

PEMODELAN KEMISKINAN DI JAWA TIMUR DENGAN *STRUCTURAL EQUATION MODELING-PARTIAL LEAST SQUARE*

¹Gangga Anuraga dan ²Bambang Widjanarko Otok

^{1,2}Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember,
Surabaya

Alamat e-mail : ¹ganuraga@gmail.com

ABSTRAK

Structural Equation Modeling (SEM) menggambarkan hubungan kausalitas antar variabel yang tidak bisa dijelaskan pada analisis regresi biasa, sehingga dapat diketahui seberapa baik suatu variabel indikator menentukan variabel laten. Seringkali dalam pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang berdekatan (*neighboring*). Sehingga untuk mengatasinya, efek lokasi dimasukkan dalam model. Didapatkan bahwa semua indikator yang terboboti pada setiap variabel laten adalah valid dan signifikan, dengan *reliability* yang baik pada semua variabel laten kemiskinan, ekonomi, SDM, kesehatan. Nilai R-square (R^2) untuk SDM sebesar 0.988 yang artinya model mampu menjelaskan variasi dari SDM pada kasus kemiskinan di Provinsi Jatim sebesar 98.8%, Ekonomi sebesar 0.986 atau 98.6% dan model kemiskinan sebesar 0,925 atau 92.5%.

Kata Kunci : Kemiskinan, *neighboring*, Partial Least Square, *Structural Equation Modeling* (SEM)

PENDAHULUAN

Structural Equation Modeling (SEM) merupakan suatu teknik statistik yang memiliki kemampuan untuk menganalisis pola hubungan antara variabel laten dan indikatornya. Dimana metode ini merupakan pengembangan dari analisis multivariat yaitu analisis faktor dan analisis regresi. SEM mempunyai kemampuan lebih dalam menyelesaikan permasalahan yang melibatkan banyak persamaan linear pada variabel laten [1]. Variabel laten adalah suatu konstruk dalam model persamaan struktural yang tidak dapat diukur secara langsung, tetapi dapat diukur oleh satu atau lebih variabel indikatornya. SEM juga dapat menggambarkan hubungan kausalitas antar variabel yang tidak bisa dijelaskan pada analisis regresi biasa, sehingga dapat diketahui seberapa baik

suatu variabel indikator menentukan variabel laten.

Ada dua hal yang dilakukan dan dihasilkan dalam SEM, yaitu mengestimasi beberapa persamaan yang saling berhubungan secara simultan dengan output model struktural (*structural model*) dan merepresentasikan variabel-variabel laten (*construct / latent / unobserved variables*) berdasarkan variabel-variabel indikator (*manifest / observed variables*) dengan output model pengukuran (*measurement model*).

Menurut Joreskog dan Sorbom dalam [4] model persamaan struktural adalah sebagai berikut :

$$\eta = \mathbf{B}\eta + \mathbf{\Gamma}\xi + \zeta \quad (1)$$

Dimana η (eta) adalah vektor variabel random dependen endogen (*latent endogenous*) dengan ukuran $m \times 1$, ξ (xi) adalah vektor variabel random

independen eksogen (*latent exogenous*) dengan ukuran $n \times 1$, \mathbf{B} adalah matrik koefisien yang menunjukkan pengaruh variabel laten endogen terhadap variabel lainnya dengan ukuran $m \times m$ dan $\mathbf{\Gamma}$ koefisien matrik yang menunjukkan hubungan dari ξ terhadap η dengan ukuran $m \times n$, sedangkan ζ (zeta) adalah vektor random error dengan ukuran $m \times 1$, dengan nilai harapan sama dengan nol.

Asumsi dari persamaan model struktural variabel laten :

$$E(\eta) = 0, E(\xi) = 0, E(\zeta) = 0,$$

ζ tidak berkorelasi dengan ξ dan $(\mathbf{I} - \mathbf{B})^{-1}$ adalah matrik nonsingular.

Selanjutnya adalah model pengukuran atau *Confirmatory Factor Analysis Model* (CFA). Model pengukuran (CFA) dapat dituliskan sebagai berikut [4]:

$$\mathbf{y} = \mathbf{\Lambda}_y \eta + \boldsymbol{\varepsilon} \quad (2)$$

$$\mathbf{x} = \mathbf{\Lambda}_x \xi + \boldsymbol{\delta} \quad (3)$$

Merujuk pada persamaan (2) dan (3) maka matrik kovarian dalam SEM $\Sigma(\theta)$ adalah sebagai berikut [4]:

$$\Sigma(\theta) = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \quad (4)$$

Pendugaan parameter di dalam PLS meliputi 3 hal, yaitu : (1) Estimasi bobot (*Weight estimate*) digunakan untuk menciptakan skor variabel laten, (2) Estimasi jalur (*path estimate*) yang menghubungkan antar variabel laten dan estimasi loading antara variabel laten dengan indikatornya, (3) Means dan lokasi parameter (nilai konstanta regresi, intersep) untuk indikator dan variabel laten.

