



# Análisis de clases latentes como herramienta para identificar y caracterizar trayectorias en estudiantes de educación superior

**Claudia Amo Grez**

Estudiante Doctorado en Psicología  
Escuela de Psicología - Facultad de Ciencias Sociales  
Pontificia Universidad Católica de Chile

**María Verónica Santelices Etchegaray**

Facultad de Educación  
Pontificia Universidad Católica de Chile

# Deserción y Persistencia

- **El problema de la deserción en Educación Superior (ES)**
  - Aumento de cobertura ES (*Paredes, 2015*).
  - Mayor deserción en niveles socioeconómicos bajos  
(*González et al. , 2005; Santelices et al., 2013*)
  - Costos económicos y sociales.
- **Sobre los estudios de deserción y persistencia**
  - Variadas operacionalizaciones de deserción según:
    - Niveles de análisis
    - Momentos de observación
    - Tipos de deserción

# El problema

- La categorización que tradicionalmente se utiliza permite representaciones parciales del comportamiento de los estudiantante.
- Diferentes determinantes de la deserción, dependiendo de la operacionalización que se utilice (Stratton, O'Toole, & Wetzel, 2008)

## Problema de Investigación

Desarrollar una aproximación empírica que permita una perspectiva más integradora del tránsito de los alumnos por ES, mediante seguimiento en varios años y considerando distintos niveles de análisis.

# Objetivos

- General

Examinar las trayectorias en ES de quienes ingresan por primera vez a instituciones universitarias.

- Específicos

1. Identificar y caracterizar los patrones de matrícula en el Sistema de Educación Superior que mejor representan las distintas trayectorias de los alumnos.
2. Determinar la proporción de alumnos que desarrolla cada trayectoria.
3. Establecer en qué medida estas trayectorias se relacionan con factores sociodemográficos, académicos, de interés vocacional y de financiamiento.

# Metodología

- Población de interés.

Alumnos que ingresan por primera vez a ES.

- Muestra.

51.522 egresados de EM en 2006, inscritos en PSU para la admisión 2007 que ingresan ese año a 55 universidades chilenas.

- *91,7% de los casos que egresaron de secundaria en 2006 e ingresaron a universidades en 2007: cubriendo el 99,7% de los ingresos al CRUCH y el 81,6% de ingresos a universidades privadas.*

- Fuentes de Información.

Registros del DEMRE (cuestionarios inscripción PSU), SIES, Ayudas Estudiantiles del MINEDUC y Comisión Ingresos.

*\*Información consolidada en proyecto FONIDE N° F611103 (Santelices, Catalán, Horn, & Kruger, 2013)*

Tabla 1  
Características de la Muestra

<i>Género</i>		
Mujer	26.555	52%
Hombres	24.967	48%
<i>Nivel de Ingreso Familiar</i>		
Tramo1(< \$278 mil)	21.085	41%
Tramo 2 (\$278-\$835 mil)	17.612	34%
Tramo 3 (\$834-\$1.400 mil)	5.585	11%
Tramo 4 (> \$1.400 mil)	7.206	14%
Datos perdidos	34	0%
<i>Tipo de Colegio</i>		
Particular Pagado	12.253	24%
Particular Subvencionado	23.993	47%
Municipal	14.904	29%
Datos perdidos	372	1%

# Metodología

- Variables
  - **Variable Latente:** variable catorógica que identifica subpoblaciones de estudiantes con distinto tipo de trayectorias.
  - **Indicadores:** mat2008 , mat2009, mat2010, mat2011, cinst.
  - **Co-variables:** género, nivel de ingreso familiar, nivel educacional de los padres, NEM y PSU-Mat (estandarizadas), tipo de colegio, estudiar en otra región, interés vocacional (lugar de postulación a la carrera de ingreso), proporción monto beneficio/arancel (CAE, FSCU, BECAS).

