

Trabajo 3 - Redes Sociales :

Hipótesis y análisis sustantivo de la red

Juan Illanes Vásquez^{*1}

¹CICS, Facultad de Gobierno, Universidad del Desarrollo

1 Hipótesis

En el trabajo 2 se observó una particularidad. En la centralidad de grado de la red de licitaciones, 8 de 10 empresas correspondían al área de la salud (su principal rubro correspondía a licitar a empresas estatales correspondientes al área de la salud). La interpretación de la red consistía en unir las empresas que tuvieran una especialización sobre el promedio en las comunas del país, es decir, licitaran constantemente al estado y no de manera esporádica, luego se construían los enlaces por su "similaridad" es decir, la probabilidad de que, ponderando por todo el territorio nacional, dos empresas p y q coexistieran en el mes de Enero de 2025.

Se puede interpretar entonces que la centralidad de grado mostraría a aquellas empresas que tuvieran muchas empresas relacionadas entre sí, por tanto, existiendo una cantidad de empresas que prestan servicio al estado de manera continua y diversamente, contrariamente a lo que sería una empresa que está inconexa o brinda un servicio hiperespecializado en el área (o de lleno, estaría representando una preferencia patológica en el servicio estatal, como por ejemplo, corrupción).

En la literatura se ha visto que esta centralidad entre proveedores del sistema de salud pública se debe a distintos factores, que van desde una táctica de las empresas a acaparar licitaciones y generar más ganancias a través de su influencia [1], a una integración compleja para obtener suministros críticos [2] o de un equilibrio de mercado que reduce los costos generales de obtención de insumos médicos [3].

Por lo tanto, una hipótesis plausible y directa sería: Para la red de licitaciones, el grado medio de las empresas asociadas al rubro de la salud es mayor al de aquellas que brindan servicio a otras áreas del aparato estatal.

2 Metodología

Es directo en plantear una regresión que tenga como variable independiente una variable dummy que chequee si un nodo pertenece al área de la salud o no. Sin embargo, para poder proceder y someter a juicio dicha regresión, es necesario descartar sesgos asociados a la construcción de la red misma. El procedimiento de construcción de la red consiste en 3 pasos: Obtener la red de volumen X_{cp} , obtener la matriz de especialización binaria (biadyacencia) M_{cp} y calcular las similaridades ϕ_{pq} .

La obtención de X_{cp} está suministrado únicamente por los datos disponibles, por tanto, a priori no debería existir un error de construcción asociado, de igual manera el cálculo de ϕ_{pq} es estándar y directo, sin embargo, al hacer la proyección de la red bipartita es donde suceden los sesgos usuales, es decir, la matriz M_{cp} puede generar una distorsión en los cálculos posteriores, por tanto se proponen dos modelos nulos para comparar con el resultado original. Asimismo, existe una posibilidad en las etiquetas de salud sean una fuente sesgo, por tanto, se presentan 3 modelos nulos a comparar con la red original para esclarecer si el resultado es producto del azar o de un efecto real:

^{*}j.illanesva@udd.cl

- **Modelo nulo 1:** Se revuelven los 1s de la matriz M_{cp} tal que se conserve el grado asociado a las empresas.
- **Modelo nulo 2:** Se revuelven los 1s de la matriz M_{cp} tal que se conserve el grado asociado a las comunas.
- **Modelo nulo 3:** Se permutan las etiquetas de salud de los nodos.

Después de testear los modelos nulos se propone establecer la siguiente regresión para explicar las diferencias de los grados de la red de licitaciones:

$$\log(\text{Grado}) = \beta_0 \text{Salud} + \beta_1 \log(\text{MontoTotalMensual}) + \beta_2 \text{Salud} \cdot \log(\text{MontoTotalMensual}) + \varepsilon, \quad (2.1)$$

siendo **Salud** una variable dummy si el nodo es una empresa de salud, **MontoTotalMensual** es la cantidad adjudicada total mensual de la empresa (un proxy de tamaño de dicha empresa).

3 Análisis

Una vez calculada la diferencia entre el grado medio del rubro de la salud y el grado medio de las empresas no relacionadas a la salud nos da un valor de 20.18, es decir, de media, los nodos que no pertenecen al área de la salud, tiene 20 grados menos de aquellos que si pertenecen al área de la salud. Pero como se mencionó anteriormente, ese valor puede deberse al azar, por tanto se compara con los modelos nulos 1, 2 y 3. Véase las figuras (1), (2) y (3). Es claro observar que el valor observado no se explica por el azar estadístico, obteniéndose un valor empírico **p-value <1e-16** para los tres casos. Podemos proceder entonces a aplicar el modelo de regresión cuya tabla de estadísticos se puede ver en (1).