Estimasi dengan algoritma PLS berlangsung dalam tiga tahap. Langkah pertama dalam estimasi PLS terdiri dari prosedur iterasi regresi sederhana atau regresi berganda dengan memperhitungkan hubungan model struktural / inner model, model pengukuran / outer model dan estimasi bobot / *weight relation*. Kemudian hasil dari estimasi satu set bobot digunakan

untuk menghitung nilai skor variabel laten, yang mana merupakan kombinasi linier dari variabel indikator / manifest. Setelah estimasi skor variabel laten diperoleh, maka langkah kedua dan ketiga melibatkan estimasi koefisien model struktural (*inner model*) dan koefisien dari masing-masing model pengukuran (*outer model*). Pada dasarnya algoritma PLS merupakan serang-kaian regresi sederhana dan berganda dengan estimasi *ordinary least square* [5].

Terdapat dua alasan penting yang mendasari digunakannya SEM. Pertama, SEM memiliki kemampuan untuk mengestimasi hubungan antar variabel yang bersifat *multiple relationship*. Hubungan ini dibentuk dalam model struktural yang digambarkan melalui hubungan antara variabel laten endogen (dependen) dan variabel laten eksogen (independen). Kedua, SEM memiliki kemampuan untuk menggambarkan pola hubungan antara variabel laten dan variabel indikator (*manifest*). Metode SEM menggunakan fungsi *maximum likelihood* untuk meminimalkan perbedaan antara *matrix covariance* yang dibentuk dari sampel data dengan *matrix covariance* yang dibentuk dari prediksi model. Penggunaan SEM memiliki asumsi yang mendasari yaitu multivariat normal dan jumlah sampel yang besar. Penggunaan sampel yang kecil dapat menghasilkan taksiran parameter yang tidak baik bahkan tidak konvergen. Sehingga salah satu pendekatan yang dapat digunakan adalah *partial least square* (PLS). PLS merupakan metode analisis yang *powerful* karena tidak membutuhkan banyak asumsi dan ukuran sampel tidak harus besar.

Partial Least Square (PLS) pertama kali dikembangkan oleh Herman Wold tahun 1975. Model ini dikembangkan sebagai alternatif apabila teori yang mendasari perancangan model lemah. PLS merupakan analisis yang *powerfull* karena dapat digunakan pada setiap jenis

skala data serta syarat asumsi yang lebih fleksibel. PLS juga dapat digunakan ketika landasan teori model adalah *tentative* atau pengukuran setiap variabel laten masih baru, sehingga didesain untuk tujuan prediksi [7].

Metode bootstrap telah dikembangkan sebagai alat untuk membantu mengurangi ketidak-andalan yang berhubungan dengan kesalahan penggunaan distribusi normal dan penggunaannya. Bootstrap standard error dari $\hat{\theta}$ dihitung dengan standard deviasi dari B replikasi [2].

$$\widehat{se}(\hat{\theta}_B) = \sqrt{\frac{\sum_{b=1}^B (\hat{\theta}^*_{(b)} - \hat{\theta}^*_{(.)})^2}{B - 1}}$$

dengan $(\hat{\theta}^*_{(.)}) = \frac{\sum_{b=1}^B \hat{\theta}^*_{(b)}}{B}$, B adalah jumlah kumpulan resampling yang berukuran n dengan *replacement* $\hat{\theta}^*_{(b)}$ adalah statistik $\hat{\theta}$ yang dihitung dari sampel ulang ke-b ($b = 1, \dots, B$).

Efek random struktur spasial tunggal untuk menghitung korelasi diantara variabelnya yang diterapkan pada kasus kesehatan [6]. Model faktor konfirmatori diadopsi dengan korelasi spasial tunggal diantara beberapa indikator sosial [3]. Provinsi Jawa Timur terbagi atas 29 kabupaten dan 9 kota atau secara administratif terdapat 38 kabupaten / kota dan dapat dikatakan berkembang dalam bidang ekonomi. Namun demikian, dalam realitanya masih banyak masyarakat yang hidup di bawah garis kemiskinan. Pada tahun 2011 persentase kemiskinan di Provinsi Jawa Timur masih berada diatas persentase garis kemiskinan tingkat nasional sebesar 13,33%. Daerah kantong kemiskinan masih saja seputar Kabupaten Sampang, Bangkalan, Probolinggo, Sumenep, Pamekasan, Tuban, Pacitan, Kota Probolinggo, Bojonegoro dan Bondowoso. Dimana me-miliki persentase kemiskinan yang lebih tinggi dari garis kemiskinan tingkat Provinsi dan Nasional.

