de las raíces de las matrices de varianza-covarianza de los modelos. Si la correlación  $r^2$  no es cero, entonces la varianza de los parámetros estimados crece con la correlación. Por esta razón, uno de las características deseadas dentro de este tipo de diseños lineales es la orthogonalidad. La orthogonalidad se refiere a la ausencia de correlación entre los atributos de un diseño, lo cual permite que todos los parámetros sean independientemente estimables (ausencia de multicolinealidad) y además garantiza que el diseño lineal minimiza la varianza estimada de los parámetros. De esta forma, cuando se cumple la orthogonalidad en los modelos lineales, la matriz de varianzas-covarianzas está dada por la ecuación 2.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

$$MVC(\beta) = \sigma^2(X'X)^{-1} \quad (2)$$

Para el caso de modelos lineales, minimizar los errores estándar entonces es equivalente a maximizar la norma de la matriz de información que viene dada por  $(X'X)^{-1}$ .

El problema de forzar la propiedad de la ortogonalidad en los diseños es que esta debe mantenerse a lo largo de todos los atributos y todas las situaciones de elección, por lo que si se generan diseños de un gran número de elecciones, deben usarse bloques en la práctica y para garantizar la ortogonalidad estos bloques deben ser igualmente usados. Aún peor, en algunos casos para garantizar la ortogonalidad en el diseño, se producen situaciones repetidas o donde una alternativa de elección claramente domina a la otra en términos de preferencia (es decir, es preferida sin lugar a duda). Las situaciones con presencia de atributos repetidos o dominados no entregan mayor información al modelador en cuanto a los balances de atributos que se espera que el individuo realice, y pueden generar situaciones que el individuo vea como poco realistas, por lo que la dominancia debe evitarse a toda costa.

Dado que garantizar la ortogonalidad en todo el diseño y la etapa de estimación no es trivial, y más cuando se encuentra la presencia de situaciones de elección dominadas, Burgess y Street (2005) desarrollaron el diseño ortogonal óptimo en diferencias. El principio de este diseño es generar experimentos donde los atributos comunes entre alternativas nunca puedan tomar los mismos niveles durante una situación de elección, pues argumentan que la ortogonalidad que importa cuando se usan modelos de elección discreta es la que viene dada por la diferencia entre los niveles de los atributos de diferentes alternativas.

Los diseños ortogonales óptimos en diferencia suelen ser útiles cuando no se dispone de ninguna información previa sobre los coeficientes de los parámetros de los modelos a estimar. Sin embargo, Rose y Bliemer (2008) reportan que en la práctica estos diseños pueden producir experimentos con gran número de situaciones de elección, o pueden ser difíciles de encontrar si se fijan las situaciones de elección, limitan la aparición de sólo un conjunto de niveles de atributos entre alternativas – aquellos que maximicen la diferencia entre niveles -, y además promueven ciertas formas de comportamiento en las respuestas (por ejemplo, elecciones lexicográficas).

## 2.2 Diseños eficientes

Los modelos de elección discreta comúnmente usados no son modelos lineales, por lo tanto el diseño de experimentos de elección para este tipo de modelos no tiene las mismas características que para un modelo lineal. Manteniendo el mismo concepto de obtener estimaciones en los parámetros con los mínimos errores estándar posibles, la idea es enfocarse en la matriz informacional de un modelo de elección discreta si se tiene en cuenta que los errores estándares asintóticos se pueden obtener de esta matriz (Train, 2003).

$$MAVC(\beta) = - \left[ E \left( \frac{\partial^2 l(\beta)}{\partial \beta^2} \right) \right]^{-1} \quad (3)$$

## Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

A diferencia de los modelos lineales, en los modelos no lineales estimados por máxima verosimilitud la condición de primer orden no es lineal en los parámetros y por lo tanto, la matriz asintótica de varianzas-covarianzas (MAVC) depende de los parámetros que deseamos estimar. Para un modelo MNL con todos los parámetros genéricos, McFadden (1974) demostró que la matriz informacional,  $I(\beta)$ , podía representarse como se muestra en la ecuación 4.

$$I(\beta) = \frac{\partial^2 l(\beta)}{\partial \underline{\beta}^2} = Z(\beta)' \cdot P(\beta) \cdot Z(\beta) \quad (4)$$

$$Z_{ji}(\beta) = X_{jk} - \sum_j X_{jk} \cdot P_j(\beta) \quad (5)$$

Donde:  $X_{jqs}$  corresponde al nivel del atributo k, para la alternativa j;  $P_{js}$  corresponde a la probabilidad calculada para la alternativa j con los parámetros  $\beta$  y los niveles  $X_{jk}$

Para el caso de modelos MNL con parámetros que no sean genéricos, ver Rose et al (2008).

Dado que la MAVC depende de los parámetros que deseamos estimar, estos diseños parten de la base que tenemos información *a priori* que podemos utilizar para mejorar los diseños de experimentos de elección. La información *a priori* que podemos utilizar para minimizar el valor esperado del inverso de la matriz informacional, no sólo corresponde al valor esperado de los parámetros a estimar sino también se pueden utilizar los signos esperados de los parámetros, la forma del modelo a estimar y hasta información sobre las posibles interacciones con otras variables socioeconómicas.

La MAVC juega un papel importante en el diseño de experimentos de elección pues el concepto de eficiencia se asocia a la minimización de la MAVC. Un diseño será más eficiente mientras menor sea la norma de la MAVC, ya que por definición se obtendrán menores errores estándares para un mismo tamaño de muestra, o lo que es lo mismo, para obtener un mismo error estándar en los parámetros un diseño más eficiente requerirá menos observaciones.

Con el fin de comparar la eficiencia de diferentes modelos se han propuesto diferentes métricas relacionadas con la norma de la MAVC y la información *a priori* utilizada para calcular la MAVC.

La medida de eficiencia más usada es la denominada D-error (ecuación 6), el cual se obtiene del determinante de la MAVC normalizada por la potencia de la razón del número K de parámetros a estimar en el modelo. Otra medida de eficiencia que se ha propuesto es el A-error que se obtiene al calcular la traza de la MAVC dividida por el número K de parámetros a estimar (ecuación 7).

$$D - \text{error} = \det \left\{ - \left[ E \left( \frac{\partial^2 l(\beta)}{\partial \underline{\beta}^2} \right) \right]^{-1} \right\}^{1/K} \quad (6)$$

$$A - \text{error} = \frac{\text{tr} \left\{ - \left[ E \left( \frac{\partial^2 l(\beta)}{\partial \underline{\beta}^2} \right) \right]^{-1} \right\}}{K} \quad (7)$$

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

Con el fin de informar con que información *a priori* fue calculada la MAVC se suele adicionar un subíndice a la definición de la medida del error calculada para el diseño.

- El subíndice z (*zero*) significa que no se utilizó información *a priori* acerca de los parámetros a estimar, o lo que es lo mismo, los parámetros se asumieron iguales a cero.
- El subíndice p (*priors*) significa que se utilizó información *a priori* fija acerca de los parámetros a estimar.
- El subíndice b (*bayesian*) significa que se incorporó incertidumbre acerca del valor *a priori* de al menos uno de los parámetros a estimar. El enfoque bayesiano incorpora incertidumbre acerca de los parámetros *a priori* calculando valores aleatorios para los parámetros *a priori* de acuerdo a una distribución entregada por el modelador. El cálculo de la medida de error bayesiano, es el valor esperado de la medida del error sobre las realizaciones aleatorias efectuadas. La ecuación 8 muestra el D-error bayesiano (Bliemer et al., 2008) y la ecuación 9 el A-error bayesiano.

$$D_b - \text{error} = \int_{\beta} \left\{ \det \left\{ - \left[ E \left( \frac{\partial^2 l(\beta)}{\partial \beta^2} \right) \right]^{-1} \right\}^{1/K} \right\} f(\beta, \varphi) d\beta \quad (6)$$

$$A_b - \text{error} = \int_{\beta} \left\{ \frac{\text{tr} \left\{ - \left[ E \left( \frac{\partial^2 l(\beta)}{\partial \beta^2} \right) \right]^{-1} \right\}}{K} \right\} f(\beta, \varphi) d\beta \quad (7)$$

El procedimiento general para generar un diseño eficiente se ilustra en el figura 1 y comprende 3 etapas en términos generales. El proceso preliminar es la primera etapa que se inicia con la definición de los criterios para realizar el diseño. Estos criterios preliminares están relacionados con los supuestos de modelación y las características del diseño a realizar. Los supuestos de modelación que deben definirse para realizar un diseño eficiente comprenden el tipo de modelo a utilizar (por ejemplo, MNL, NL, ML), las funciones de utilidad de cada alternativa, los atributos y niveles a considerar, y los parámetros *a priori* obtenidos de estudios preliminares o de encuestas piloto. Por otra parte, también hay que definir las restricciones a incluir en el diseño para lograr un experimento de elección realista, se debe definir el tipo de diseño sobre el que se va a generar el primer diseño (ortogonal, ortogonal en diferencias, bayesiano, etc), y determinar la cantidad de situaciones de elección dentro del ejercicio y las que se presentarán a cada uno de los encuestados (por lo tanto, en esta etapa es donde se define si se usan bloques o no).

Luego de completar la etapa preliminar se pasa a la segunda etapa, que es donde se genera el diseño. Una vez generado el diseño, se realiza la última etapa, que comprende la evaluación de los diseños. La evaluación consiste en calcular las probabilidades de elección, obtener la matriz asintótica de varianzas y covarianzas, ya sea por métodos numéricos o mediante derivaciones analíticas, y por último el cálculo de la eficiencia del diseño. Inicialmente el primer diseño se guarda como el mejor diseño, hasta que se encuentre un diseño que tenga mejores

### Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

condiciones de eficiencia que el mejor diseño (inicialmente, el primer diseño). Luego de evaluar la eficiencia del modelo, se procede a actualizar el diseño un conjunto de veces de acuerdo a un criterio de parada establecido. Actualizar el diseño en este contexto se refiere a variar los niveles de los atributos.