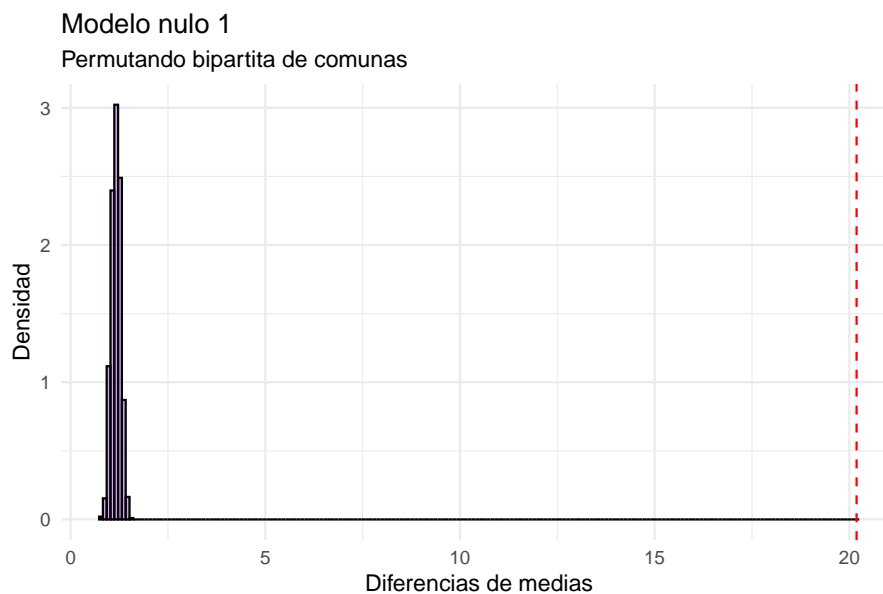


Fig. 1: Modelo nulo al reconectar las empresas manteniendo constante el grado de las comunas. La diferencia observada se observa claramente fuera del histograma de valores obtenidos, correspondiente a 1000 permutaciones. El z-score obtenido al considerar la desviación estándar y la media empírica es de 158.93, por tanto, fuera de toda duda razonable de azar.

	Variable dependiente:			
	(1)	(2)	(3)	(4)
	log(Grado)			
Salud	0.193*** (0.016)		0.189*** (0.015)	-0.064 (0.084)
log(MontoTotalMensual)		0.044*** (0.005)	0.042*** (0.005)	0.034*** (0.006)
Interacción				0.036*** (0.012)
Constante	1.505*** (0.006)	1.229*** (0.036)	1.214*** (0.036)	1.273*** (0.041)
Observaciones	5,880	5,880	5,880	5,880
R ²	0.026	0.013	0.037	0.038
R ² ajustado	0.025	0.012	0.037	0.038
Residual Std. Error	0.433 (df = 5878)	0.436 (df = 5878)	0.431 (df = 5877)	0.430 (df = 5876)
F Statistic	154.306*** (df = 1; 5878)	74.685*** (df = 1; 5878)	112.807*** (df = 2; 5877)	78.416*** (df = 3; 5876)
Note:				
*p<0.1; **p<0.05; ***p<0.01				

Tabla 1: Regresión utilizando las variables mencionadas en la metodología. Se puede apreciar que pertenecer a un área de la salud es significativo en las regresiones (1) y (3) pero no en la (4), naturalmente eso se debe a que el modelo (4) ofrece una interacción entre el MontoTotalMensual, de modo que resulta en un cambio de pendiente más que de un cambio absoluto en el intercepto. Sin embargo, a pesar de ser todos los coeficientes significativos, el R^2 es realmente bajo, explicando casi un 4% únicamente.