Berdasarkan uraian diatas, maka dilakukan penelitian menggunakan SEM dengan pendekatan *Partial Least Square* (PLS) dalam membentuk model struktural yang diterapkan pada kasus kemiskinan di Kabupaten/Kota Provinsi Jawa Timur tahun 2011, dimana variabel laten endogen yang digunakan antara lain kemiskinan, ekonomi, kesehatan, dan SDM. Dengan demikian tujuan dalam penelitian ini yaitu menyusun model untuk data kemiskinan di Kabupaten/Kota Provinsi Jawa Timur tahun 2011 dengan metode SEM *Partial Least Square*. Sehingga akan didapatkan suatu model kemiskinan di Jawa Timur.

METODE PENELITIAN

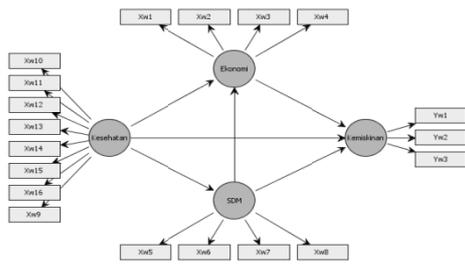
Sumber Data dan Variabel Penelitian

Data yang digunakan pada penelitian ini merupakan publikasi “Data dan Informasi Kemiskinan Kab/Kota Tahun 2011” hasil olah data triwulanan Survei Sosial Ekonomi Nasional (SUSENAS). Informasi yang didapatkan antara lain adalah persentase penduduk miskin, indeks kedalaman kemiskinan, indeks keparahan kemiskinan, bidang kesehatan, pengeluaran rumah tangga, ketenagakerjaan, bidang pendidikan dan fasilitas perumahan. Variabel-variabel yang digunakan dalam penelitian ini disajikan dalam Tabel 1.

Metode Analisis

Metode dan analisis yang digunakan dalam mencapai tujuan penelitian adalah melakukan analisis pemodelan SEM PLS. Dan langkah-langkahnya dapat diuraikan sebagai berikut :

- a. Menyusun model konseptual berbasis teori.
- b. Mengkonstruksi diagram jalur (*path diagram*)
- c. Pemodelan SEM PLS (uji validitas dalam *outer model* dan model struktural / *inner model*).



Gambar 1 Diagram Jalur (Path Diagram)

Tabel 1 Variabel Laten dan Indikator

Kemiskinan :	
Persentase penduduk miskin (Y_1)	
Indeks kedalaman kemiskinan (Y_2)	
Indeks keparahan kemiskinan (Y_3)	
Ekonomi :	
Persentase penduduk miskin usia 15 tahun keatas yang tidak bekerja (X_1)	
Persentase penduduk miskin usia 15 tahun keatas yang bekerja di sektor pertanian (X_2)	
Persentase rumah tangga yang pernah membeli beras raskin (X_3)	
Persentase pengeluaran per kapita untuk non makanan (X_4)	
SDM :	
Persentase penduduk miskin usia 15 tahun keatas yang tidak tamat SD (X_5)	
Angka Melek Huruf penduduk miskin usia 15-55 tahun (X_6)	
Angka Partisipasi Sekolah penduduk miskin usia 13-15 tahun (X_7)	
Rata-rata lama sekolah (X_8)	
Kesehatan :	
Persentase Perempuan pengguna alat KB di rumah tangga miskin (X_9)	
Persentase Balita di rumah tangga miskin yang proses kelahirannya ditolong oleh tenaga kesehatan (X_{10})	
Persentase Balita di rumah tangga miskin yang telah diimunisasi (X_{11})	
Persentase rumah tangga miskin dengan luas lantai perkapita $\leq 8 \text{ m}^2$ (X_{12})	
Persentase rumah tangga miskin yang menggunakan air bersih (X_{13})	
Persentase rumah tangga miskin dengan jamban sendiri/bersama (X_{14})	
Persentase rumah tangga miskin yang mendapatkan pelayanan jamkesmas (X_{15})	
Angka harapan hidup (X_{16})	

Keterangan :

Nilai Y_{w1} dapat diperoleh dengan cara mengalikan matrik terboboti (W) untuk wilayah Jawa Timur dengan indikator Y_1 . Adanya efek lokasi pada indikator di masing-masing variabel laten ini yang membedakan dengan penelitian SEM PLS umumnya.

$$Y_{w1} (38 \times 1) = W_{(38 \times 38)} \cdot Y_1 (38 \times 1)$$

dan dengan cara sama diperoleh juga indikator terboboti pada masing-masing variabel laten.