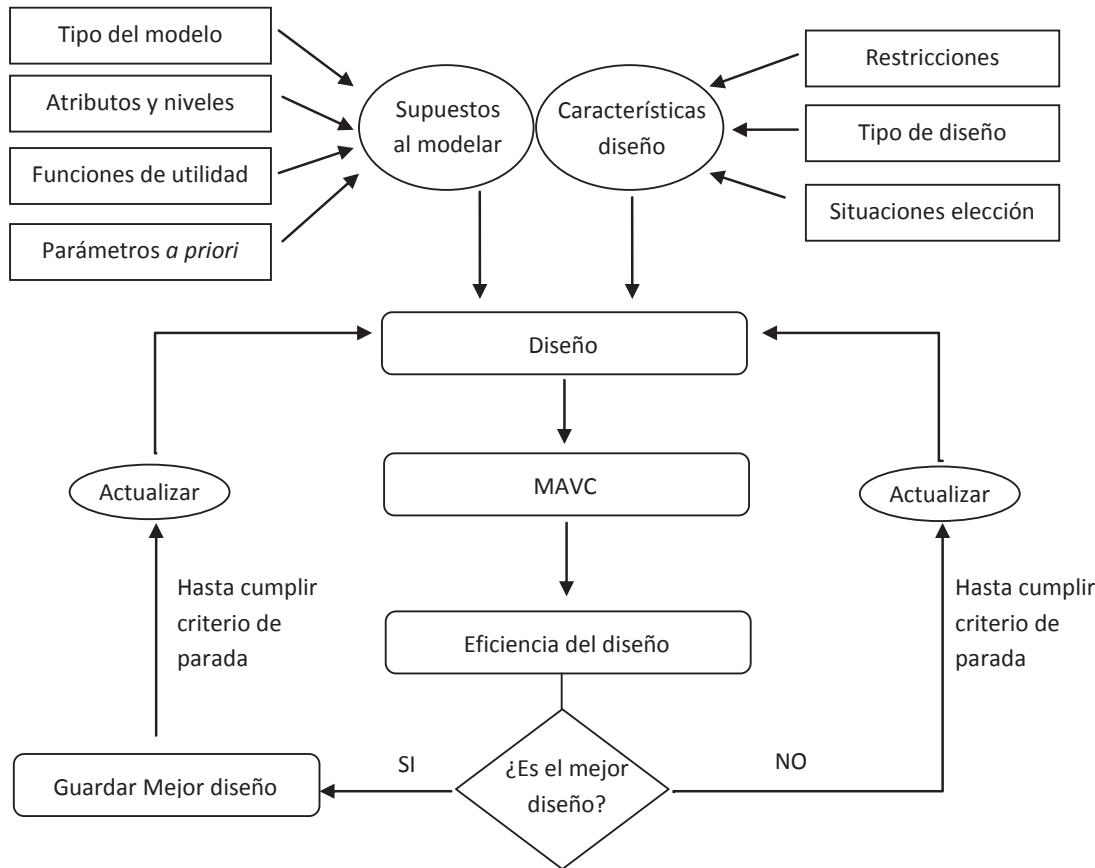


Figura 1. Esquema para realizar un diseño eficiente.

### 3. Modelos de Elección de Horario y sus diseños experimentales

#### 3.1 Modelos de elección del horario de inicio de viajes

Conocer las horas del día que los individuos eligen para realizar los viajes y los atributos que consideran para efectuar dicha elección es crucial para cuantificar detalladamente la demanda por transporte a lo largo del día como se discutió en la parte introductoria de este artículo. El crecimiento de la congestión en las ciudades y la implementación de estrategias y/o políticas para manejar la demanda de transporte ha hecho que el cambio del horario de inicio de los viajes sea una opción popular para evitar transportarse por ciertos lugares en períodos congestionados, o para experimentar costos (monetarios o de tiempos de viaje) menores en

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

períodos diferentes al horario de viaje deseado ante la ausencia de congestión. Los cambios en la hora de viaje son posibles cuando el desplazamiento de la hora de inicio del viaje no acarree inconvenientes en la realización de la actividad en el destino y se pueda finalizar la actividad que se estaba realizando en el origen.

Cuando existe congestión, el incentivo para cambiar de horario de inicio de viajes es transportarse en un periodo de mayor demanda y experimentar menores tiempos de viaje. Cuando se implementan estrategias para manejar la demanda a diferentes horas del día los incentivos para cambiar de horario pueden ser diferentes a los incentivos cuando existe congestión y dependen de la medida implementada. Por ejemplo, cuando existe tarificación vial, los individuos pueden estar incentivados a cambiar de horario de inicio de los viajes con el fin de experimentar menores costos en sus viajes. Las políticas de flexibilización de los horarios de trabajo pueden afectar la estructura temporal de la demanda ya que las personas están dispuestas a trabajar más o menos horas con el fin de evitar viajar en periodos congestionados o experimentar un costo menor al viajar (Bianchi et al., 1998, Ettema et al., 2007). Otro tipo de incentivo es el que se presenta al implementar restricciones de circulación por patente desde cierta hora del día. En Bogotá, algunas personas deciden iniciar sus viajes al trabajo mucho más temprano de lo usual, y salir mucho más tarde de lo usual para utilizar su vehículo y evitar la restricción vehicular.

Los modelos de elección del horario de partida tratan de estimar cómo se distribuyen los viajes a lo largo del día y la investigación en este campo está avanzando. Los modelos de elección de horario contribuyen a entender cómo los viajeros organizan sus viajes y actividades en el tiempo (Hess et al., 2007a). Además permiten vincular las variables socioeconómicas de los individuos junto con la percepción de los niveles de servicio de las redes y los modos de transporte.

El modelo de elección de la hora de inicio de los viajes más conocido y sobre el que se basan la mayoría de estudios realizados en el área es el SM desarrollado por Small (1982). El SM supone que los individuos eligen el periodo de inicio de sus viajes teniendo en cuenta el tiempo de viaje y la utilidad que les produce llegar temprano, tarde o a la hora, respecto a una hora deseada de llegada al destino. La importancia de este trabajo radica en la inclusión de términos relacionados con la programación de las actividades dentro de las funciones de utilidad de un modelo de elección discreta. Específicamente nos referimos al término que indica la cantidad de tiempo que una persona llega tarde o temprano a su destino respecto a una hora de llegada deseada al destino. Este término se conoce en la literatura como la demora horaria (*SD, Schedule Delay*), y es un desarrollo motivado por el trabajo de Vickrey (1969). Diversos estudios han comprobado que el SM de Small (1982), el cual se puede reescribir como las ecuaciones 8 a 12, permite realizar de forma exitosa los *trade-offs* entre el tiempo de viaje y los términos asociados a la demora horaria.

$$V_i = \beta_{TT} TT_i + \beta_{SDE} SDE_i + \beta_{SDL} SDL_i + \delta_L d_L \quad (8)$$

$$SDE_i = \text{Max}\{-SD_i, 0\} \quad (9)$$

$$SDL_i = \text{Max}\{0, SD_i\} \quad (10)$$

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

$$d_L = \begin{cases} 1 & \text{si } SDL_i > 0 \\ 0 & \text{si } SDL_i = 0 \end{cases} \quad (11)$$

$$SD_i = Hr\ Llegada\ observada - Hr\ Llegada\ preferida \quad (12)$$

Donde el subíndice  $i$  se refiere a la alternativa (periodo de hora de inicio del viaje).  $TT_i$  indica el tiempo de viaje del individuo al salir en el periodo  $i$ .  $SD_i$  denota la demora horaria,  $SDE_i$  y  $SDL_i$  las demoras horarias si el individuo llega temprano o tarde, respectivamente.  $d_L$  es una penalidad por llegar tarde al destino. Todos los coeficientes se asumen negativos por la desutilidad que se produce al llegar al destino a un horario diferente al deseado.

En un sistema de transporte por lo general, existe cierta variabilidad en los tiempos de viaje principalmente asociada a la misma operación del sistema o a la congestión presente. La variabilidad en los tiempos de viaje no sólo depende de la relación entre el volumen de tráfico y la capacidad de la infraestructura, sino también de la distribución del tráfico en los diferentes períodos del día (de Jong et al., 2003). Los individuos incorporan la variabilidad en los tiempos de viaje en sus decisiones de viaje basados en la percepción del nivel de congestión y la experiencia de viajes anteriores. Con el fin de incorporar incertidumbre asociada a los sistemas de transporte, Noland y Small (1995) extendieron el SM inicial de Small (1982) y definieron la utilidad esperada al salir a cierta hora del día. La extensión de Noland y Small (1995) se puede reescribir como se muestra en la ecuación 13.

$$E(U_i) = \beta_{TT} E(TT_i) + \beta_{SDE} E(SDE_i) + \beta_{SDL} E(SDL_i) + \delta_L P_L \quad (13)$$

Para estimar modelos que incorporen la incertidumbre de los tiempos de viaje dentro de los experimentos de elección, se debe presentar a los encuestados una noción de variabilidad de los tiempos de viaje. La forma de presentar la variabilidad en los tiempos de viaje de forma eficiente ha sido un tema de discusión y se puede encontrar una buena discusión en Hensher et al (2011). Cuando se incorpora incertidumbre,  $E(TT_i)$  indica el valor esperado de los tiempos de viaje presentados al individuo.  $E(SDE_i)$  y  $E(SDL_i)$  los valores esperados de las demoras horarias si el individuo llega temprano o tarde, respectivamente.  $P_L$  es la probabilidad asociada al número de veces que el individuo llegaría tarde según los tiempos de viaje presentados.