- **Técnica:** Análisis de Clases Latentes (Collins & Lanza, 2010)

$$P(\mathbf{Y} = \mathbf{y}) = \sum_{c=1}^C \gamma_c \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho^{I(y_j=r_j)}(j, r_j | c)$$

Análisis de Clases Latentes con Covariables:

$$P(\mathbf{Y} = \mathbf{y} | X = x) = \sum_{c=1}^C \gamma_c(x) \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho^{I(y_j=r_j)}(j, r_j | c)$$

Donde  $\gamma_c(x)$  es un modelo de regresión logística multinomial

$$\gamma_c(x) = P(L = c | X = x) = \frac{e^{\beta_{0c} + \beta_{1c}x}}{1 + \sum_{c'=1}^{C-1} e^{\beta_{0c'} + \beta_{1c'}x}}$$

$$\ln\left(\frac{\gamma_c}{\gamma_C}\right) = \beta_{0c} + \beta_{1c}x$$

# Metodología

- Estrategia de Análisis
  - Comparación de modelos con distinto número de clases
  - Selección del modelo base
    - Indicadores de ajuste
    - Interpretación de la sustantiva de la solución
    - Doble validación cruzada de los modelos
  - Estimación modelos con covariables
    - Modelo A (toda la muestra)
    - Modelo B (solo alumnos con ingreso al CRUCH)

## Syntax Mplus Modelo A

```
Title: modelo A

data: file = IND&COV_03.11.15.csv;

variable: names = G MRUN IDLCA CODINST MAT2008 MAT2009
                 MAT2010 MAT2011 CINST CINT2TP SEX
                 ING1 ING2 ING3 ING4
                 EDUF1 EDUF2 EDUF3 EDUF4 EDUF5
                 COLE1 COLE2 COLE3
                 NEM PSUMAT NEMZ PSUMATZ
                 PREF1 PREF2 PREF3 PREF4 PREF5
                 PREFE REGION REGOTRA
                 PPBECA PPCRED PPCAE
                 ACREDIN TINST2;
missing = all (9999);
usevariables = IDLCA MAT2008 MAT2009
              MAT2010 MAT2011 CINST
              SEX ING2 ING3 ING4
              EDUF2 EDUF3 EDUF4 EDUF5
              COLE1 COLE2 REGOTRA
              NEMZ PSUMATZ
              PPBECA PPCRED PPCAE;
categorical = MAT2008 MAT2009
             MAT2010 MAT2011 CINST;
auxiliary = IDLCA;
classes = c(4); !numero de clases

analysis: type = mixture;
          starts = 1500 100;
          stiterations = 50;
          process = 4 (starts);

model: %overall%           !modelo con convariables
      c ON SEX ING2 ING3 ING4
        EDUF2 EDUF3 EDUF4 EDUF5
        COLE1 COLE2 REGOTRA
        NEMZ PSUMATZ
        PPBECA PPCRED PPCAE;|
```

