

**PENERAPAN *RESPONSE BASED UNIT SEGMENTATION*  
*IN PARTIAL LEAST SQUARE (REBUS-PLS)*  
UNTUK ANALISIS DAN PENGELOMPOKAN WILAYAH  
(Studi Kasus: Kesehatan Lingkungan Perumahan di Provinsi Jawa Tengah)**

**Febriana Sulistya Pratiwi<sup>1</sup>, Sudarno<sup>1</sup>, Agus Rusgiyono<sup>1</sup>**

<sup>1, 1, 1</sup> Departemen Statistika FSM Universitas Diponegoro

[febrianasulistyp@gmail.com](mailto:febrianasulistyp@gmail.com)

**ABSTRACT**

Residential environmental health is a complex problem that depends on several dimensions. One of the statistical method that can be used to analyze the relation between complex dimensions is Structural Equation Modeling (SEM) with a variant/component based approach or Partial Least Square. The purpose of this study is to develop a structural model of the relation between household economy, education, housing facilities, and residential environmental health in Central Java Province in 2018 based on 12 valid and reliable indicators. In the structural equation model there is a significant positive effect path that is the influence of household economy towards education and towards housing facilities, and influence housing facility on the residential environment health. In SEM analysis it is generally assumed that the data taken comes from a homogeneous population but often the data consists of several segments. Therefore, we need a method to detect heterogeneity problems, namely Response Based Unit Segmentation in Partial Least Square (REBUS-PLS). Based on the dendrogram produced, by forming 2 classes/segments,  $R^2$  values as the accuracy of the prediction model on the local model had a higher value (except  $R^2$  values for Education in local model 2) than  $R^2$  values on the global model. In addition, the Goodnes of Fit value as a measure of model suitability for each local model is also had a higher value, so that it indicates the goodness of the model in the local model is better than the global model.

**Keywords:** environmental health, SEM, PLS, REBUS-PLS

**1. PENDAHULUAN**

Salah satu kebutuhan mendasar yang sangat diperlukan bagi kelangsungan hidup seseorang atau rumah tangga adalah rumah atau tempat tinggal. Sebagai upaya pencegahan penyakit dan/atau gangguan kesehatan dari faktor risiko lingkungan, memperhatikan kesehatan lingkungan perumahan merupakan hal yang amat perlu guna mewujudkan kualitas lingkungan yang sehat dari berbagai aspek (BPS, 2019). Berkaitan dengan kesehatan lingkungan perumahan, Mayasari dan Ritohardoyo (2012) menyatakan kualitas perumahan dan permukiman mencakup mengenai kondisi bangunan rumah, lingkungan rumah dan manusia yang ada di dalamnya. Demikian pula komponen manusia secara individu juga akan berpengaruh. Sektor pendidikan, pendapatan, pekerjaan juga dapat mempengaruhi usaha manusia dalam membangun kualitas permukimannya.

Salah satu metode statistika yang dapat digunakan untuk menganalisis keterkaitan antar dimensi kompleks yang mempengaruhi kesehatan lingkungan perumahan khususnya di Provinsi Jawa Tengah adalah *Structural Equation Modelling* (SEM). SEM mampu menjelaskan hubungan tersebut dan dapat mengeksplorasi indikator-indikator lain yang dapat menjelaskan pengaruhnya terhadap dimensi kesehatan lingkungan. Oleh sebab keterbatasan yang ada karena asumsi data yang harus berdistribusi normal, indikator harus dalam bentuk reflektif, model harus berdasar teori, dan adanya *indeterminacy*, maka sebagai alternatif dikembangkan SEM dengan pendekatan berbasis varian atau *Partial Least Square* (SEM-PLS) untuk mengatasi permasalahan pemenuhan asumsi tersebut. Salah satu penelitian sebelumnya yang menggunakan SEM-PLS adalah yang dilakukan oleh Eva Umami Nikmaty Sholihah dan Mutiah Salamah (2015) yang menggunakan metode tersebut untuk pemodelan derajat kesehatan kabupaten/kota di Jawa Timur berdasarkan Indeks Pembangunan Kesehatan Masyarakat.