Los modelos de elección de horario de inicio de viajes no sólo deben incorporar la información de los horarios de los viajes, los niveles de servicio y la incertidumbre asociada a ellos, sino también la duración de las actividades (Ettema et al., 2004). El tiempo de participación en las actividades que realiza una persona durante un día, así como la información descriptiva sobre los orígenes-destinos también es relevante pues condiciona la realización de viajes, el orden de participación en actividades, y la elección de las horas de inicio de los viajes. Para estimar buenos modelos idealmente debe conocerse la secuencia de actividades a realizar, los lugares dónde se realizaran, y la flexibilidad de horarios con que se cuenta para realizar cada actividad. Para incorporar la duración de las actividades y eventualmente su secuencialidad, el enfoque de modelación no sólo debe considerar los viajes por sí solos. La realización de otras actividades impone restricciones sobre la elección del viaje, por lo que es ideal considerar *tours* o sucesión de varios viajes que tengan en cuenta la relación entre las diferentes actividades durante el día.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

De Jong et al (2003) y Hess et al (2007b), con la idea de incluir el efecto de la duración de las actividades en el destino, presentan ejemplos de SM que incluyen penalidades explícitas relacionadas con una menor o mayor duración de las actividades. La función de utilidad genérica al incluir estas penalidades por la duración de las actividades ( $PTD_i$  y  $PTI_i$  en las ecuaciones 14 a 16) se muestra en la ecuación 14.

$$V_i = \beta_{TT} TT_i + \beta_C costo_i + \beta_{SDE} SDE_i + \beta_{SDL} SDL_i + \beta_{PTD} PTD_i + \beta_{PTI} PTI_i \quad (14)$$

$$PTD_i = \text{Max}\{-PT_i, 0\} \quad (15)$$

$$PTI_i = \text{Max}\{0, PT_i\} \quad (16)$$

$$PT_i = \text{Duración actividad observada} - \text{Duración actividad preferida} \quad (17)$$

Otro aspecto que no se debe dejar de lado al modelar la elección de la hora de inicio de los viajes son las variables socioeconómicas del individuo, la información sobre las actividades a realizar en los destinos, y la información de otras elecciones que puedan interactuar con la elección de hora (por ejemplo, elección de ruta y de modo).

La información socioeconómica es importante para la programación de las actividades de un individuo debido a las restricciones inherentes impuestas por las responsabilidades propias de cada persona dentro una familia. Pertenecer a una familia influencia las decisiones individuales. Por ejemplo, si un padre debe llevar, y recoger a sus hijos en la escuela debe considerarse dentro de la programación de actividades. En este caso, transportar a los hijos es un determinante de la duración y de la elección de la hora de partida de los viajes del padre sea cual sea el propósito. Para estimar modelos de elección de horario debería recolectarse información socioeconómica como la edad, el sexo, el ingreso familiar, el ingreso individual, el estatus marital, la ocupación, el nivel de escolaridad, el número de niños en el hogar, las horas de trabajo, y la flexibilidad en el horario.

En la discusión que se realizó en la introducción se mostró que la elección de horario se relaciona con otras elecciones que determinan la distribución temporal y espacial de la demanda por transporte (es decir, decisiones de destino, modo, y ruta). Dada la anterior relación, parece buena idea estimar la elección de horario de los viajes simultáneamente con otras elecciones de transporte. En este contexto, la mayoría de estudios (Hess et al., 2007b, de Jong et al., 2003, Hendrickson and Plank, 1984, Day et al., 2010, Bhat, 1998) ha estudiado la relación entre las elecciones de modo y hora de inicio de los viajes, y ha encontrando evidencias para modelar conjuntamente estas elecciones. Adicionalmente, Heydecker y Addison (2005) demuestran que también existe alta correlación entre la elección de ruta y el horario mediante la formulación de un modelo de equilibrio de tráfico cuando se implementan políticas de tarificación.

Modelar la elección de hora de inicio de los viajes es complejo y existe gran cantidad de factores que influyen en esta elección, por lo que resultaría buena idea estimar modelos con diferentes tipos de datos que permitan incorporar diferentes dimensiones que pueden influenciar la elección de horario de inicio de viajes. Los estudios académicos relacionados con la elección de horario durante los últimos años, utilizan datos PD para estimar SM y sería

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

interesante evaluar la posibilidad de incluir variables sobre actitudes de las personas que tal vez pueden mejorar la modelación del comportamiento de las personas.

A pesar de que el enfoque predominante para modelar la elección de la hora de inicio de los viajes son las encuestas PD, no se tiene consenso en el diseño a utilizar para recolectar los datos del SM de forma adecuada. En las siguientes secciones se describen los principales diseños experimentales que se han utilizado en estudios que usan SM para modelar la elección de horario de inicio de viajes y se identifican los problemas no abordados por estos diseños.

### 3.2 Diseños experimentales en estudios anteriores de elección de horario

No existe consenso respecto al diseño experimental más apropiado para generar encuestas PD basadas en el SM. En estudios previos, las situaciones de elección se han generado con diferentes tipos de diseño.

Como se describió en la sección anterior, Small et al (1995, 1999) formularon el SM que ha servido de inspiración para muchos estudios en el área en el marco de un proyecto para evaluar el valor del tiempo de viaje bajo condiciones de congestión. En este proyecto se desarrollaron dos diseños: el primero, para evaluar el *trade-off* entre 4 atributos (tiempo de viaje, variabilidad, hora de inicio de los viajes, y costo); y el segundo, enfocado en evaluar el costo frente a tiempos de viaje sujetos a congestión y en flujo libre (3 atributos). Todos los atributos considerados tenían 3 niveles y la muestra fue segmentada basándose en los tiempos de viaje que experimentaban los encuestados con el fin de darle mayor realismo al experimento. Utilizaron 9 segmentos, y a cada uno de ellos le asignaron diferentes tiempos medios de viaje. Por último, para evaluar el desempeño del modelo en predicción, utilizaron un amplio rango de valores de los coeficientes en experimentos de simulación (coeficientes entre 1 a 5 veces más, en valor absoluto, que los menores valores asumidos para los coeficientes).

Los 7 pasos del procedimiento de diseño utilizado por Small et al (1999) se describen a continuación:

1. Desarrollar los procedimientos para realizar la simulación;
2. Determinar el rango de los coeficientes para la simulación;
3. Seleccionar el diseño PD;
4. Obtener los valores para las variables (costo, tiempo de viaje, etc)
5. Experimento de simulación para cada segmento de tiempo de viaje;
6. Revisar los diseños y los valores;
7. Si es necesario, volver al paso 3.

El último paso no se realiza una vez que el diseño encontrado recupera los coeficientes dentro de un rango de coeficientes preestablecido. La selección del diseño PD dentro del diseño factorial completo ( $3^4$  en el primer ejercicio, y  $3^3$  en el segundo) a evaluar se realizó utilizando bloques basados en criterios para evitar la dominancia de las situaciones de elección, considerando que a cada encuestado se le presentarían 6 situaciones de elección entre dos alternativas por experimento (es decir, 12 en total por encuestado).

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

A pesar que los estudios de Small et al (1995, 1999) obtuvieron resultados más que aceptables y la formulación de su modelo es por lejos la más utilizada, su procedimiento de diseño para obtener los datos no ha sido utilizado con la misma popularidad. Sólo algunos estudios (Asensio and Matas, 2008) se han basado en el procedimiento de diseño descrito por Small et al (1999). Tal vez la popularidad que han adquirido otras técnicas de diseño de experimentos y el hecho de sólo considerar criterios de dominancia para seleccionar los diseños a evaluar en la simulación ha influido en que este procedimiento de diseño no sea el más popular.

Por ejemplo, algunos estudios (Börjesson, 2008) ocupan un diseño ortogonal en diferencia para obtener sus situaciones de elección, y otros (Bajwa et al., 2008, Saleh and Farrell, 2005, Tseng and Verhoef, 2008) un diseño factorial fraccional. Hess et al. (2007b) utilizaron diferentes bases de datos obtenidas con diseños experimentales diferentes. El experimento de elección realizado en Londres, descrito como APRIL dentro de su artículo, utilizó un diseño ortogonal; mientras que el diseño experimental de la base de datos holandesa con la que se trabaja en este artículo utilizó un diseño adaptado manualmente basado en lo que se conoce como diseño Bradley.

El diseño utilizado en este experimento holandés, también utilizado en de Jong et al (2003) y Ettema et al (2007), resulta interesante pues considera la elección de horario en los *toures* que incluyen los viajes de ida y de salida del trabajo. Considerar los *toures* que realizan los encuestados permite incluir la duración de la actividad dentro de los modelos de elección de horario de inicio de los viajes, como se mostró en la sección anterior.

Este diseño contempló la realización de dos experimentos de elección por encuestado, con o sin tarificación vial, cada uno con 8 situaciones de elección (es decir, 16 situaciones de elección por individuo). Dentro de cada situación de elección existían 6 alternativas: 3 alternativas viajando en el mismo modo de transporte establecido por la persona pero con horas de viaje diferente, 1 alternativa de viajar usando un modo de transporte alternativo pero a la misma hora de viaje reportada, 1 opción de no viajar, y 1 opción de no preferencia. Una de las tres primeras alternativas era realizar el viaje a una hora muy similar a la hora reportada por la persona, el cual se utilizó como la alternativa de referencia para adaptar los ejercicios a la realidad de cada encuestado.

Para generar las situaciones de elección, se incluyeron 29 variables dentro del diseño que tenían entre 3 y 7 niveles. Dada la cantidad de atributos y los niveles considerados, era muy complicado adoptar un diseño factorial fraccional u ortogonal específico para este caso. Un diseño de este estilo no estaba disponible en las tablas publicadas por Kocur et al (1982) o en los software comúnmente utilizados en la época, por lo que se optó por generar un diseño ortogonal fraccional 8<sup>3</sup>. Este diseño contó entonces con 512 situaciones de elección en total y 29 variables por situación. Dado que los atributos no tenían 8 niveles en todo caso, lo que se hizo fue colapsarlos para que tuvieran esta cantidad de niveles cada uno, lo cual generó que algunos niveles tuvieran mayor probabilidad de aparecer en el ejercicio que otros. Por último, para escoger las situaciones de elección para cada individuo se dividieron las 512 situaciones entre las 8 situaciones a asignar por individuo, y se asignó un número aleatorio entre 1 y 64 para la primera situación, entre 65 y 128 para la segunda, y así sucesivamente para las 8

### Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

situaciones de elección. Finalmente, este procedimiento de diseño se personalizó para cada individuo mediante el software WinMint.

Los procedimientos de diseño anteriormente mencionados, no utilizan criterios de eficiencia para seleccionar los diseños a presentar a los encuestados por lo que en teoría, deberían requerir un mayor número de muestra para obtener parámetros significativos. Además, salvo la parte de simulación incluida dentro del procedimiento de diseño de Small et al (1999), estos procedimientos de diseño no tienen en cuenta la información previa que puede tenerse de los parámetros.