# RESULTADOS

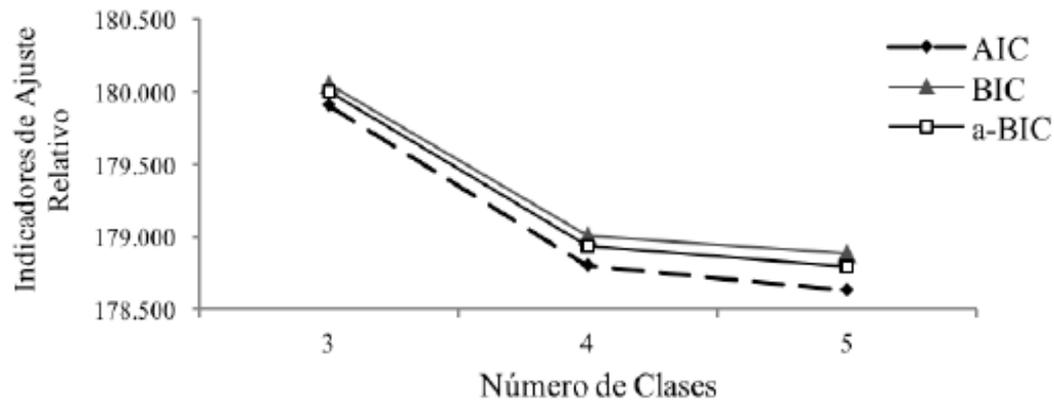
Selección del modelo base

## • Evaluación del Ajuste

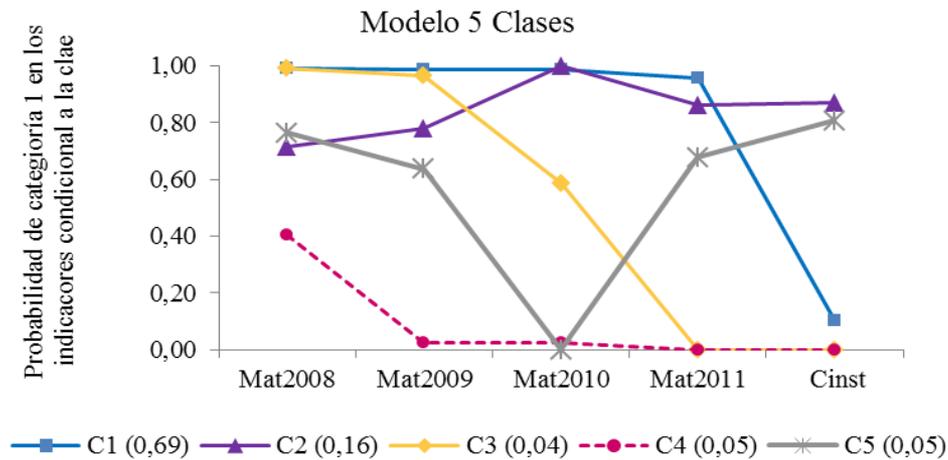
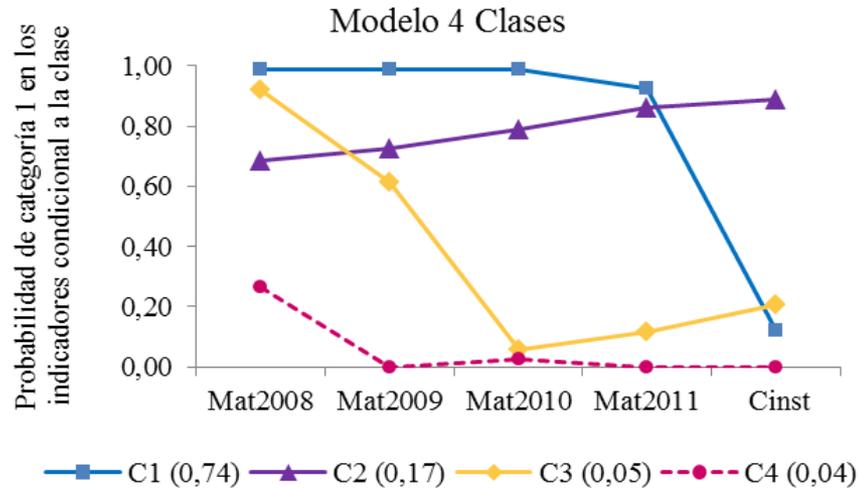
Tabla 2: Comparación de indicadores de ajuste en modelos con distinto número de clases

n° clases	n° P.L	gl	LL	AIC	BIC	a-BIC	$G^2$	$\chi^2$	Entropía
1	5	26	-106.277	212.563	212.608	212.592	34.286 *	294.467 *	1,00
2	11	20	-93.905	187.832	187.929	187.894	9.542 *	11.430 *	0,89
3	17	14	-89.936	179.905	180.056	180.002	1.604 *	1.344 *	0,82
4	23	8	-89.377	178.801	179.004	178.931	487 *	457 *	0,84
5	29	2	-89.286	178.631	178.888	178.795	306 *	258 *	0,85

Nota. N=51522. W=32. \* p < ,05.



- Evaluación Interpretabilidad



- Doble validación cruzada

Tabla 2

*Ajuste relativo AIC en doble validación cruzada*

	Modelo a partir de SM1		Modelo a partir de SM2	
	Calibración SM1	Validación SM2	Calibración SM2	Validación SM1
4c	89.490	89.329	89.334	89.496
5c	89.403	<b>89.266</b>	89.239	<b>89.422</b>