Pada umumnya, dalam analisis SEM para peneliti mengasumsikan data yang diambil berasal dari populasi yang sama (homogen), tetapi observasi seringkali terdiri atas beberapa

segmen atau kelas. Bila asumsi ini digunakan maka dalam beberapa hal akan berakibat tidak validnya kesimpulan yang diambil sehingga interpretasi hasil penelitian menjadi kurang tepat dan tidak rasional. Maka dari itu, diperlukan suatu metode yang dapat mendeteksi problem heterogenitas diantaranya menggunakan metode baru untuk pendeteksian *unobserved heterogeneity* dalam PLS Path Modeling yaitu *Response Based Unit Segmentation in Partial Least Square* (REBUS-PLS). REBUS-PLS dapat mensegmentasi unit observasi sekaligus mengestimasi parameter dari masing-masing model lokal pada segmen yang terbentuk (Trincher, 2007), sebagaimana Vinzi *et al.* (2008) yang telah menggunakan REBUS-PLS untuk mendeteksi heterogenitas pada model SEM-PLS dimensi kualitas hidup (*quality of life*) negara di dunia.

Ketersediaan data statistik perumahan diharapkan dapat membantu membuat kebijakan mengenai perumahan dan permukiman yang sesuai dengan kebutuhan masyarakat untuk mewujudkan kualitas lingkungan yang sehat dari berbagai aspek. Berdasarkan uraian tersebut, akan dikaji bagaimana pengaruh dimensi ekonomi, sosial, dan sarana prasarana dalam mempengaruhi kesehatan lingkungan. Dimensi-dimensi tersebut digambarkan dengan kondisi ekonomi rumah tangga, pendidikan sebagai status sosial, serta fasilitas perumahan sebagai segi ketersediaan sarana pendukung kesehatan lingkungan perumahan dengan indikator yang valid dan reliabel. Selanjutnya, pengelompokan/segmentasi wilayah terhadap hasil struktur model kesehatan lingkungan perumahan dianalisis dengan REBUS-PLS.

## 2. TINJAUAN PUSTAKA

*Partial Least Square* (PLS) termasuk dalam analisis SEM yang berbasis *variance (component base)*. Analisis *Partial Least Square* (PLS) termasuk dalam teknik statistika multivariat yang melakukan perbandingan antara variabel dependen berganda dan variabel independen berganda.

### 2.1 Estimasi Parameter Model *Partial Least Square* (PLS)

Menurut Sanchez (2013), pemodelan jalur PLS mengikuti prosedur berurutan yang dapat dibagi dalam tiga tahap utama.

#### a. Estimasi Bobot (*Weight Estimate*)

Pada tahap awal, dilakukan proses inisialisasi yaitu menentukan nilai sembarang untuk *outer weight*. Bobot untuk seluruh indikator dibuat sama, yaitu 1 (satu).

Langkah selanjutnya yaitu melakukan aproksimasi eksternal dengan rumus:

$$\hat{Y}_j = \sum_{h=1}^H w_{jh} X_{jh} \quad (1)$$

Kemudian dilakukan aproksimasi internal dengan rumus sebagai berikut:

$$\hat{Z}_j = \sum_{i=1}^I e_{ji} \hat{Y}_i \quad (2)$$

dimana  $\hat{Z}_j$  adalah data variabel laten dugaan yang akan diestimasi ulang dan  $e_{ji}$  merupakan hubungan antara variabel laten  $j$  dengan variabel laten  $i$ . Bobot *inner model* ( $e_{ji}$ ) dipilih melalui tiga skema yaitu skema jalur (*path scheme*), skema centroid (*centroid scheme*), dan skema faktor (*factor scheme*). Setelah didapatkan nilai *inner weight*, dilakukan estimasi sesuai persamaan (2).