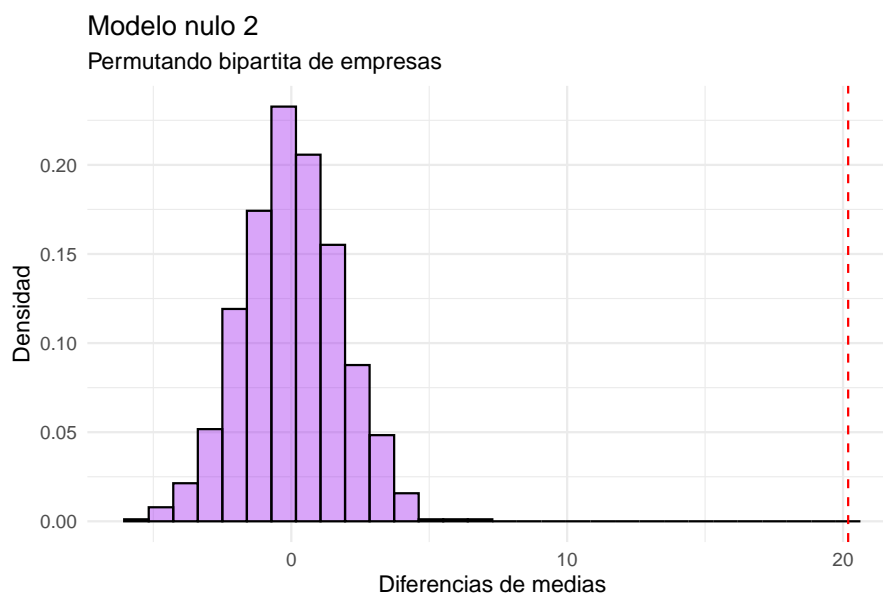


Fig. 2: Modelo nulo al reconectar las comunas manteniendo constante el grado de las empresas. La diferencia observada se observa claramente fuera del histograma de valores obtenidos, correspondiente a 1000 permutaciones. El z-score obtenido al considerar la desviación estándar y la media empírica es de 11, por tanto, fuera de toda duda razonable de azar.

4 Resultados y Conclusión

Es posible apreciar en la tabla (1) las significancias asociadas a la interacción entre el Monto Total Mensual y la etiqueta Salud, por tanto, confirmando una diferencia positiva del resto, esto es, cuando aumenta un 1% el Monto Total Mensual en un rubro distinto al de salud, aumenta un 3.4% el grado del nodo, pero si el nodo pertenece al área de la salud, aumentaría un 7%, es decir, casi el doble.

La hipótesis por tanto quedaría confirmada, existe una diferencia, pero hay que cogerlo con matices extremos. Si bien la tabla de regresión dio un resultado positivo, con casi 6 mil observaciones, el ajuste es de un R^2 de tan solo 0.038, por tanto existe una gran parte del fenómeno que no es posible ser explicado por el tamaño de la empresa o la pertenencia a un área específica de las licitaciones, es claro al observar el gráfico de los residuos para ambos casos (7). No obstante, deja abierta la puerta para explorar otras hipótesis y alternativas al fenómeno, como he recalado en los trabajos anteriores, un factor que podemos considerar son las influencias externas o la corrupción, quedando entonces como trabajo futuro de analizar, o en su defecto, ampliar el rango temporal para observar efectos globales, periodicos, o su contrafactual, es decir, aquellas empresas que no lograron licitar efectivamente y se descartaron para el estudio de este trabajo.

Referencias

- [1] Niclas Frederic Sturm, Cristian Candia, Bruno Damásio, and Flávio L. Pinheiro. High earnings through firm influence: the role of hierarchical structures in public procurement. *EPJ Data Sci.*, 14(1):27, March 2025.
- [2] Kamran Bagheri Lankarani, Leila Zarei, Esmaeil Alinezhad, and Adel Sadeghdoost. A novel social-network-analysis-based approach for analyzing complex network of actors involved in accessibility of anti-cancer medications in Iran. *Health Res Policy Sys*, 22(1):178, December 2024.
- [3] Massimiliano Ferraresi, Gianluca Gucciardi, and Leonzio Rizzo. Savings from public procurement centralization in the healthcare system. *European Journal of Political Economy*, 66:101963, 2021.