Tabla 3

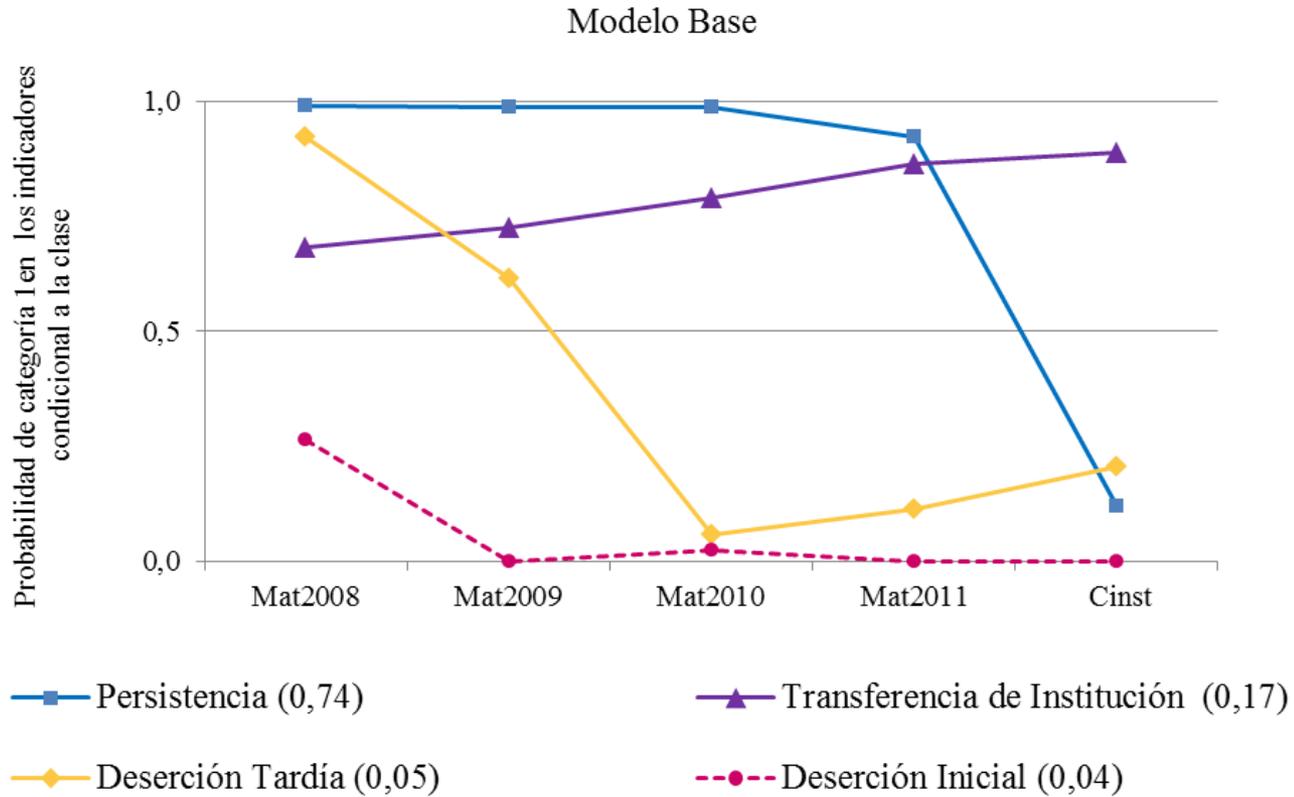
*Resultados calibración comparación de soluciones en SM1 y SM2*

	Modelo 4 Clases				Modelo 5 Clases				
	SM1				SM1				
	C1	C2	C3	C4	C1	C2	C3	C4	C5
Prevalencia	0,73	0,18	0,05	0,04	0,73	0,18	0,05	0,04	0,01
Mat2008	<b>0,99</b>	<b>0,68</b>	<b>0,93</b>	0,24	<b>0,99</b>	<b>0,70</b>	<b>1,00</b>	0,18	0,00
Mat2009	<b>0,99</b>	<b>0,74</b>	<b>0,60</b>	0,00	<b>0,99</b>	<b>0,74</b>	<b>0,56</b>	0,02	<b>1,00</b>
Mat2010	<b>0,99</b>	<b>0,79</b>	0,05	0,02	<b>0,99</b>	<b>0,81</b>	0,06	0,03	0,36
Mat2011	<b>0,92</b>	<b>0,86</b>	0,10	0,00	<b>0,92</b>	<b>0,89</b>	0,08	0,00	0,00
Cinst	0,12	<b>0,86</b>	0,21	0,00	0,11	<b>0,84</b>	0,20	0,00	<b>0,88</b>
	SM2				SM2				
	C1	C2	C3	C4	C1	C2	C3	C4	C5
Prevalencia	0,75	0,16	0,05	0,04	0,71	0,15	0,05	0,05	0,04
Mat2008	<b>0,99</b>	<b>0,69</b>	<b>0,91</b>	0,29	<b>0,99</b>	<b>0,71</b>	<b>0,77</b>	0,41	<b>0,99</b>
Mat2009	<b>0,99</b>	<b>0,72</b>	<b>0,63</b>	0,00	<b>0,99</b>	<b>0,78</b>	<b>0,61</b>	0,03	<b>0,96</b>
Mat2010	<b>0,99</b>	<b>0,79</b>	0,07	0,03	<b>0,98</b>	<b>1,00</b>	0,00	0,03	<b>0,57</b>
Mat2011	<b>0,92</b>	<b>0,87</b>	0,13	0,01	<b>0,95</b>	<b>0,86</b>	<b>0,69</b>	0,00	0,00
Cinst	0,13	<b>0,93</b>	0,21	0,00	0,11	<b>0,93</b>	<b>0,82</b>	0,00	0,00