Jika tahapan aproksimasi *inner weight* atau internal selesai, langkah selanjutnya adalah memperbarui *outer weight*. Afifah dan Sunaryo (2013) menyatakan dalam memperoleh *weight estimate* ini, dibedakan atas *mode A* (tipe indikator reflektif) dan *mode B* (tipe indikator formatif).

#### *Mode A*

Estimasi untuk *mode A* diperoleh menggunakan metode OLS dengan cara meminimumkan jumlah kuadrat  $e_{jh}$ . Dari persamaan  $X_{jh} = w_h Z_j + e_{jh}$  diperoleh  $e_{jh} = X_{jh} - w_h Z_j$ . Selanjutnya, jumlah kuadrat  $e_{jh}$  diturunkan terhadap  $w_h$

$$\hat{w}_h = E \left[ \frac{\sum_{j=1}^J (X_{jh} Z_j)}{\sum_{j=1}^J (Z_j^2)} \right] = \frac{E \left[ \sum_{j=1}^J (X_{jh} Z_j) \right]}{E \left[ \sum_{j=1}^J (Z_j^2) \right]} = \frac{Cov(X_{jh}, Z_j)}{Var(Z_j)} \quad (3)$$

*Mode B*

Pada *mode B* untuk tipe indikator formatif, dari persamaan  $Z_j = w_{jh} X_{jh} + \varepsilon_j$  diperoleh  $\varepsilon_j = Z_j - w_{jh} X_{jh}$ . Dengan menghitung  $\varepsilon_j^T \varepsilon_j$  maka bobot untuk *mode B* adalah:

$$\hat{w}_{jh} = (X_{jh}^T X_{jh})^{-1} X_{jh}^T Z_j \quad (4)$$

dimana  $Z_j$  adalah matriks berisi data variabel laten ke- $j$ ,  $X_{jh}$  adalah vektor indikator ke- $h$  variabel laten ke- $j$  dan  $w_{jh}$  adalah *outer weight* indikator ke- $h$  variabel laten ke- $j$ .

Langkah-langkah di atas dilakukan hingga konvergen. Jika dalam setiap iterasi terdapat  $s = 1, 2, \dots, S$ , maka iterasi hingga  $S$  adalah konvergen. Kriteria konvergensi dalam hal ini adalah  $|w_{jh}^{s-1} - w_{jh}^s| < 10^{-5}$  (Sanchez, 2013).

b. Estimasi Koefisien Jalur (*Path Coefficient*)

Untuk koefisien jalur model struktural diduga dengan *ordinary least square* pada regresi berganda  $Y_j$  dan  $Y_i$  yang bersesuaian.

$$\beta_{ji} = (Y_i^T Y_i)^{-1} Y_i^T Y_j \quad (5)$$

c. Estimasi Koefisien *Loadings* (*Outer Model*)

Koefisien *loading* diperoleh dengan menghitung korelasi antara variabel laten dengan masing-masing indikatornya.

$$\lambda_{jh} = cor(X_{jh}, Y_j) \quad (6)$$

## 2.2 Evaluasi Model *Partial Least Square* (PLS)

a. Evaluasi Model Pengukuran atau *Outer Model*

Pada model dengan indikator reflektif, Yamin dan Kurniawan (2011) menyatakan pemeriksaan individual dalam evaluasi *convergent validity* ini diantaranya:

**Tabel 1.** Ukuran Evaluasi Model Pengukuran Reflektif

Ukuran Evaluasi	Parameter	Kriteria
<i>Item reliability</i> atau validitas indikator	<i>Loading factor</i> ( <i>standardized loadings</i> )	Indikator dapat dinyatakan valid jika nilai <i>loading factor</i> > 0,5
<i>Internal consistency</i> atau <i>construct reliability</i>	<i>Composite reliability</i> ( $\rho$ ) atau <i>Dillon-Glodstein's</i> ( <i>D.G. rho</i> )	> 0,7 berarti dapat diterima dan > 0,8 dan 0,9 berarti sangat memuaskan
<i>Average variance extracted</i> (AVE)	<i>Average variance extracted</i> (AVE)	Ukuran <i>convergent validity</i> yang baik memiliki AVE minimal 0,5

b. Evaluasi Model Struktural atau *Inner Model*

Yamin dan Kurniawan (2011) menyatakan model struktural dapat dievaluasi dengan ukuran sebagai berikut, diantaranya:

**Tabel 2.** Ukuran Evaluasi Model Struktural

Ukuran Evaluasi	Kriteria
Signifikansi <i>Path coefficient</i>	Nilai <i>t test</i> ( <i>critical ratio</i> ) yang diperoleh dari proses <i>bootstrapping</i> > 1,96 ( $\alpha = 5\%$ )
$R^2$ ( <i>R-square</i> )	Nilai $R^2$ ini dalam tiga klasifikasi, yaitu nilai $R^2$ 0,67; 0,33; dan 0,19 sebagai substansial, moderat, dan lemah.
<i>Goodness of fit</i> (GoF)	GoF kecil jika bernilai 0,1; GoF moderat yang bernilai 0,25; dan GoF besar dengan nilai 0,36

### 2.3 Pengujian Hipotesis

PLS tidak mengasumsikan data harus mengikuti distribusi tertentu seperti berdistribusi normal multivariat. Pengujian hipotesis untuk menilai signifikansi *path coefficient* dilihat dari nilai *t-test (critical ratio)* yang diperoleh dari *resampling method* dengan proses *bootstrapping* (Yamin dan Kurniawan, 2011).

Pada model pengukuran (*outer model*), pengujian parameter  $\lambda$  menggunakan hipotesis sebagai berikut:

$H_0 : \lambda_h = 0$ , (*loading factor* tidak signifikan dalam mengukur variabel laten)

$H_1 : \lambda_h \neq 0$  (*loading factor* signifikan dalam mengukur variabel laten)

dengan  $h = 1, 2, 3, \dots, n$  merupakan variabel *manifest/indikator*.

Untuk statistik uji yang digunakan adalah statistik uji *t (T-statistics)* sebagai berikut:

$$T = \frac{\hat{\lambda}_h}{se^*(\hat{\lambda}_h)} \quad (7)$$

dimana  $Se^*(\hat{\lambda}_h)$  merupakan *standard error bootstrap* dari koefisien  $\hat{\lambda}_h$ .

$H_0$  ditolak jika nilai *t-statistic* atau  $|CR| > 2$  yang menunjukkan *loading factor* tersebut signifikan, sehingga dapat dikatakan valid (Yamin dan Kurniawan, 2011).

Pada model struktural (*inner model*), pengujian hipotesis dilakukan terhadap parameter beta ( $\beta$ ) dan parameter gamma ( $\gamma$ ). Parameter beta ( $\beta$ ) merupakan parameter yang menunjukkan pengaruh antar variabel endogen. Pengujian parameter  $\beta$  menggunakan hipotesis sebagai berikut:

$H_0 : \beta_{ji} = 0$ , (variabel laten endogen tidak signifikan dalam mengukur variabel laten endogen lainnya)

$H_0 : \beta_{ji} \neq 0$ , (variabel laten endogen signifikan dalam mengukur variabel laten endogen lainnya)

dengan  $j = 1, 2, 3, \dots, l$  merupakan banyaknya variabel laten endogen dependen, dan  $i = 1, 2, 3, \dots, m$  merupakan banyaknya variabel laten endogen independen.

Dengan statistik uji *t (T-statistics)*:

$$T = \frac{\hat{\beta}_{ji}}{se^*(\hat{\beta}_{ji})} \quad (8)$$

dimana  $Se^*(\hat{\beta}_{ji})$  merupakan *standard error bootstrap* dari koefisien  $\hat{\beta}_{ji}$ .

Pada level signifikansi 5% dengan tes dua arah (*two-tailed*) keputusan yang diambil adalah menolak  $H_0$  saat ukuran dari nilai empiris *t* yang dihasilkan  $|T - statistics| > 1,96$  atau nilai *p-value*  $< \alpha (0,05)$ .