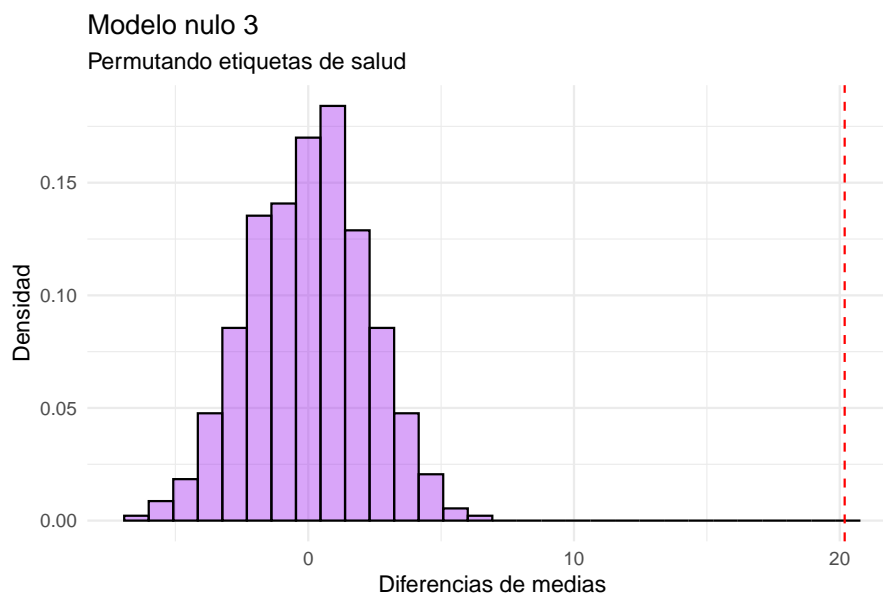


Fig. 3: Modelo nulo al permutar las etiquetas de salud. En el mismo ejercicio de calcular el z-score asociado a la diferencia observada utilizando las media y desviación estandar empírica se obtiene un valor de 9.29, es decir, fuera de toda duda razonable asociada al azar.

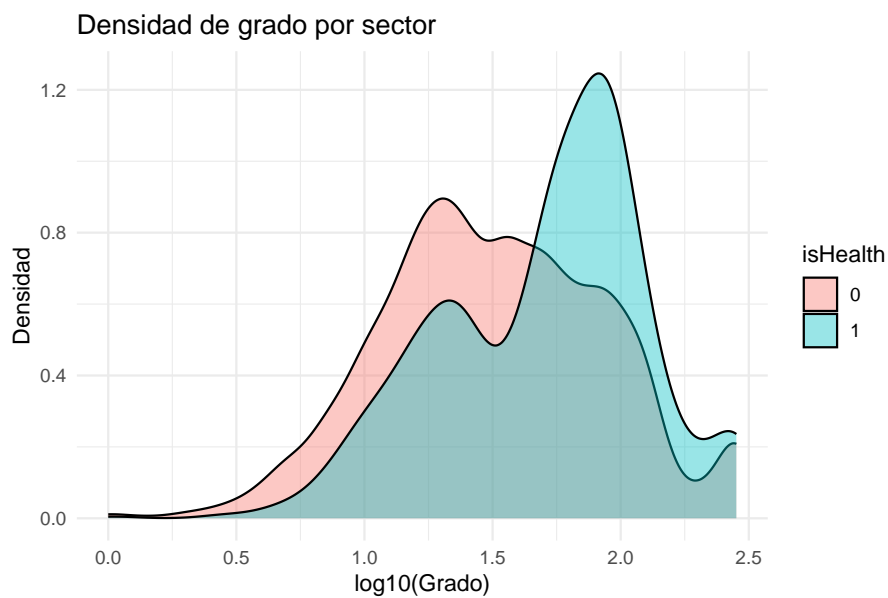


Fig. 4: Densidad de grado por sector. Es posible observar una moda más a la derecha en el caso de salud que el de no salud. A vista, se puede corroborar la hipótesis, pero es necesario confirmarlo con la regresión.

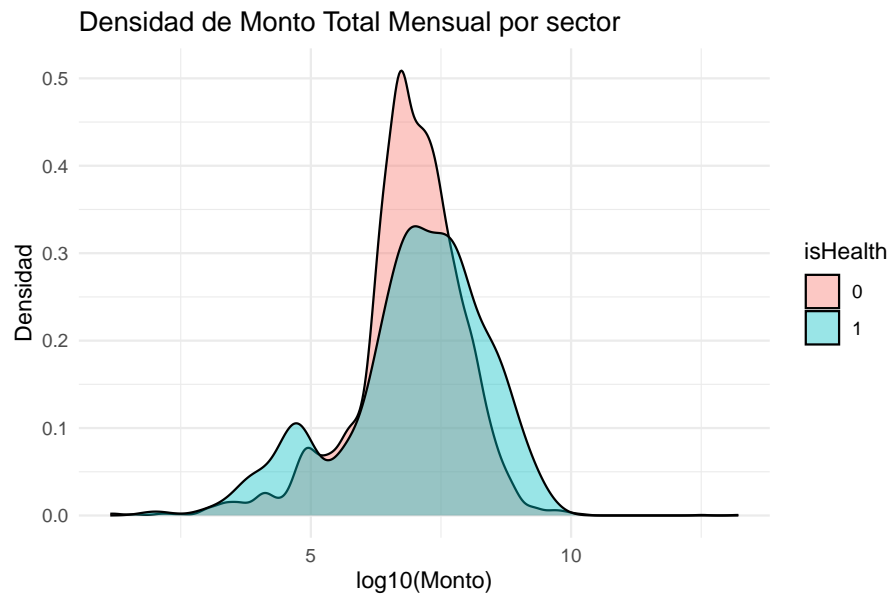


Fig. 5: Densidad de Monto Mensual Total. Se utiliza como métrica de control para evitar sesgos asociados al tamaño de la empresa o de los montos comunes que se tratan en el área de la salud. Se puede observar una leve diferencia mostrando montos más caros que el caso no de salud, pero la media de ambos coincide visualmente, así como su desviación.