*Soluciones Similares*

*Soluciones Diferentes*

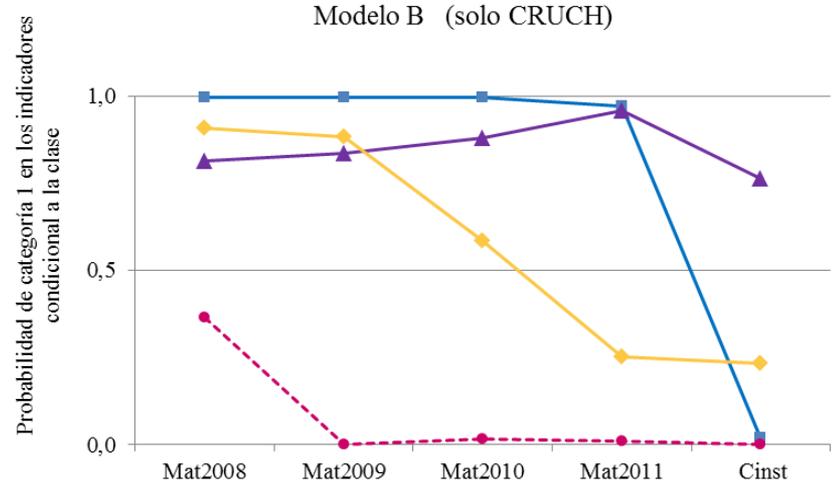
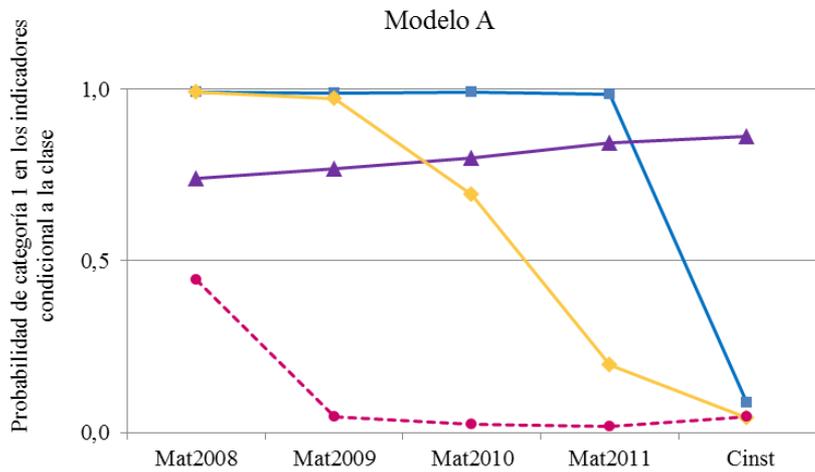
- Se selecciona el modelo con 4 clases como Modelo Base



# RESULTADOS

Modelo de 4 clases con co-variables

- Clases muy similares al modelo base



■ Persistencia (0,65)      ▲ Transferencia de Institución (0,21)  
◆ Deserción Tardía (0,08)      ● Deserción Inicial (0,06)

■ Persistencia (0,68)      ▲ Transferencia de Institución (0,22)  
◆ Deserción Tardía (0,08)      ● Deserción Inicial (0,03)

Tabla 6

Resultados regresión logística multinomial expresados en Odds y Odds Ratio (Modelo A)