Parameter gamma ( $\gamma$ ) merupakan parameter yang berhubungan dengan pengaruh variabel laten eksogen terhadap variabel laten endogen. Pengujian parameter  $\gamma$  analog dengan pengujian hipotesis parameter  $\beta$ .

### 2.4 Response Based Unit Segmentation in Partial Least Square (REBUS-PLS)

REBUS-PLS adalah algoritma iteratif yang dapat mengelompokkan atau mensegmentasi unit observasi sekaligus mengestimasi parameter dari masing-masing model lokal pada segmen yang terbentuk tanpa membutuhkan segala jenis asumsi distribusi baik pada variabel *manifest* atau pada variabel laten (Trincherá, 2007).

#### Algoritma REBUS-PLS

Ukuran untuk menilai jarak antara unit dan model dalam REBUS-PLS dihitung dari residual model *communality* yaitu menggunakan *Closeness Measure (CM)*. Trincherá (2007) menyatakan nilai CM didefinisikan sebagai berikut.

$$CM_{ik} = \sqrt{\frac{\sum_{q=1}^Q \sum_{p=1}^{P_q} [e_{ipqk}^2 / Com(\xi_{qk}, x_{pq})]}{\sum_{i=1}^N \sum_{q=1}^Q \sum_{p=1}^{P_q} [e_{ipqk}^2 / Com(\xi_{qk}, x_{pq})]} \times \frac{\sum_{j=1}^{J^*} [f_{ijk}^2 / R^2(\xi_j, \{\xi_{q^*}'s \text{ explaining } \xi_j\})]}{\sum_{i=1}^N \sum_{j=1}^J [f_{ijk}^2 / R^2(\xi_j, \{\xi_{q^*}'s \text{ explaining } \xi_j\})]} \quad (9)$$

dengan  $Com(\xi_{qk}, x_{pq})$  merupakan indeks komunal indikator ke- $p$  dari blok ke- $q$  di dalam kelas laten ke- $k$ ;  $e_{ipqk}^2$  merupakan ukuran residual model untuk unit ke- $i$  di dalam kelas laten ke- $k$  yang merujuk variabel *manifest* ke- $p$  di dalam blok ke- $q$ ;  $f_{ijk}$  menunjukkan residual model struktural untuk unit ke- $i$  di dalam kelas laten ke- $k$  yang merujuk blok endogen ke- $j$ ;  $N$  adalah jumlah dari unit; dan  $m_k$  adalah jumlah komponen yang diekstraksi. Oleh karena semua blok harus reflektif, angka ini akan selalu sama dengan 1.

Untuk mendapatkan model lokal yang lebih baik daripada model global, ukuran kedekatan yang dipilih ditentukan sesuai dengan struktur indeks *Goodness of Fit* (GoF). Indeks GoF sebagaimana menurut Trinchera (2007) didefinisikan sebagai berikut.

$$GoF = \sqrt{\frac{\sum_{q=1}^Q \sum_{p=1}^{P_q} Cor^2(x_{pq}, \xi_q)}{\sum_{q=1}^Q P_q} \times \frac{\sum_{j=1}^J R^2(\xi_j, \{\xi_{q^*}'s \text{ explaining } \xi_j\})}{J}} \quad (10)$$

Dengan  $Q$  adalah jumlah variabel laten endogen dan eksogen,  $J$  adalah jumlah variabel laten endogen, dan  $P_q$  adalah jumlah variabel *manifest* dalam blok. Indeks rata-rata *communality* dapat dianggap sebagai indeks yang mengukur kualitas model pengukuran yang mana diperoleh sebagai rata-rata korelasi kuadrat yang menghubungkan setiap variabel *manifest* ( $x_{pq}$ ) dengan variabel laten koresponden ( $\xi_q$ ) di semua blok  $Q$ .