Con el fin de incluir algunos criterios de eficiencia en la selección del diseño a utilizar y abordar el problema de correspondencia entre los atributos de diseño y los atributos modelables dentro de un SM (este tema se profundizará en la próxima sección), Koster y Tseng (2010) presentaron un procedimiento de diseño de 8 pasos que se describe a continuación:

1. Escoger los niveles de los atributos basados en observaciones realistas del tiempo de viaje (distribuciones);
2. Definir las restricciones de los niveles de los atributos y los coeficientes *a priori*;
3. Definir el número de bloques; el número de situaciones de elección y los criterios de parada del procedimiento;
4. Definir la máxima diferencia dentro de la distribución de tiempos que se muestran al encuestado;
5. Generar aleatoriamente un diseño PD candidato;
6. Calcular los datos del modelo dado los atributos de diseño, los coeficientes *a priori*, y los puntos de referencia de los viajeros;
7. Calcular la correlación de las diferencias y los otros criterios de diseño para evaluar el diseño;
8. Repetir desde el paso 5 hasta que el criterio de parada se haya alcanzado.

El criterio de parada establecido por Koster y Tseng (2010) fue de 30 minutos sin encontrar un mejor diseño. Es decir, si en 30 minutos el procedimiento no encuentra un mejor diseño, entonces el procedimiento se detiene.

### 3.3 Problemas en diseño experimental de modelos de elección de horario

Los diseños experimentales mencionados en la sección anterior tienden a simplificar la complejidad de la tarea de elección y utilizan los experimentos de simulación para aumentar el realismo de las situaciones de elección. Dentro de los problemas principales que dificultan el desarrollo de experimentos de elección para SM se encuentran (a) la dependencia de algunos niveles de los atributos con los niveles de otros atributos del diseño, (b) los atributos de diseño no son exactamente los mismos que se muestran a los encuestados, o algunos atributos de los modelos son obtenidos a partir de los atributos que se muestran a los encuestados (Koster and Tseng, 2010), (c) para garantizar el realismo del experimento, el diseño debe ser personalizado y basado en las características específicas de los viajes de cada encuestado.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

La dependencia entre atributos del diseño a la que nos referimos en este artículo no sólo es la que se presenta cuando los niveles de un atributo en la alternativa j pueden generarse a partir del nivel del mismo atributo en la alternativa i. Esta dependencia ha sido estudiada anteriormente por Rose et al (2008), y producen diseños que se conocen como pivoteados a partir de alternativas de referencia. La dependencia que nos interesa reportar en este artículo es la que se presenta cuando un atributo dentro de la alternativa j depende del nivel de otro atributo de la misma alternativa j, que a su vez es también parte del diseño.

En un ejercicio de elección de horario de inicio de viajes, las variables que se presentan a los encuestados generalmente son el tiempo de viaje, el costo del viaje, la hora de inicio del viaje, y eventualmente la variabilidad del tiempo de viaje (representada algunas veces por una serie de tiempos de viaje alternativos, horas de llegada diferentes, o diferencia de horas de llegada/tiempos de viajes posibles). La práctica común para diseñar un ejercicio de este tipo indica que estos atributos se obtienen al asignar niveles que varían porcentualmente o en valor absoluto respecto a un valor base establecido (en algunos casos el valor de referencia es reportado por el individuo, y en otros casos es un promedio luego de realizar una segmentación de la muestra).

El tiempo de viaje, la variabilidad del tiempo de viaje y el costo son función de la hora en la que se realiza el viaje, que a su vez es otro atributo que varía dentro del diseño. Si se asignan diferentes niveles de variación a ciertos atributos y se ignora la interdependencia entre ellos, pueden generarse situaciones irreales o incluso absurdas para los encuestados.

Suponiendo el caso extremo de generar un experimento con dos atributos, tiempo de viaje y hora de inicio de los viajes, en tres niveles. Los niveles del tiempo de viaje son: 20 minutos – 30 minutos – 40 minutos y 7:00am – 8:00am – 10:00am para las horas de inicio de los viajes. Se sabe que el periodo punta está entre las 7:45am y las 8:45am. Al generar el diseño, pueden presentarse situaciones de elección con alternativas que muestren que viajar a las 8:00am tardaría 20 minutos, mientras que viajar a las 7:00am demoraría 40 minutos. Lo anterior sería contra intuitivo y podría generar un problema si el encuestado no toma en serio la encuesta que está realizando.

Mediante la inclusión de restricciones al momento de generar el diseño, y un procedimiento de diseño en varias etapas como el que se propone en este artículo, la interdependencia entre los atributos de un experimento puede controlarse. Controlar la interdependencia es de suma importancia para el realismo y el compromiso del participante dentro de la encuesta. Debe impedirse a toda costa que el individuo encuentre situaciones de elección contra intuitivas o simplemente poco probables en la realidad.

Por otra parte, formular modelos de elección con atributos diferentes a los que se presentan a los encuestados, caso común en un SM, es un problema desde el punto de vista de la generación del diseño. El problema consiste en que la relación entre los atributos presentados y los atributos modelados no es uno a uno, entonces se presenta un problema de equivalencia, que en el contexto de un diseño eficiente es muy importante definir para que el modelador tenga claro cuál atributo debe intentar optimizar y cómo realizar la equivalencia entre atributos. Este problema ha sido indicado anteriormente por Koster y Tseng (2010) pero es un

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

problema que no ha sido tratado en profundidad dentro de la literatura de diseño de encuestas PD.

En el SM más sencillo los atributos mínimos que se presentan al encuestado son el tiempo de viaje, el costo, y la hora en la que se realiza el viaje (notar que no se incluye variabilidad del tiempo de viaje, ni la duración de las actividades). Al evaluar la ecuación 8 puede verse que la hora en que se realiza el viaje no se estima dentro del SM sino que se usa para derivar la demora horaria definida en las ecuaciones 9 a la 12. Es decir, en el SM lo que interesa es estimar la desutilidad que produce llegar tarde o temprano respecto a una hora deseada de referencia, equivalente a la utilidad marginal de la demora horaria por llegar antes o después de la hora deseada de referencia, y no específicamente el atributo presentado al encuestado que es la hora de inicio del viaje.

Si se incluye la variabilidad del tiempo de viaje y se realiza el experimento en un contexto de *tour*, donde se incluyan varios viajes dentro de la elección, el problema empeora porque se incrementan los atributos en el SM que no tienen equivalencia uno a uno con los atributos presentados al encuestado. En este caso habrá que estimar valores esperados del tiempo de viaje, de la demora horaria, probabilidad de llegar tarde, penalidad asociada a la mayor o menor duración de las actividades o la demora horaria respecto a otro viaje presentado.

En el contexto de un diseño eficiente, la información *a priori* con la que se puede contar es la relacionada con los términos que se estiman en el modelo (por ejemplo, la demora horaria por llegar temprano o tarde) por lo que puede inferirse que estos son los términos indicados a optimizar. Intuitivamente, el diseño debería realizarse controlando los atributos del modelo y luego utilizar las ecuaciones descritas para determinar los niveles de los atributos a presentar al encuestado. En un SM sencillo, sin variabilidad del tiempo de viaje y sólo dentro del contexto de un viaje, tal vez resulte directo optimizar los niveles de los atributos del modelo porque las ecuaciones que vinculan los niveles de los atributos son sencillas; sin embargo al incluir incertidumbre en el SM, los niveles de variación del atributo del modelo muchas veces no se encuentran en la misma escala (unidades) que los niveles de los atributos presentados a la persona. Por esta razón, en modelos complejos hay que ser cuidadosos en cuanto al nivel del atributo a optimizar, pues pueden generarse variaciones no deseadas en los niveles de los atributos presentados al encuestado (a veces muy grandes, o dentro del rango de indiferencia para las personas).

Una alternativa para tratar este problema de equivalencia entre atributos y evitar variaciones no deseadas en los niveles de los atributos presentados a los encuestados, es la que se propone en este artículo. La idea es optimizar el diseño inicialmente controlando los niveles de los atributos a presentar al encuestado en lugar de definir los niveles de los atributos dentro del modelo. Lo que se convierte en este caso son los parámetros *a priori* a considerar para los atributos a presentar, que se obtienen a partir de los parámetros *a priori* de los atributos del modelo teniendo en cuenta las ecuaciones que permiten relacionar los atributos.

## Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

Por último, con el fin de aumentar el realismo de la encuesta y hacer que el individuo recuerde el contexto de la situación de elección que enfrenta usualmente, es deseable presentar una alternativa igual o muy similar a la elección reportada por el individuo. Dotar al encuestado de un contexto general dentro del ejercicio, le permite al individuo incluir restricciones no tenidas en cuenta por el modelador y que pueden influir en el comportamiento real del encuestado. En conclusión, incluir una alternativa de referencia puede otorgarle mayor relevancia al ejercicio. Además, en la literatura se ha encontrado que las encuestas PD que se generan a partir de una respuesta PR previa, suelen presentar mayor eficiencia en el diseño (Rose et al., 2008, Caussade et al., 2005, Train and Wilson, 2009).

### 4. Diseño del experimento PD propuesto

El procedimiento de diseño por etapas que se describe en esta sección permite obtener diseños eficientes para modelos que cuenten con atributos interdependientes, y presenten problemas de equivalencia entre los atributos del modelo y aquellos presentados al individuo. Además, permite incluir restricciones a los atributos con el fin de lograr mayor realismo y evitar la dominancia entre las alternativas de las situaciones de elección presentadas a los encuestados. Es importante notar que una condición necesaria para realizar estos diseños es contar con la información de referencia de cada encuestado previamente a la realización de este experimento.

El procedimiento de diseño contempla 6 etapas que se describen a continuación:

i. Definición de las características preliminares del diseño

Esta etapa contempla todas las actividades previas al desarrollo de un diseño eficiente, como:

- Definir el contexto del experimento y los atributos a presentar;
- Identificar las restricciones y la dependencia entre los niveles de los atributos;
- Definir el atributo que se optimiza primero. En el caso de un SM, se propone optimizar primero los cambios a presentar en el atributo *Hora de inicio/llegada del viaje*;
- Definir los coeficientes *a priori* para todos los atributos dentro del diseño;
- Definir el número de situaciones de elección;
- Definir el número de bloques.

ii. Proceso de optimización para obtener el diseño PD genérico.

Dado que las personas enfrentan diferentes situaciones de elección, el enfoque que se propone en esta etapa consiste en generar un diseño genérico inicial donde los niveles de los atributos en su mayoría representen cambios relativos (porcentuales), no absolutos, respecto a un punto de referencia.