Clase de referencia	Persistencia (P)			Transferencia de Institución (TI)			Deserción Tardía (DT)			Deserción Inicial (DI)		
	TI	DT	DI	P	DT	DI	P	TI	DI	P	TI	DT
Constante	3,21 ***	6,23 ***	3,82 ***	0,31 ***	1,94 ***	1,19	0,16 ***	0,51 ***	0,61 **	0,26 ***		1,63 **
Variables												
<b>Proporción "monto beneficio"/"monto arancel" (1 = 100% de cobertura)</b>												
Becas	1,90 ***	0,66 ***	2,73 ***	0,53 ***	0,35 ***	1,44 *	1,51 ***	2,87 ***	4,12 ***	0,37 ***	0,70 *	0,24 ***
Crédito FSCU	1,55 ***		2,26 ***	0,64 ***	0,54 ***	1,46 ***		1,85 ***	2,69 ***	0,44 ***	0,69 ***	0,37 ***
Crédito CAE	5,74 ***	1,60 ***	9,64 ***	0,17 ***	0,28 ***	1,68 *	0,63 ***	3,60 ***	6,04 ***	0,10 ***	0,60 *	0,17 ***
<b>Antecedentes Académicos Previos (puntajes estandarizados)</b>												
Puntaje NEM	1,95 ***	1,59 ***	1,90 ***	0,51 ***	0,82 ***		0,63 ***	1,23 ***	1,19 **	0,53 ***		0,84 **
PSU Matemáticas	1,59 ***	1,83 ***	2,06 ***	0,63 ***	1,15 **	1,30 ***	0,55 ***	0,87 **	1,13 *	0,48 ***	0,77 ***	0,89 *
<b>Ingreso Familiar (ref: &lt; \$278 mil)</b>												
Tramo 2 (\$278-\$835 mil)	1,33 ***	1,46 ***	2,05 ***	0,75 ***		1,54 ***	0,68 ***		1,40 ***	0,49 ***	0,65 ***	0,71 ***
Tramo 3 (\$834-\$1.400 mil)	1,42 ***	1,54 ***	2,09 ***	0,70 ***		1,47 **	0,65 ***		1,35 *	0,48 ***	0,68 **	0,74 *
Tramo 4 (> \$1.400 mil)	1,49 ***		1,74 ***	0,67 ***					1,45 *	0,57 ***		0,69 *
<b>Educación de los padres (ref: sin estudios o básica incompleta)</b>												
Básica com. /Media incom.			1,29 *			1,56 **				0,77 *	0,64 **	
Media com. / CFT incom.			1,91 ***			2,17 ***			1,63 **	0,52 ***	0,46 ***	0,61 **
CFT com. / Univ. incom		1,34 *	2,40 ***		1,61 **	2,87 ***	0,75 *	0,62 **	1,79 **	0,42 ***	0,35 ***	0,56 **
Universitaria comp.		1,34 *	2,44 ***		1,62 **	2,94 ***	0,74 *	0,62 **	1,82 **	0,41 ***	0,34 ***	0,55 **
<b>Región de estudio respecto del hogar (ref: hogar y estudios en misma región)</b>												
Estudia en otra región	0,62 ***	1,18 *		1,62 ***	1,92 ***	1,46 ***	0,85 *	0,52 ***	0,76 **		0,68 ***	1,31 **
<b>Tipo colegio (ref: Municipal)</b>												
Particular Pagado												
Particular Subvencionado			1,34 ***			1,29 ***			1,38 ***	0,75 ***	0,77 ***	0,73 ***
<b>Género(ref: mujer)</b>												
Hombre	0,70 ***		0,72 ***	1,43 ***	1,38 ***		0,72 ***	0,75 ***		1,39 ***		1,34 ***

\*p < ,05. \*\* p < ,01. \*\*\* p < ,001.

Tabla 7

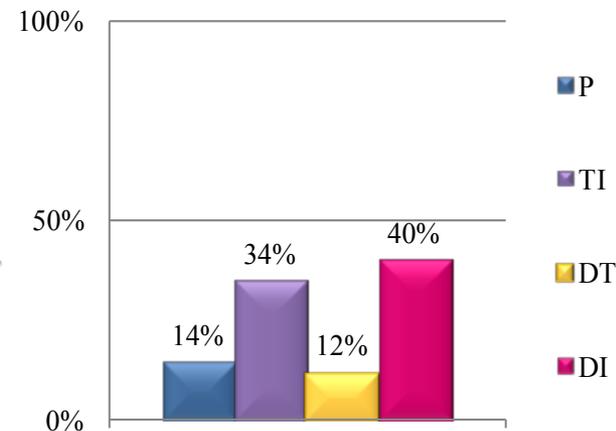
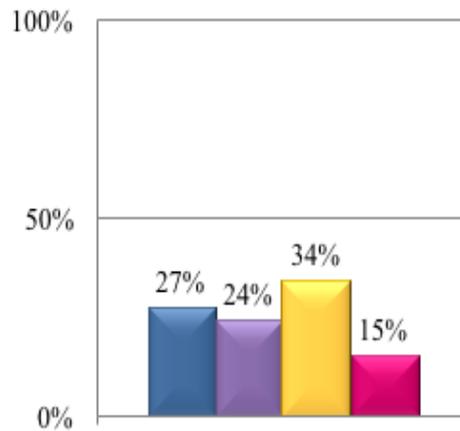
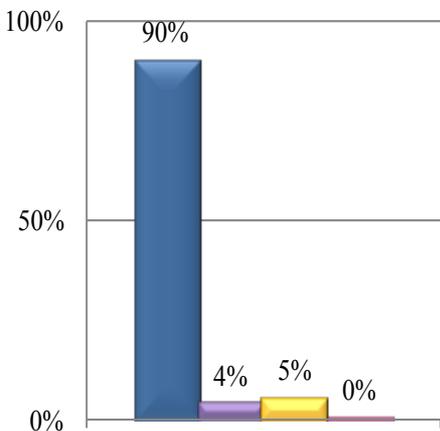
Resultados Factor Interés Vocacional en Regresión Logística Multinomial Modelo B Expresados en Odds Ratio.