Langkah pertama dari algoritma REBUS-PLS yaitu perhitungan model global menggunakan PLS-PM biasa. Kemudian dihitung *communality* dan residual struktural dari masing-masing unit dari model global. Jumlah kelas ( $K$ ) diperoleh selama iterasi berturut-turut dan komposisi awal kelas diperoleh berdasarkan hasil *hierarchical cluster analysis*. Selanjutnya yaitu dihitung model lokal pada tiap kelas/segmen yang terbentuk dengan menggunakan PLS-PM biasa. Parameter spesifik segmen yang dihitung pada langkah sebelumnya digunakan untuk menghitung *communality* dan *residual structural* dari model lokal sehingga nilai CM setiap unit dari model lokal dapat diperoleh. Selanjutnya, setiap unit observasi dimasukkan ke dalam kelas yang menunjukkan nilai CM yang lebih kecil. Setelah komposisi kelas diperbarui, sejumlah  $K$  model lokal baru diperkirakan. Algoritma akan terus menerus melakukan iterasi hingga tidak ada perubahan unit observasi dalam komposisi kelas yang mana dianggap sebagai *stopping rule* (Trinchera, 2007).

### 3. METODOLOGI PENELITIAN

#### 3.1 Sumber Data dan Variabel Penelitian

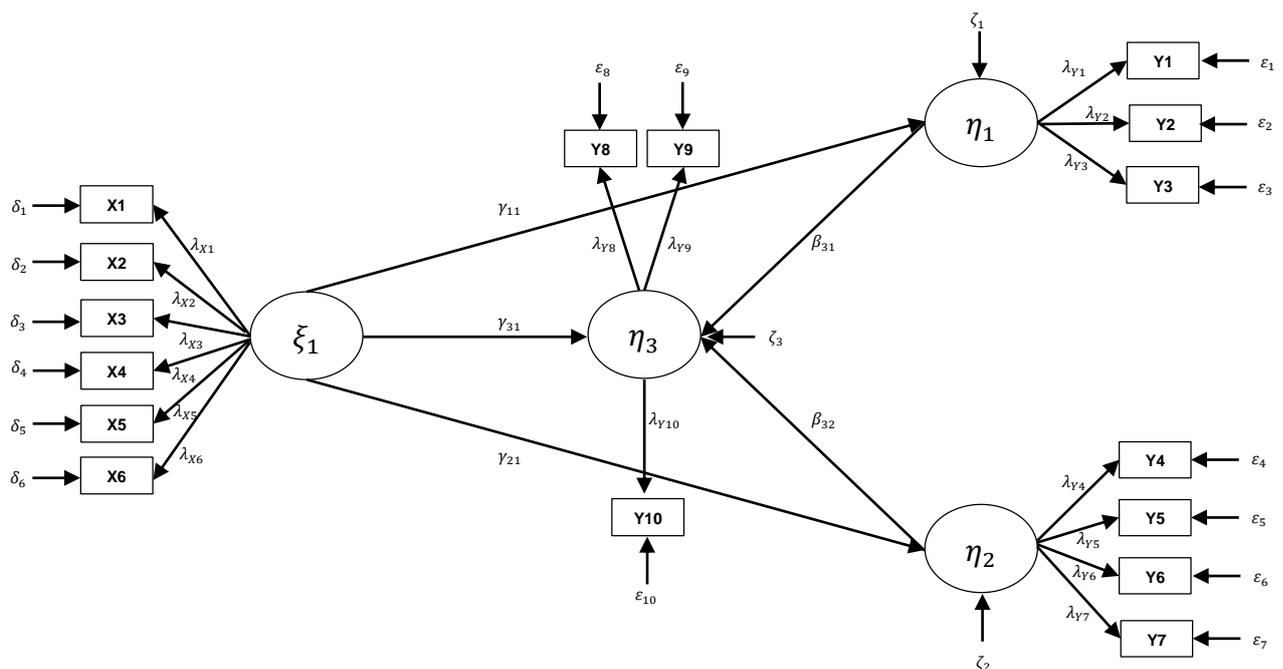
Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari publikasi BPS Provinsi Jawa Tengah. Data memiliki 35 unit observasi yang masing-masing merupakan kabupaten/kota di Provinsi Jawa Tengah. Sumber data yang digunakan dalam publikasi BPS tersebut adalah data hasil Survei Sosial Ekonomi Nasional (Susenas) tahun 2018. Pelaksanaan Susenas Maret 2018 mencakup 27.520 rumah tangga sampel yang meliputi wilayah perkotaan dan perdesaan dan tersebar di seluruh kabupaten/kota di Jawa Tengah.