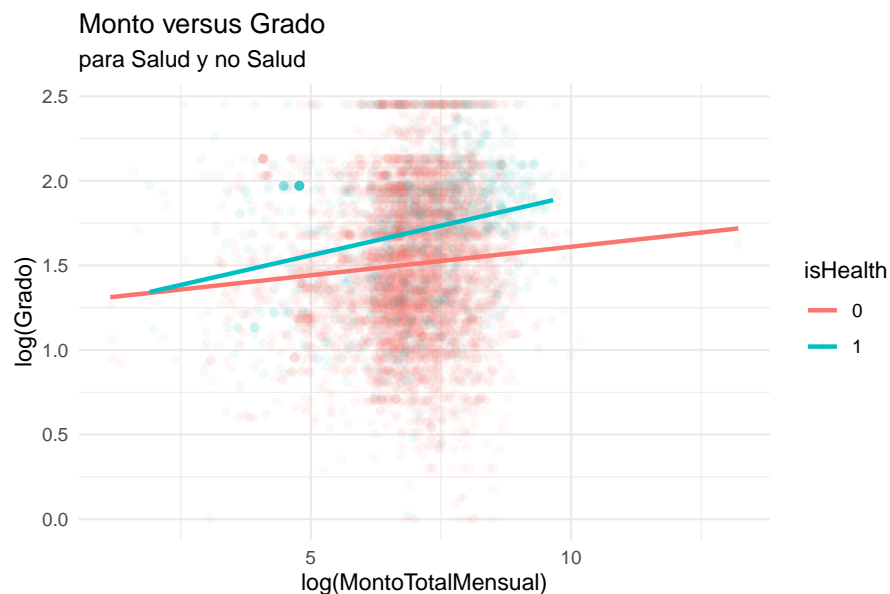


Fig. 6: Gráfico de regresión por sector. Es la visualización de la regresión obtenida de la tabla (1). Se puede observar la diferencia de pendientes en el área de la salud por sobre aquellos nodos no pertenecientes a la salud, aunque la dispersión y ajuste es pobre, explicando solo un 4% del fenómeno.

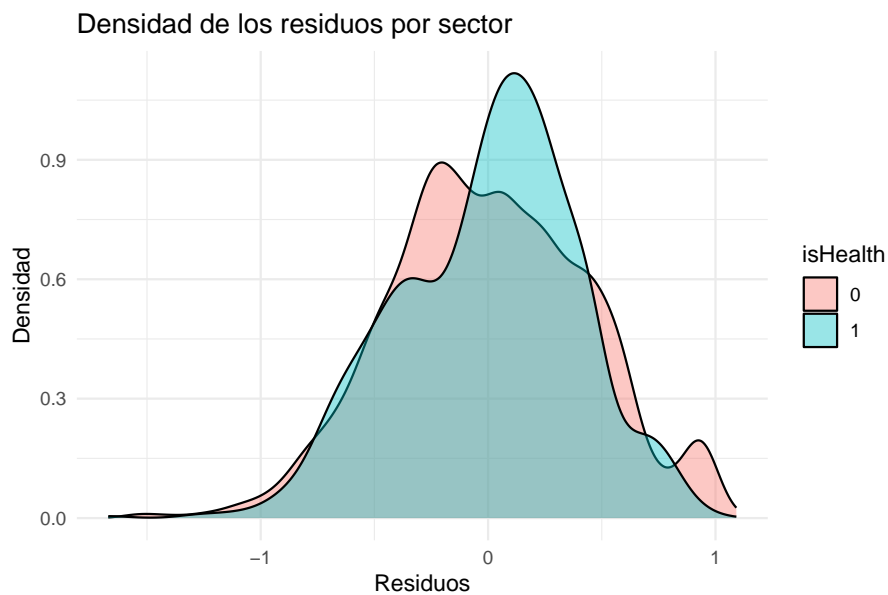


Fig. 7: Densidad de los residuos para ambos casos. Es notable observar la dispersión de los residuos, así como una no normalidad aparente, si se compara con el gráfico (4), no es demasiado lo que cambia, más allá de poseer una media nula.