El diseño genérico es común para todos los encuestados si se desea, o se pueden crear un diseño diferente si se cuenta con diversos segmentos predefinidos e identificados dentro de la muestra.

### Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

En esta etapa se optimizan los niveles de los atributos que condicionan los niveles de otros atributos y los niveles de aquellos atributos que no tengan relaciones de dependencia.

Las actividades a desarrollar dentro de esta etapa son:

- Definir el criterio de eficiencia con el que se pretenden evaluar los diseños;
- Definir los criterios de parada del procedimiento;
- Seleccionar un diseño PD candidato aleatoriamente o mediante heurísticas. En esta actividad deben incluirse las restricciones para evitar la dominancia entre alternativas;
- Calcular las probabilidades y la MAVC dados los atributos de diseño y los coeficientes *a priori*;
- Calcular los criterios de eficiencia para evaluar el diseño;
- Volver a escoger otro diseño PD candidato hasta que se cumpla el criterio de parada.

iii. Personalización de las situaciones de elección.

Pasar de los niveles de cambio relativo por atributo que se encuentran en el diseño genérico de la etapa anterior, a los niveles de los atributos personalizados de acuerdo a la situación de elección de cada encuestado es el objetivo de la etapa que se describe a continuación.

La finalidad de esta etapa es obtener mayor realismo en las situaciones de elección y mayor eficiencia en el diseño total de la encuesta, realizando las siguientes actividades:

- Adaptar las situaciones de elección usando la alternativa de referencia entregada. En esta etapa se deben usar los niveles de los atributos del mejor diseño genérico obtenido en la etapa anterior, y convertirlos en niveles personalizados para cada encuestado.
- Definir diferentes niveles de los atributos no optimizados basados en observaciones realistas (por ejemplo, mediciones del tiempo de viaje, costo observado, etc) o respecto a los valores reportados por el individuo;
- Incluir condiciones de dependencia entre los atributos (Restricciones que tienen en cuenta el nivel del atributo condicionante escogido para cada situación de elección en el mejor diseño genérico de la etapa 2);
- Incluir otras restricciones si son necesarias (por ejemplo, umbrales de diferencia entre los atributos de diferentes alternativas, restricciones de variación de los niveles dentro del diseño, máxima diferencia dentro de los puntos extremos que definen la distribución de tiempos de viaje que se muestran al encuestado cuando se decide presentar la variabilidad del tiempo de viaje mediante tiempos de viaje diferente, etc)

iv. Proceso de optimización para obtener el diseño PD final.

Este proceso de optimización es muy similar al realizado en la segunda etapa, sin embargo presenta dos diferencias fundamentalmente.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

La primera diferencia se relaciona con los atributos que varían para optimizar el diseño. En esta etapa los niveles de los atributos que se optimizan son diferentes a los que se modifican en la etapa 2 de este procedimiento.

La segunda diferencia se relaciona con la matriz de varianza y covarianza que se utiliza. En esta parte se deriva y optimiza una matriz de varianza y covarianza a partir del total de datos en la muestra, teniendo en cuenta los atributos del modelo a estimar obtenidos a partir de los atributos personalizados a presentar al encuestado. Es decir, aquí no se está optimizando un diseño común para todos los encuestados, se está optimizando un diseño que contiene tantas filas como situaciones de elección por encuestado multiplicada por el número de participantes.

Las actividades dentro de esta etapa son:

- Definir el criterio de eficiencia con el que se pretenden evaluar los diseños;
  - Definir los criterios de parada del procedimiento;
  - Seleccionar un diseño PD candidato aleatoriamente o mediante heurísticas. En esta actividad deben incluirse las restricciones para evitar la dominancia entre alternativas;
  - Calcular las probabilidades y la MAVC dados los atributos de diseño y los coeficientes *a priori*;
  - Calcular los criterios de eficiencia para evaluar el diseño;
  - Volver a escoger otro diseño PD candidato hasta que se cumpla el criterio de parada.
- v. Experimento de simulación
- La finalidad de esta etapa de simulación es testear si el mejor diseño obtenido en la etapa anterior puede recuperar un amplio rango de coeficientes. Cabe destacar que la simulación debe realizarse con la muestra total que participará en el diseño, variando solamente los coeficientes a recuperar y simulando las elecciones del total de encuestados.
- vi. Volver a 2 si no se recuperan los parámetros en un amplio rango de valores para los coeficientes.

## 5. Aplicación empírica del diseño propuesto

Como parte de un proyecto académico para estudiar la elección de la hora de inicio de los viajes en la ciudad de Santiago de Chile se realizó una encuesta dentro de la que se implementó el diseño descrito en este artículo. Dentro de los objetivos principales del proyecto académico se encontraba el estudio de la elección del modo y del horario de inicio de los viajes basado en la situación actual de la ciudad y el análisis del comportamiento de viaje de los individuos ante la hipotética implementación de estrategias y políticas para manejar la demanda de transporte debido al incremento en la congestión en la ciudad. Específicamente, las estrategias que se consideraron dentro del estudio fueron un esquema de flexibilización de

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

horarios en los centros de empleo y la implementación de la tarificación vial en la ciudad durante horas de la mañana y de la tarde.

La encuesta utilizada para recopilar los datos para este proyecto fue realizada en tres etapas y aplicada a cerca de 500 empleados en diferentes centros de empleo de Santiago. La primera etapa, fue una entrevista asistida por computador (CAPI – *Computer assisted personal interview*) en el lugar de trabajo de la persona y recolectó información socioeconómica, del empleo de la persona, de los factores que inciden en la elección de horario, y de la programación de actividades para el día en el que se realizaría la segunda parte de la encuesta. La segunda etapa consistió en un diario de viajes utilizando un marco recordatorio de actividades (Ampt and Ortúzar, 2004) que fue recopilada vía web. Por último, la tercera etapa de la encuesta también fue recolectada mediante una entrevista asistida por computador en los lugares de trabajo de las personas, corresponde a experimentos PD enfocados hacia los viajes relacionados con el trabajo y donde se utilizó el procedimiento de diseño descrito en este artículo.

2 experimentos PD fueron presentados a cada participante.

- Experimento enfocado sólo en los viajes hacia el trabajo;
- Experimento enfocado en los viajes basados en el trabajo. Es decir, considera el viaje hacia el trabajo y el viaje después del trabajo

Cabe destacar que se diseñaron experimentos de elección diferentes para usuarios de transporte privado y de transporte público. Estos ejercicios de elección fueron diseñados para 360 de los 500 participantes. No se realizó experimentos PD a personas que realizaran algún viaje basado en el hogar o en el trabajo caminando o en bicicleta; ni tampoco se tuvo en cuenta para esta etapa a las personas que reportaran viajes basados en el trabajo cambiando de transporte público a transporte privado (o viceversa) en los diferentes viajes.

Las encuestas PD fueron programadas en MS Excel, y se realizaron en los lugares de trabajo de los participantes con la ayuda de un encuestador que presentaba y contextualizaba el experimento, tomaba las respuestas a cada situación de elección y anotaba datos relevantes respecto a la reacción del encuestado frente al experimento (por ejemplo, atributos no considerados, comentarios, etc.)

La presentación de ambos experimentos contenía 5 columnas. En la primera columna se entregó información acerca de los atributos y en las cuatro restantes se presentaron 4 opciones de elección por viaje (Ver figura 2 y 3). Entre la segunda y la cuarta columna, las opciones correspondieron a viajar en el mismo modo reportado pero a diferente hora: a la hora reportada por el encuestado (alternativa de referencia), antes, y después que la hora reportada. La quinta columna siempre correspondió a viajar en un modo de transporte diferente llegando al destino a una hora muy similar a la reportada. El modo de transporte alternativo para los usuarios de transporte privado fue el transporte público, mientras que para el caso de los usuarios de transporte público fue el transporte privado, si lo tenían disponible, o un servicio de transporte público tipo taxi colectivo en caso de no contar con la disponibilidad de viajar en auto. Un aspecto importante a mencionar es que la posición de las

### Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

alternativas de elección de hora dentro de la pantalla fue fijada aleatoriamente para evitar hábito en las elecciones de las personas.

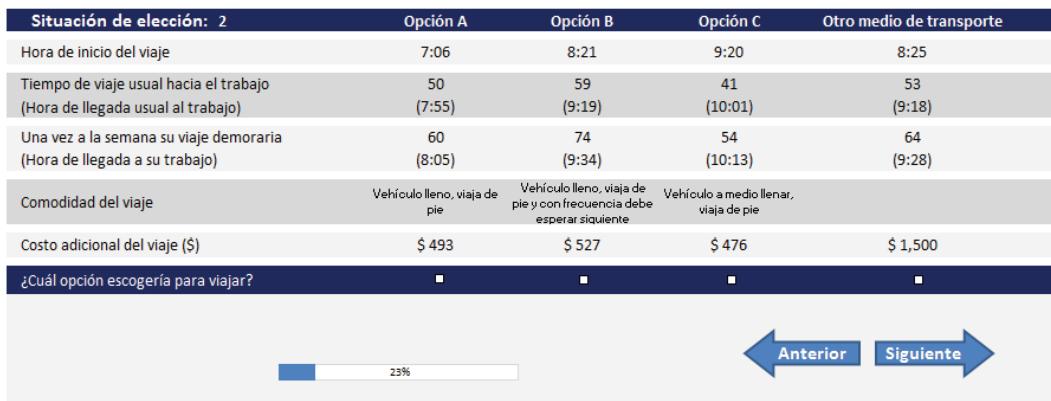


Figura 2. Ejemplo de la presentación del experimento de elección enfocado en viajes hacia el trabajo para usuarios de transporte público