**Tabel 3.** Variabel Konstruk dan Indikator

Variabel Laten (Konstruk)	Variabel <i>Manifest</i> (Indikator)
1. $X_1$	Persentase Rumah Tangga dengan Status Kepemilikan Bangunan Tempat Tinggal Milik Sendiri

Ekonomi Rumah Tangga ( $\xi_1$ )	$X_2$	Persentase Rumah Tangga dengan Jenis Lantai Terlalu Besar Bukan Tanah
	$X_3$	Persentase Rumah Tangga dengan Luas Lantai Bangunan Tempat Tinggal $\geq 150 \text{ m}^2$
	$X_4$	Persentase Rumah Tangga Jenis Dinding Terlalu Besar adalah Tembok
	$X_5$	Persentase Rumah Tangga yang Memiliki Telepon
	$X_6$	Persentase Rumah Tangga yang Memiliki Komputer/Laptop
	2. Pendidikan ( $\eta_1$ )	$Y_1$
	$Y_2$	Rata-rata Lama Sekolah (tahun) Penduduk Usia 25 Tahun ke Atas
	$Y_3$	Persentase Penduduk Usia 15 Tahun ke Atas dengan Pendidikan Tertinggi yang Ditamatkan Jenjang SMA/Sederajat
3. Fasilitas Perumahan ( $\eta_2$ )	$Y_4$	Persentase Rumah Tangga dengan Sumber Penerangan Utama adalah Listrik PLN
	$Y_5$	Persentase Rumah Tangga dengan Penggunaan Fasilitas Tempat Buang Air Besar Sendiri
	$Y_6$	Persentase Rumah Tangga dengan Tempat Pembuangan Akhir Tinja adalah Tangki Septik
	$Y_7$	Persentase Rumah Tangga dengan Jenis Kloset berupa Leher Angsa
4. Kesehatan Lingkungan Perumahan ( $\eta_3$ )	$Y_8$	Persentase Rumah Tangga yang Memiliki Akses terhadap Air Minum Layak
	$Y_9$	Persentase Rumah Tangga yang Memiliki Akses terhadap Sanitasi Layak
	$Y_{10}$	Persentase Rumah Tangga yang Menempati Rumah Layak Huni

Berikut merupakan kerangka konseptual yang mengilustrasikan pengaruh antar variabel laten serta indikator dari masing-masing konstruk/laten dalam penelitian ini.



**Gambar 1.** Kerangka Konseptual Penelitian

### 3.2 Tahapan Analisis Data

Metode analisis yang digunakan dalam penelitian ini secara sistematis menggunakan PLS-PM (*Partial Least Square Path Modeling*) serta REBUS-PLS (*Response Based Unit Segmentation in Partial Least Square*) menggunakan *software* XLSTAT 2014. Adapun analisis data dilakukan melalui beberapa tahapan sebagai berikut.