Variables	Persistencia (P)			Transferencia de Institución (TI)			Deserción Tardía (DT)			Deserción Inicial (DI)		
	TI	DT	DI	P	DT	DI	P	TI	DI	P	TI	DT
<b>Interés Vocacional (categoría de referencia: ingreso a la primera preferencia en la postulación)</b>												
2a Pref.	0,76 ***			1,32 ***	1,54 ***		0,65 ***	0,70 *				1,42 *
3a Pref.	0,72 ***			1,39 ***	1,51 **	1,41 *	0,66 **				0,71 *	
4a Pref.	0,59 ***			1,71 ***	1,97 ***	1,55 *	0,51 ***				0,64 *	
5a a 8a Pref.	0,61 ***			1,64 ***	1,44 *	2,09 ***	0,69 *				0,48 ***	

Nota: relación interés vocacional con trayectorias, controlando por covariables del Modelo A

\* p < ,05. \*\* p < ,01. \*\*\* p < ,001.

# Casos Hipotéticos Sistema Nacional



*Beneficio Estudiantil:* **CAE (100%)**

*NEM:* **601 pts.**  
*PSU-Mat:* **573 pts.**

*Ingreso Familiar:* **\$278-\$835mil**

*Educación Familiar:* Media Completa  
/CFT incompleto

*Región:* Misma Región  
*Colegio:* Particular  
Subvencionado  
*Sexo:* Mujer

**Beca (100%)**

**450 pts.**  
**426 pts.**

**< \$278mil**

**Sin Educación Formal  
o Básica Incompleta**

Misma Región  
Particular  
Subvencionado  
Mujer

**Sin Beneficio**

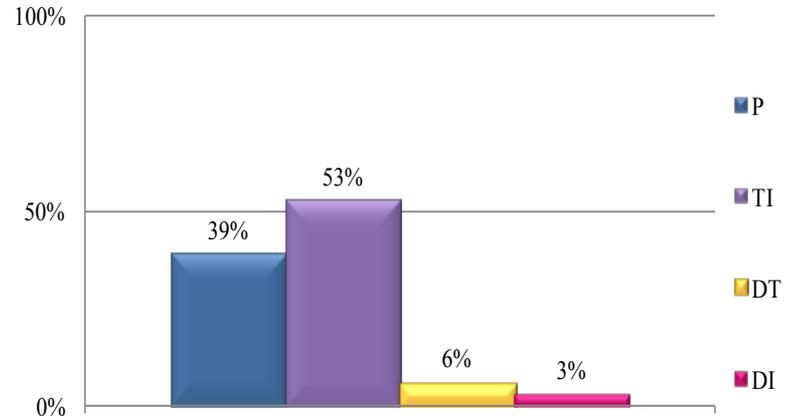
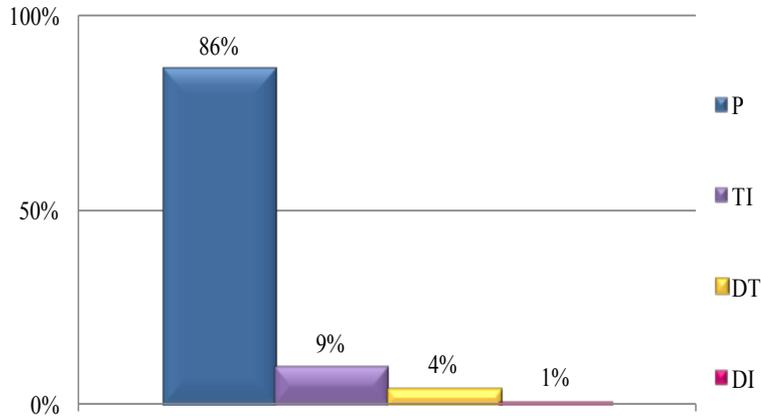
**450**  
**426**