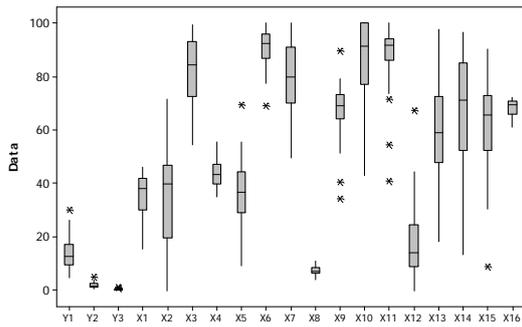
Tabel 2 Struktur data SEM PLS

No Kab/Kota	Y_{w1}	...	Y_{w3}	...	X_{w1}	...	X_{w16}
1. Pacitan
2. Ponorogo
.
.
38. Batu

HASIL PENELITIAN

Kemiskinan di Jawa Timur

Persebaran kantong kemiskinan di Jawa Timur masih saja seputar Kabupaten Sampang, Bangkalan, Probolinggo, Sumenep, Pamekasan, Tuban, Pacitan, Kota Probolinggo, Bojonegoro dan Bondowoso. Dan terdapat keragaman yang cukup tinggi di beberapa indikator pada variabel laten kemiskinan, ekonomi, SDM dan Kesehatan. Artinya terdapat kesenjangan antara kabupaten/kota di Provinsi Jawa Timur, sesuai dengan karakteristik wilayah itu sendiri. Untuk memperjelas gambaran kemiskinan di Jawa Timur tahun 2011 berdasarkan indikator-indikator yang memengaruhi disajikan dalam bentuk *Boxplot* pada Gambar 2 sebagai berikut :



Gambar 2 Boxplot Indikator Kemiskinan

SEM Partial Least Square

1. Uji validitas dan reliabilitas pada Outer Model

Berdasarkan Tabel 3 menunjukkan bahwa estimasi nilai loading pada masing-masing variabel laten adalah signifikan, hal ini ditunjukkan dengan nilai t-statistik yang lebih besar dari t-tabel 1,960 (2-tailed) pada tingkat signifikansi alpha 0.05.

Tabel 3 Uji signifikansi t-statistik bootstrap 400

Variabel laten	$\hat{\lambda}_{jk}$	$se(\hat{\lambda}_{jk})$	$t = \frac{\hat{\lambda}_{jk}}{se(\hat{\lambda}_{jk})}$
Kesehatan	0.996	1.00E-03	9.96E+02
	0.994	1.87E-03	5.32E+02
	0.995	1.50E-03	6.63E+02
	0.749	7.98E-02	9.39E+00
	0.98	5.51E-03	1.78E+02
	0.946	1.97E-02	4.80E+01
	0.974	8.01E-03	1.22E+02
	0.997	1.08E-03	9.23E+02
SDM	0.943	1.90E-02	4.96E+01
	0.997	9.12E-04	1.09E+03
	0.987	5.34E-03	1.85E+02
	0.988	4.12E-03	2.40E+02
Ekonomi	0.97	8.65E-03	1.12E+02
	0.934	1.73E-02	5.40E+01
	0.997	8.82E-04	1.13E+03
	0.989	3.25E-03	3.04E+02
Kemiskinan	0.996	1.28E-03	7.78E+02
	1.000	8.26E-05	1.21E+04
	0.997	1.19E-03	8.38E+02

Tabel 4 Hasil Uji Reliabilitas (Composite Reliability)

Variabel Laten	Reliabilitas (Composite Reliability)	Keterangan
Kesehatan	0.989	Reliabel
SDM	0.989	Reliabel
Ekonomi	0.986	Reliabel
Kemiskinan	0.998	Reliabel

2. Uji validitas dan reliabilitas pada Outer Model

Berdasarkan Tabel 3 menunjukkan bahwa estimasi nilai loading pada masing-masing variabel laten adalah signifikan, hal ini ditunjukkan dengan nilai t-statistik yang lebih besar dari t-tabel 1,960 (2-tailed) pada tingkat signifikansi alpha 0.05.

Tabel 4 menunjukkan bahwa semua blok indikator yang mengukur konstruk Kesehatan, SDM, Ekonomi, dan Kemiskinan > 0.7. Sehingga variabel laten kesehatan, SDM, ekonomi, dan kemiskinan dikatakan memiliki reliabilitas yang baik.