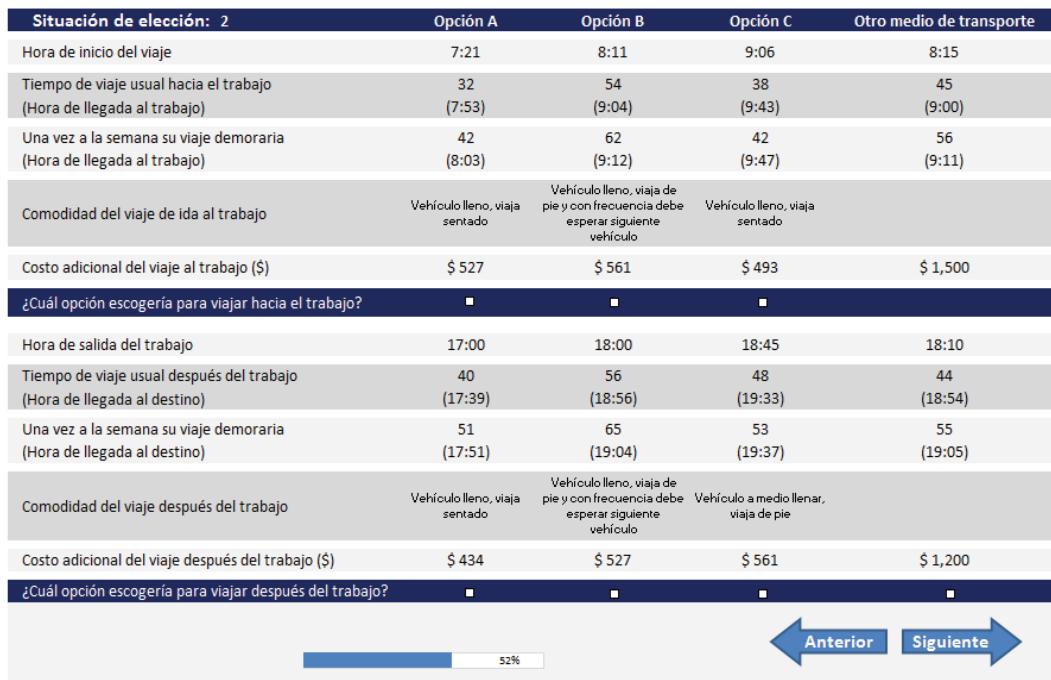


Figura 3. Ejemplo de la presentación del experimento de elección enfocado en viajes basados en el trabajo para usuarios de transporte público

Dado que el primer experimento se enfocaba sólo en los viajes hacia el trabajo, las 4 opciones de elección representaban las 4 alternativas disponibles para el individuo.

En el caso del segundo experimento, dado que se presentaban los dos viajes principales basados en el trabajo y el objetivo principal del experimento es estudiar la elección de hora, las alternativas de elección del encuestado se fijaron en 10. 9 de las alternativas correspondieron

## Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

a viajar en el mismo modo pero a horas diferentes y si se escogía la alternativa de cambio de modo, debía hacerse en ambos viajes.

El diseño experimental para generar las situaciones de elección de los experimentos mencionados se realizó utilizando el procedimiento descrito en la sección anterior. Por simplicidad de presentación y limitación de espacio, en este artículo describiremos principalmente el diseño del ejercicio más simple (es decir, el experimento enfocado en los viajes al trabajo).

### i. Características preliminares del diseño

Se decidió realizar un diseño eficiente de los experimentos de elección basado en el SM y utilizando el procedimiento descrito en la sección anterior. Para la segunda etapa se utilizó el software de diseño NGENE ([www.choice-metrics.com](http://www.choice-metrics.com)) y la tercera etapa se programó usando macros en Visual Basic para Excel porque los software que conocemos no permiten incluir restricciones de dependencia entre atributos de la misma alternativa.

#### - Contexto del experimento y definición de atributos a presentar

Antes de presentar las situaciones de elección a los encuestados, se decidió entregar una descripción del experimento con el fin de crear un contexto común entre todos los encuestados. Al principio de la encuesta PD a cada encuestado se le presentaba la siguiente descripción de la situación:

*Debido al crecimiento de la población y del parque automotor en la Región Metropolitana, la congestión en el Metro, los buses, y la producida por los autos que circulan en la ciudad es cada vez mayor. De hecho, se pronostica que dentro de 10 años la duración promedio de los viajes será dos o tres veces mayor que la duración actual.*

*Por estos motivos, estamos realizando una investigación acerca de las preferencias de los usuarios del sistema de transporte santiaguino, con miras a evaluar diferentes alternativas de gestión que permitan mitigar los efectos de la congestión esperada en la ciudad a futuro.*

*En esta encuesta, le presentaremos una serie de situaciones de elección (basadas en los viajes reportados por Ud. anteriormente), en que Ud. debe elegir la opción que le parezca más atractiva en cada caso, teniendo en cuenta que el gobierno ha decidido implementar un sistema de cobro a los automovilistas que circulen por el centro de la ciudad en horarios de congestión.*

Además se les decía a las personas que tenían la opción de modificar sus horas de entrada y de salida de trabajo siempre y cuando cumplieran con un número de horas de trabajo fijas y preestablecidas a la semana.

Los atributos a presentar dentro del experimento del viaje hacia el trabajo fueron: hora de inicio del viaje, tiempo de viaje, variabilidad del tiempo de viaje, comodidad del tiempo de viaje, costo.

## Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

Es importante resaltar que se decidió presentar la variabilidad del tiempo de viaje como un mayor tiempo de viaje posible en un viaje durante un día a la semana, en lugar de otras presentaciones comunes (ver Hensher et al, 2011) porque consideramos que nuestra prioridad era el compromiso entre atributos para elegir la hora de viaje y el modo. En grupos focales, encontramos que las personas efectivamente pueden manejar los formatos de presentaciones comunes que se han utilizado en la literatura para considerar la variabilidad del tiempo de viaje (por ejemplo, 5 tiempos de viaje alternativos). Sin embargo, cuando en el experimento se presentaban atributos con este tipo de presentación las personas tienden a fijarse mayormente en ese atributo dentro del ejercicio. En el contexto de un estudio acerca de la valoración de la incertidumbre del tiempo de viaje resultaría adecuado incluir los 5 tiempos de viaje, pero en el contexto de un ejercicio para la elección de horario pensamos que es mejor considerar la variabilidad usando un atributo más simple que no distraiga el foco del experimento.

- Restricciones y relaciones de dependencia entre atributos

En el diseño de este experimento se consideró que los atributos costo y tiempo de viaje dependen de la hora del día en que se realicen los viajes. Debido a que decidimos mantener lo más simple la presentación de este atributo, en el contexto de este experimento se decidió calcular los niveles de este atributo como un porcentaje promedio de la variación de las mediciones de los tiempos de viaje realizadas sobre la red de transporte de Santiago.

Adicionalmente, haciendo uso del software NGENE se incluyeron restricciones para impedir la dominancia entre las alternativas del experimento durante la segunda etapa del experimento.

Las restricciones incluidas en la tercera etapa se describirán más adelante.

- Atributo que se optimiza primero

En el caso de este experimento se propone optimizar primero en el diseño genérico a los cambios del atributo Hora de inicio del viaje junto con la comodidad y la variabilidad del tiempo de viaje.

- Niveles y coeficientes *a priori* de los atributos

Para generar el diseño genérico de cambios relativos respecto a un punto de referencia, se decidió utilizar 5 niveles por atributo.

Los valores de los coeficientes *a priori* del costo, el tiempo de viaje y la comodidad se obtuvieron de estudios previos realizados en Chile. Los valores de los coeficientes *a priori* para las demoras horarias fueron obtenidas a partir de los valores de la confiabilidad del tiempo de viaje ( $\beta_{SDE}/\beta_{TT}$  y  $\beta_{SDL}/\beta_{TT}$ ) reportados en estudios internacionales. Dado que no se encontraron estimaciones en Chile para los coeficientes de demoras horarias, se decidió incluir incertidumbre en los valores de todos los parámetros *a priori* (enfoque bayesiano).

Los niveles de los atributos y el rango de los valores *a priori* utilizados se muestran en la tabla 1.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

Tabla 1. Niveles de los atributos y valores de los parámetros a priori dentro del diseño genérico

Atributo	Nivel					Parámetro a priori	
	1	2	3	4	5	Min	Max
Cambio del tiempo de viaje*	1	1.05	1.1	1.15	1.2	-7	-0.7
Cambio hora de inicio del viaje: alternativa actual y cambio de modo	-10	-5	0	5	10	-	-
Cambio hora de inicio del viaje: alternativa viaja antes	-30	-45	-60	-75	-90	0.01	0.24
Cambio hora de inicio del viaje: alternativa viaja después	30	45	60	75	90	-0.36	-0.015
Variabilidad del tiempo de viaje transporte privado	0.1	0.15	0.2	0.25	0.3	-9	-0.3
Variabilidad del tiempo de viaje transporte público	0.15	0.2	0.25	0.3	0.35	-9	-0.3
Cambio del costo*	0.7	0.85	1	1.15	1.3	-0.3	-0.017
Comodidad	0.7	0.85	1	1.1	1.2	-1.5	-0.85

Dado que la mayoría de los valores de los coeficientes *a priori* reportados en la tabla 1 se refieren a cambios en los atributos y no a los atributos en sí, los valores entregados por la literatura se “adaptaron” multiplicando los valores de los parámetros *a priori* por valores medios de referencia dentro de la muestra.

- Número de bloques y de situaciones de elección

Luego de realizar una prueba piloto, tener en cuenta la duración del experimento, y comparar el número de situaciones de elección en otros estudios similares, se definió que los encuestados enfrentarán en total 13 situaciones de elección. 5 situaciones de elección del primer experimento, y 8 del segundo.

Utilizando el software NGENE, se generó el diseño genérico común para el primer experimento a partir de 50 situaciones de elección y 10 bloques. Las situaciones de elección del segundo experimento, se obtuvieron a partir de un diseño genérico común con 40 situaciones de elección y 5 bloques.

## ii. Proceso de optimización para obtener el diseño PD genérico

Para el proceso de optimización se utilizó el software NGENE. Se incluyeron las restricciones de dominancia entre alternativas, por las características del diseño se escogió el  $D_b$ -error como el

---

\* Notar que los atributos tiempo de viaje y costo serán fijos entre alternativas para el diseño genérico pues no se optimizan en la segunda etapa por su relación de dependencia respecto a la hora de inicio de los viajes. Los niveles que se presentan en estos casos, son variaciones del nivel base que se incorporan con el fin de variar los escenarios entre situaciones de elección.

criterio de comparación, y se fijó el criterio de parada en 1 hora sin encontrar un mejor diseño al correr el programa en un computador Core 2 Duo de 1.86MHz. Luego del proceso de optimización, los  $D_b$ -error medios de los mejores diseños genéricos son los que se muestran en la tabla 2.