**< \$278mil**

**Sin Educación Formal  
o Básica Incompleta**

**Otra Región**  
**Municipal**  
**Hombre**

# Casos Hipotéticos CRUCH



Beneficio Estudiantil: **CAE (100%)**

NEM: **601**

PSU-Mat: **573**

Ingreso Familiar: **\$278-\$835mil**

Educación Familiar: **Media Completa o CFT incompleto**

Colegio: **Particular Subvencionado**

Región: **Misma Región**

Lugar de Preferencia: **1º preferencia**

Sexo: **Mujer**

**Sin Beneficio**

601

573

\$278-\$835mil

Media Completa /CFT incompleto

Particular Subvencionado

**Otra Región**

**3º preferencia**

**Hombre**

# Discusión y Conclusión

- Novedades en las trayectorias identificadas:
  - 86% de los alumnos se mantienen vinculados al sistema.
  - Transferencia de Institución (vinculo reingreso con deserción institucional)
  - 2 tipos de Deserción (distintos momentos de salida)
- Sobre los factores asociados:
  - Distintos factores se relacionarían con distintas trayectorias (Ej. Transferencia de Institución con interés vocacional y estudiar en otra región)
  - Gran relevancia del financiamiento en trayectorias de Persistencia, Deserción Tardía y Deserción Inicial
    - CAE
    - BECAS
  - Relevancia de la dimensión académica podría ser aún mayor si se analizara el rendimiento de primer año (Allen et al., 2008)

- Ventajas de la propuesta de análisis:
  - Permite relacionar niveles de análisis.
  - Permite integrar varios momentos de observación.
  - No requiere levantar información adicional.
- Limitaciones de representación:
  - No se representa a los alumnos que no se inscriben en PSU (Ej: Admisión especial)
  - Menor representación de egresados de enseñanza técnico profesional y de colegios municipales (González et al., 2014)
- Futuros desarrollos
  - Análisis por tipo de institución.
  - Inclusión del nivel carreras o áreas disciplinares.
  - Análisis por cohortes.

# Referencias

- Allen, J., Robbins, S. B., Casillas, A., & Oh, I. S. (2008). Third-year college retention and transfer: Effects of academic performance, motivation, and social connectedness. *Research in Higher Education*, 49(7), 647–664. <http://doi.org/10.1007/s11162-008-9098-3>
- Collins, L. M., & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. (D. Balding, N. A. C. Cressie, G. M. Fitzmaurice, I. M. Johnstone, G. Molenberghs, D. W. Scott, ... S. Weisberg, Eds.). Nueva Jersey y Canadá: John Wiley & Sons.
- González, L., Uribe, D., & González, S. (2005). Estudio sobre la repitencia y deserción en la educación superior chilena. Santiago de Chile. Recuperado a partir de [https://www.inacap.cl/tportal/portales/tp4964b0e1bk102/uploadImg/File/REPITENCIA\\_DESERCION\\_L\\_E\\_Gonzalez\\_2005.pdf](https://www.inacap.cl/tportal/portales/tp4964b0e1bk102/uploadImg/File/REPITENCIA_DESERCION_L_E_Gonzalez_2005.pdf)
- González, M., Elizalde, L., & Rolando, R. (2014). Transición de a Educación Superior. Santiago de Chile. Recuperado a partir de [http://www.mifuturo.cl/images/Estudios/Estudios\\_SIES\\_DIVESUP/transicion\\_cohorte\\_2006\\_a\\_ed\\_superior\\_2014.pdf](http://www.mifuturo.cl/images/Estudios/Estudios_SIES_DIVESUP/transicion_cohorte_2006_a_ed_superior_2014.pdf)
- Paredes, R. (2015). Desafíos de la experiencia de financiamiento de la educación superior en Chile. En A. Bernasconi (Ed.), *La Educación Superior de Chile. Transformación, desarrollo y crisis*. (Primera Ed, pp. 219–257). Ediciones Universidad Católica de Chile.
- Santelices, V., Catalán, X., Horn, C., & Kruger, D. (2013). Determinantes de Deserción en la Educación Superior Chilena, con Énfasis en Efecto de Becas y Créditos. Santiago de Chile. Recuperado a partir de <http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2062&tm=2>
- Stratton, L. S., O'Toole, D. M., & Wetzel, J. N. (2008). A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*, 27, 319–331. <http://doi.org/10.1016/j.econedurev.2007.04.003>