3. Persamaan struktural / inner model

Berdasarkan hasil yang diperoleh diketahui terdapat tiga persamaan sebagai berikut :

$$\text{Kemiskinan} = 1.338 \text{ Ekonomi} + 1.148 \text{ SDM} - 1.564 \text{ Kesehatan}$$

$$\text{Ekonomi} = 1.377 \text{ SDM} - 0.387 \text{ Kesehatan}$$

$$\text{SDM} = 0.994 \text{ Kesehatan}$$

Model kemiskinan (η_{w1}) dapat diketahui bahwa ekonomi memiliki pengaruh positif dan signifikan terhadap variabel laten kemiskinan, SDM juga memiliki pengaruh positif tetapi tidak signifikan, sedangkan variabel kesehatan memiliki pengaruh negatif dan tidak signifikan terhadap model kemiskinan. Untuk model ekonomi (η_{w2}) diketahui bahwa variabel laten SDM memiliki pengaruh yang positif dan signifikan

terhadap variabel laten ekonomi pada selang kepercayaan 95% maupun secara t-statistik, dengan estimasi koefisien jalur didapatkan sebesar 1.377, sedangkan variabel laten kesehatan memiliki pengaruh negatif dan tidak signifikan dengan estimasi koefisien jalur sebesar 0.387. Dan untuk model SDM (η_{w3}) dapat diketahui bahwa variabel laten kesehatan memiliki korelasi yang positif dan signifikan terhadap SDM pada selang kepercayaan 95%, maupun secara t-statistik dengan estimasi koefisien jalur sebesar 0,994.

Nilai uji evaluasi terhadap model persamaan struktural pada SEM PLS dapat diketahui dari nilai *goodness of fit* atau R^2 . Hasil pengolahan data penelitian ini dengan menggunakan R memberikan nilai R-square (R^2) sebagaimana nampak pada Tabel 5 berikut:

Tabel 5 R-Square (R^2)

Variabel Laten	R-square	Keterangan
SDM	0.988	Baik
Ekonomi	0.986	Baik
Kemiskinan	0.925	Baik

Nilai R-square (R^2) untuk SDM sebesar 0.988 yang artinya model mampu menjelaskan variasi dari SDM pada kasus kemiskinan di Provinsi Jatim sebesar 98.8%, Ekonomi sebesar 0.986 atau 98.6% dan model kemiskinan sebesar 0,925 atau 92.5%.

KESIMPULAN DAN SARAN

Kesimpulan

Didapatkan model struktural untuk kasus kemiskinan di Jawa Timur adalah sebagai berikut :

$$\begin{aligned} \text{Kemiskinan} = & 1.338 \text{ Ekonomi} \\ & + 1.148 \text{ SDM} \\ & - 1.564 \text{ Kesehatan} \end{aligned}$$

$$\begin{aligned} \text{Ekonomi} = & \\ & 1.377 \text{ SDM} - 0.387 \text{ Kesehatan} \\ \text{SDM} = & 0.994 \text{ Kesehatan} \end{aligned}$$

Dengan nilai R-square (R^2) untuk SDM sebesar 0.988 yang artinya model mampu menjelaskan variasi dari SDM pada kasus kemiskinan di Provinsi Jatim sebesar 98.8%, Ekonomi sebesar 0.986 atau 98.6% dan model kemiskinan sebesar 0.925 atau 92.5%.

Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah pemodelan persamaan struktural SEM PLS dengan mengembangkan indikator kemiskinan yang digunakan.

DAFTAR PUSTAKA

[1] Bollen, K., (1989), *Structural Equations With Latent Variables*, John Wiley & Sons, Inc, New York.

[2] Efron, B. and Tibshirani, R.J., (1993), *An Introduction to the Bootstrap*, Chapman and Hall, Inc., New York.

[3] Hogan, J. dan Tchernis, R., (2004), *Bayesian factor analysis for spatially correlated data with application to summarizing area-level material deprivation from census data*, J. Amer, Statist Assoc.99,314-324.

[4] Schumacker dan Lomax., (2004), *A Beginner's Guide To Structural Equation Modeling*, Lawrence Erlbaum Associates Inc, New Jersey.

[5] Tenenhaus, M., Vinzi, dan Chatelin., (2002), *State-of-Art On PLS Path Modeling Trough The Available Software.4t*

[6] Wang, F dan Wall, M., (2003), *Generalized common spatial factor model*, Biostatistics, page:569-582.

[7] Yamin dan Kurniawan, (2011), *Mengolah Data Penelitian dengan Partial Least Square Path-Modeling*, Salemba Infotek, Jakarta.