Tabla 2.  $D_b$ -error medios de los mejores diseños genéricos para cada experimento

Experimento	D <sub>b</sub> -error por tipo de usuario	
	Transporte privado	Transporte público
1 – Viajes al trabajo	1.31	1.12
2 – Viajes basados en el trabajo	1.70	1.42

### iii. Personalización de las situaciones de elección

Esta etapa se realizó en Microsoft Excel usando los diseños genéricos de la etapa anterior y la información entregada por el encuestado en las dos primeras partes de la encuesta de este proyecto.

El primer paso de este proceso consistió en asignar los niveles de cada atributo para cada situación de elección de cada persona a encuestar. Cada bloque se asignó aleatoriamente y se buscó repetir el uso de cada bloque un número similar de veces en la muestra total.

- Niveles de los atributos no optimizados en el diseño genérico y restricciones de dependencia

El cálculo del tiempo de viaje para las alternativas que no implican cambios considerables de modo ni hora, se basó en el valor reportado por los encuestados en las dos primeras partes de la encuesta. El tiempo de viaje para la alternativa cambio de modo se basó en el valor obtenido al dividir la distancia que recorrería el encuestado si usará el modo alternativo y el promedio de la velocidad medida en la red de transporte durante la recolección de datos, para la hora en que inicia el viaje según la alternativa personalizada. El cálculo del tiempo de viaje para las alternativas con cambios considerables de hora fue optimizado dentro del diseño y los valores se pivotearon a partir del tiempo de viaje de la alternativa donde se usa el mismo modo y se viaja a una hora muy similar a la reportada.

Para calcular los tiempos de viaje base y los niveles de variación de los tiempos de viaje, se introdujo una restricción respecto a la hora de inicio del viaje. La restricción utilizada para las alternativas de viaje donde se llega al lugar de destino a una hora muy cercana a la hora reportada, es diferente a las otras 2 alternativas donde el individuo llega a su destino en una hora diferente a la reportada (antes o después).

Para las alternativas cambio de modo y salir a la hora base reportada, se tuvo en cuenta que al aplicar el nivel de variación de la hora del inicio del viaje, el individuo podía realizar su viaje en un periodo del día diferente al reportado (notar que el periodo puede cambiar porque a la hora del día reportada se le adiciona el nivel del atributo cambio en la hora del día obtenido

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

del diseño genérico para cada alternativa). Los períodos del día considerados para calcular el tiempo de viaje de usuarios de auto eligiendo estas alternativas fueron: horario punta (entre 7:30-9:00hrs), fuera de punta (antes de las 6:30hrs, y después de las 10hrs), y periodo valle (entre 6:30-7:30hrs, y 9:00-10:00hrs). El tiempo de viaje base entonces era afectado por un valor igual a 1 si la persona no cambia de periodo del día, menor a 1 si la alternativa implica que la persona viajaría en un periodo menos congestionado, y mayor a uno si se le ofrecía viajar en un periodo de mayor congestión al reportado. Finalmente, el valor que aparecería en la encuesta corresponde al tiempo base afectado por el factor de la hora del día y multiplicado por el nivel del escenario cambio de tiempo de viaje del diseño genérico.

Para el cálculo del tiempo de viaje de las alternativas que implican cambios de hora considerables, se definieron 5 niveles de variación que dependían de la hora del día en la cual se realizan los viajes. La restricción de hora evalúa en qué hora del día se propone realizar el viaje, y toma los 5 niveles de variación del tiempo de viaje que corresponden a viajar a esa hora del día.

Adicionalmente, se consideraron diferentes niveles de variación del tiempo de viaje para estas alternativas de acuerdo a la duración del viaje realizado por el encuestado. Por esta razón, viajes con duración menor a 50 minutos, tuvieron diferentes niveles de variación del tiempo de viaje respecto a los viajes de mayor duración. No es lo mismo proponer un cambio de 50% en un viaje de 10 minutos que un cambio de 50% en un viaje de 120 minutos. Un resumen de los 5 niveles de variación de los tiempos de viaje se muestra en la tabla 3.

Para calcular el tiempo de viaje de estas alternativas, se debía seleccionar un nivel en el procedimiento de optimización del diseño y luego multiplicar este nivel por el tiempo de viaje de la alternativa base (es decir, la alternativa de viajar en el mismo modo y a una hora similar a la reportada).

Tabla 3. Niveles de variación del tiempo de viaje utilizados en la segunda parte del diseño

Alt	Niveles	Duración < 50 minutos					Duración > 50 minutos				
		1	2	3	4	5	1	2	3	4	5
Antes	TT1	0.8	0.85	0.9	0.95	1	0.9	0.92	0.94	0.96	0.98
	TT2	0.7	0.75	0.8	0.85	0.9	0.8	0.85	0.9	0.95	1
	TT3	0.55	0.6	0.65	0.7	0.75	0.7	0.8	0.85	0.9	0.95
	TT4	1.02	1.04	1.06	1.08	1.1	1.02	1.04	1.06	1.08	1.1
	TT5	1	1.05	1.1	1.15	1.2	1	1.05	1.1	1.15	1.2
Después	TT1	0.8	0.85	0.9	0.95	1	0.9	0.92	0.94	0.96	0.98
	TT2	0.7	0.75	0.8	0.85	0.9	0.8	0.85	0.9	0.95	1
	TT3	0.55	0.6	0.65	0.7	0.75	0.7	0.8	0.85	0.9	0.95
	TT4	1.05	1.1	1.15	1.2	1.25	1.02	1.04	1.06	1.08	1.1
	TT5	1.25	1.3	1.35	1.4	1.45	1	1.05	1.1	1.15	1.2

El cálculo del costo por viaje se realizó a partir del valor personalizado del atributo hora de inicio del viaje para cada alternativa. Dentro del ejercicio, los encuestadores indicaron a los

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

participantes que si eran usuarios de transporte privado el costo indicado se refería al valor adicional a pagar por concepto de tarificación vial. Por otra parte, si el participante era usuario de transporte público, correspondería a la tarifa completa de viaje.

Para el caso del atributo costo no se realizaron variaciones en los niveles por simplicidad y para evitar situaciones poco realistas. De acuerdo a un esquema de cobro por períodos del día, como el que actualmente rige en la ciudad para el sistema de transporte público, sería poco creíble encontrar situaciones de elección donde la persona tenga dos alternativas de viaje dentro de un mismo período del día pero con costo de viaje diferente.

Este atributo pudo haber sido optimizado aplicando un procedimiento similar al del tiempo de viaje pero adicionando una restricción de diferencia de costos para la selección del nivel de variación del atributo. La optimización de este atributo no se realizó porque se consideró que se iba a obtener variación suficiente del atributo costo al adicionar la variación fija entre alternativas que se incluyó en el diseño genérico y la dependencia respecto a la hora de inicio de los viajes.

El procedimiento que se siguió para asignar los costos fue multiplicar la variación fija entre alternativas incluida dentro del diseño genérico por el costo correspondiente al período del día en que viajaría la persona según lo propuesto en la alternativa y lo indicado en la tabla 4.

Tabla 4. Valores del costo por período del día utilizados en la segunda parte del diseño del experimento 1

Hora	-6:30	6:30-	7:00-	7:30-	8:00-	8:30-	9:00-	9:30-	10:00-	10:30-
	7:00	7:30	8:00	8:30	9:00	9:30	10:00	10:30		
T. Privado	500	800	1000	1200	1500	1500	1200	1000	800	500
T. Público	510	560	620	660	660	620	580	560	540	510

- Restricciones para aumentar el realismo del diseño

Luego de realizar la personalización inicial y en conjunto con la definición de los niveles de los atributos a utilizar en el segundo proceso de optimización se definieron tres restricciones adicionales para aumentar el realismo del diseño.

En el experimento 1, la primera restricción incluida se refiere al atributo de la hora de inicio de los viajes para la alternativa que implica salir antes de la hora reportada hacia el trabajo. Se consideró que no sería atractivo para una persona que reportaba que iniciaba su viaje antes de las 7:30 hrs, presentar una alternativa de viaje que implique tener que salir hacia su trabajo antes de las 6:00 hrs. Por esta razón, a estas personas se les incluyó una corrección en la hora de inicio del viaje (ecuación 18), de tal manera que los cambios a presentar sólo para esta alternativa fueran menores a los propuestos en la tabla 1.

$$Hr\ inicio_B = Hr\ reportada + Cambio\ de\ hr_B \cdot \frac{(Hr\ inicio_A - Hr\ referencia)}{90} \quad (18)$$

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

Donde el subíndice B se refiere a la alternativa de salir antes, el subíndice A a la alternativa de salir a una hora similar a la actual, cambio de hr se refiere al nivel del atributo dentro del diseño genérico y hr referencia es la hora mínima fijada para que el individuo inicie su viaje. En este proyecto, la hora de referencia utilizada fue las 6:10 am.

Las dos restricciones restantes son umbrales de diferencia permitidos entre atributos de diferentes alternativas (Ortúzar and Rodríguez, 2002). Estas restricciones son aplicadas sobre el tiempo de viaje y las horas de inicio de los viajes entre diferentes alternativas de una misma situación de elección. Diferencias de tiempos u horas entre alternativas menores a 5 minutos tienden a ser consideradas irrelevantes dentro de una comparación. Por esta razón, cuando se encontró una diferencia menor a este umbral respecto al atributo de la alternativa de referencia, el atributo de la alternativa no referencial fue colapsado a tener una diferencia de al menos 5 minutos en valor absoluto.

#### iv. Proceso de optimización para obtener el diseño PD final

El proceso de optimización para la obtención del diseño final fue programado en Visual Basic para Excel y en líneas generales es muy similar al realizado en la segunda etapa, con las diferencias indicadas en la etapa descriptiva del procedimiento. El criterio de comparación entre diseños continuó siendo el el  $D_b$ -error medio, utilizando 150 *halton draws*. Los parámetros a priori utilizados se muestran en la tabla 5 y el criterio de parada fue fijado en 30 minutos si es que no se encuentra un mejor diseño.

Tabla 5. Valores de los parámetros *a priori* utilizados en la optimización final

Atributo	Tiempo de viaje	ESDE	ESDL	Costo	Comodidad
Max	-0.012	-0.0072	-0.0144	-0.00017	-0.00038
Min	-0.12	-0.24	-0.36	-0.003	-0.00666

Luego del proceso de optimización, los  $D_b$ -error medios de los mejores diseños genéricos son los que se muestran en la tabla 6.

Tabla 6.  $D_b$ -error medios de los mejores diseños para cada experimento

Experimento	$D_b$ -error por tipo de usuario	
	Transporte privado	Transporte público
1 – Viajes al trabajo	5.72E-05	9.55E-06
2 – Viajes basados en el trabajo	9.20E-06	4.91E-06

Notar que los  $D_b$ -error medios obtenidos en esta etapa no son comparables con los de la etapa intermedia porque en este último caso se está optimizando la matriz de varianza covarianza total de la muestra. Los  $D_b$ -error medios para los diseños genéricos deberían estar escalados por el número de encuestados N que utiliza el mismo diseño.

## Draft version

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

### v. Experimento de simulación

Luego de obtener los diseños, estos fueron sometidos a un procedimiento de simulación donde se testeó si era posible recuperar el rango de parámetros *a priori* que se utilizaron para optimizar los diseños generados para el proyecto. En el 100% de los casos, las muestras generadas recuperaron los parámetros de diseño.

## 6. Conclusiones y Discusión

Este artículo presenta el diseño de una encuesta en desarrollo en el marco de un proyecto académico que se realiza en Santiago. Se ha presentado la metodología para generar un diseño eficiente de un experimento PD basado en un SM, y para el momento de la conferencia se espera tener resultados de las estimaciones con la muestra total recolectada.

La metodología de diseño presentada es novedosa y contribuye al estado del arte y la práctica de generación de diseños de encuestas de preferencias declaradas. La principal novedad de la metodología es permitir obtener experimentos PD realistas que incluyan restricciones de dependencia entre atributos y sea capaz de generar diseños donde los atributos que se usan en la estimación del modelo son función de aquellos atributos presentados directamente a los individuos.

Es importante mencionar que una de las principales bondades de la metodología de diseño es la eficiencia y el realismo que se puede lograr, sin embargo presenta la limitación de requerir una encuesta de preferencias reveladas que se haya realizado previamente al estudio para poder realizar la personalización de las situaciones de elección. Lo anterior podría ser una limitación importante para su aplicación en términos de tiempo y costo de la encuesta. Además, debe reconocerse que como la optimización del diseño se realiza utilizando la muestra completa, se espera que exista una pérdida de eficiencia en el modelo a estimar asociada a que la encuesta tal vez no sea contestada por todas las personas con las que se realizó el diseño.

Por último, este proyecto lo consideramos importante porque es la primera experiencia en Chile conocida por los autores, para modelar la hora de inicio de los viajes usando un SM estimado a partir de datos PD y PR. Bianchi et al (1998), estudiaron previamente la elección de horario en el metro de Santiago pero utilizaron otro tipo de modelo basado en una encuesta de jerarquización.

## 7. Referencias

- AMPT, E. & ORTÚZAR, J. D. D. 2004. On Best Practice in Continuous Large-scale Mobility Surveys. *Transport Reviews*, 24, 337-363.
- ASENSIO, J. & MATAS, A. 2008. Commuters' valuation of travel time variability. *Transportation Research Part E: Logistics and Transportation Review*, 44, 1074-1085.
- BAJWA, S., BEKHOR, S., KUWAHARA, M. & CHUNG, E. 2008. Discrete Choice Modeling of Combined Mode and Departure Time. *Transportmetrica*, 4, 155-177.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

- BHAT, C. R. 1998. Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice. *Transportation Research Part B: Methodological*, 32, 455-466.
- BHAT, C. R. & STEED, J. L. 2002. A continuous-time model of departure time choice for urban shopping trips. *Transportation Research Part B: Methodological*, 36, 207-224.
- BIANCHI, R., JARA-DÍAZ, S. R. & ORTÚZAR, J. D. D. 1998. Modelling new pricing strategies for the Santiago Metro. *Transport Policy*, 5, 223-232.
- BLIEMER, M. C. J., ROSE, J. M. & HESS, S. 2008. Approximation of bayesian efficiency in experimental choice designs. *Journal of Choice Modelling*, 1(1), 98-127.
- BÖRJESSON, M. 2008. Joint RP-SP data in a mixed logit analysis of trip timing decisions. *Transportation Research Part E: Logistics and Transportation Review*, 44, 1025-1038.
- BURGESS, L. & STREET, D. J. 2005. Optimal designs for choice experiments with asymmetric attributes. *Journal of Statistical Planning and Inference*, 134, 288-301.
- CAUSSADE, S., ORTÚZAR, J. D. D., RIZZI, L. I. & HENSHER, D. A. 2005. Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological*, 39, 621-640.
- DAY, N., HABIB, K. N. & MILLER, E. J. 2010. Analysis of work trip timing and mode choice in the Greater Toronto Area. *Canadian Journal of Civil Engineering*, 37, 695-705.
- DE JONG, G., DALY, A., PIETERS, M., VELLAY, C., BRADLEY, M. & HOFMAN, F. 2003. A model for time of day and mode choice using error components logit. *Transportation Research Part E: Logistics and Transportation Review*, 39, 245-268.
- ETTEMA, D., ASHIRU, O. & POLAK, J. 2004. Modeling Timing and Duration of Activities and Trips in Response to Road-Pricing Policies. *Transportation Research Record: Journal of the Transportation Research Board*, 1894, 1-10.
- ETTEMA, D., BASTIN, F., POLAK, J. & ASHIRU, O. 2007. Modelling the joint choice of activity timing and duration. *Transportation Research Part A: Policy and Practice*, 41, 827-841.
- HENDRICKSON, C. & PLANK, E. 1984. The flexibility of departure times for work trips. *Transportation Research Part A: General*, 18, 25-36.
- HENSHER, D. A., GREENE, W. H. & LI, Z. 2011. Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings. *Transportation Research Part B: Methodological*, 45, 954-972.
- HESS, S., DALY, A., ROHR, C. & HYMAN, G. 2007a. On the development of time period and mode choice models for use in large scale modelling forecasting systems. *Transportation Research Part A: Policy and Practice*, 41, 802-826.
- HESS, S., POLAK, J., DALY, A. & HYMAN, G. 2007b. Flexible substitution patterns in models of mode and time of day choice: new evidence from the UK and the Netherlands. *Transportation*, 34, 213-238.
- HEYDECKER, B. G. & ADDISON, J. D. 2005. Analysis of Dynamic Traffic Equilibrium with Departure Time Choice. *TRANSPORTATION SCIENCE*, 39, 39-57.
- KOCUR, G., ADLER, T., HYMAN, W. & AUNET, B. 1982. Guide to Forecasting Travel Demand with Direct Utility Assessment. Washington DC: United States Department of Transportation, Urban Mass Transportation Administration.
- KOSTER, P. & TSENG, Y. Y. 2010. Stated Choice Experimental Designs for Scheduling Models. In: HESS, S. & DALY, A. (eds.) *Choice Modelling: The state-of-the-art and the state-of-practice*. Bingley, UK: Emerald.
- LOUVIERE, J. J., HENSHER, D. A. & SWAIT, J. D. 2000. *Stated Choice Methods: Analysis and Application*, Cambridge, University Press, Cambridge
- MCFADDEN, D. 1974. Conditional logit analysis of qualitative choice behaviour. In: ZAREMBKA, P. (ed.) *Frontiers of Econometrics*. New York: Academic Press.

**Draft version**

Arellana et al (2011). Stated preference design for departure time models including dependency between attribute levels

- NOLAND, R. B. & SMALL, K. A. 1995. Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation Research Record*, 1493, 150-158.
- ORTÚZAR, J. D. D. & RODRÍGUEZ, G. 2002. Valuing reductions in environmental pollution in a residential location context. *Transportation Research Part D: Transport and Environment*, 7, 407-427.
- ORTÚZAR, J. D. D. & WILLUMSEN, L. G. 2011. *Modelling Transport*, Chichester, John Wiley and Sons.
- ROSE, J. M. & BLIEMER, M. C. J. 2008. Stated Preference Experimental Design Strategies. In: HENSHER, D. A. & BUTTON, K. J. (eds.) *Handbook of Transport Modelling*. Second Edition ed. Amsterdam: Elsevier.
- ROSE, J. M., BLIEMER, M. C. J., HENSHER, D. A. & COLLINS, A. T. 2008. Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological*, 42, 395-406.
- SALEH, W. & FARRELL, S. 2005. Implications of congestion charging for departure time choice: Work and non-work schedule flexibility. *Transportation Research Part A: Policy and Practice*, 39, 773-791.
- SMALL, K. 1982. The scheduling of consumer activities: Work Trips. *The American Economic Review*, 72, 467-479.
- SMALL, K., NOLAND, R. & KOSKENOJA, P. 1995. Socio-economic Attributes And Impacts Of Travel Reliability: A Stated Preference Approach. Research Reports, California Partners for Advanced Transit and Highways (PATH), Institute of Transportation Studies, UC Berkeley.
- SMALL, K. A., NOLAND, R., CHU, X. & LEWIS, D. 1999. Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation. *NCHRP Report 431*. Transportation Research Board.
- STREET, D. J. & BURGESS, L. 2004. Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. *Journal of Statistical Planning and Inference*, 118, 185-199.
- TRAIN, K. 2003. *Discrete Choice Methods with Simulation*, Cambridge, University Press, Cambridge.
- TRAIN, K. & WILSON, W. 2009. Monte Carlo analysis of SP-off-RP data. *Journal of Choice Modelling*, 2(1), 101-117.
- TSENG, Y. Y., KOSTER, P., PEER, S., KNOCKAERT, J. & VERHOEF, E. 2011. Discrete choice analysis for trip timing decisions of morning commuters: estimations from joint SP/RP-GPS data. *International Choice Modelling Conference 2011*. Leeds.
- TSENG, Y. Y. & VERHOEF, E. 2008. Value of time by time of day: A stated-preference study. *Transportation Research Part B: Methodological*, 42, 607-618.
- VICKREY, W. 1969. Congestion Theory and Transport Investment. *The American Economic Review*, 59, 251-260.