1. Melakukan estimasi parameter dengan PLS-PM yang meliputi:
  - a. Konseptualisasi model dengan merancang *outer model* dan *inner model*.
  - b. Mengkonstruksi diagram jalur (*path diagram*).
  - c. Mengkonversi diagram jalur tersebut ke bentuk model persamaan matematis.
  - d. Mengestimasi parameter model yaitu estimasi bobot (*weight estimate*), estimasi jalur (*path estimate*), serta estimasi untuk mendapatkan *loading (outer model)*.
  - e. Mengevaluasi *outer model* hingga indikator telah valid dan reliabel yang kemudian dilanjutkan dengan evaluasi *inner model*.
  - f. Kembali mengkonstruksi diagram jalur jika melakukan perbaikan *outer model*nya.
  - g. Melakukan pengujian hipotesis dengan *resampling bootstrap standar error*.
  - h. Menginterpretasi hasil analisis PLS-PM
2. Melakukan penerapan metode REBUS-PLS apabila telah memenuhi kriteria secara statistik dengan langkah-langkah sebagai berikut.
  - a. Mengestimasi model PLS.
  - b. Menghitung *communality* dan residual struktural dari tiap unit model global.
  - c. Membentuk *hierarchical cluster analysis* berdasarkan *communality* dan *residual structural*.
  - d. Memilih jumlah kelas/segmen ( $K$ ) berdasarkan dendogram yang diperoleh.
  - e. Mengelompokkan setiap *case* ke dalam kelas/segmen menurut analisis *cluster*.
  - f. Mengestimasi model pada setiap kelas/segmen (model lokal)
  - g. Menghitung nilai CM (*Closeness Measure*) untuk setiap *case* pada model lokal.
  - h. Mengelompokkan setiap *case* pada model lokal. Setelah proses stabilitas komposisi kelas/segmen tercapai maka dilanjutkan ke langkah berikutnya.
  - i. Mendeskripsikan kelas/segmen menurut perbedaan diantara model lokal.
  - j. Melakukan evaluasi dan interpretasi output berdasarkan hasil REBUS-PLS.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Evaluasi Model Pengukuran (*Outer Model*)

Hasil evaluasi *item reliability* dengan nilai *loading factor* untuk tiap indikator reflektif pada model dapat disimpulkan bahwa dalam mengukur konstruk Ekonomi Rumah Tangga, nilai *loading factor* untuk  $X_1$  dan  $X_3$  bernilai  $\leq 0,5$  sehingga indikator tersebut dikatakan tidak valid. Selain itu, dalam mengukur konstruk Fasilitas Perumahan, nilai *loading factor* untuk  $Y_4$  dan  $Y_5$  bernilai  $\leq 0,5$  sehingga indikator tersebut dikatakan tidak valid dan perlu dihapus. Selanjutnya, kerangka konseptual dalam bentuk diagram jalur dikonstruksi kembali karena adanya penghapusan indikator yang tidak memenuhi kriteria tersebut, yaitu  $X_1, X_3, Y_4, Y_5$ . Setelah estimasi parameter, dilakukan tahap evaluasi model.

### 4.2 Evaluasi Model Pengukuran (*Outer Model*) II

- a. *Item Reliability*

**Tabel 4.** Evaluasi Validitas Konvergen dengan Nilai *Standardized Loadings* II

<i>Latent variable</i>	<i>Manifest variables</i>	<b>Koefisien Parameter</b>	<i>Standardized loadings</i>	<i>Critical ratio (CR)</i>	<b>Kesimpulan</b>
Ekonomi	$X_2$	$\lambda_{x2}$	0,835	25,375	Valid
Rumah Tangga	$X_4$	$\lambda_{x4}$	0,764	13,788	Valid

	$X_5$	$\lambda_{x5}$	0,837	17,690	Valid
	$X_6$	$\lambda_{x6}$	0,899	37,197	Valid
Pendidikan	$Y_1$	$\lambda_{y1}$	0,827	12,180	Valid
	$Y_2$	$\lambda_{y2}$	0,979	245,420	Valid
	$Y_3$	$\lambda_{y3}$	0,961	100,165	Valid
Fasilitas	$Y_6$	$\lambda_{y6}$	0,891	13,013	Valid
Perumahan	$Y_7$	$\lambda_{y7}$	0,956	45,877	Valid
Kesehatan	$Y_8$	$\lambda_{y8}$	0,532	2,546	Valid
Lingkungan	$Y_9$	$\lambda_{y9}$	0,764	11,358	Valid
Perumahan	$Y_{10}$	$\lambda_{y10}$	0,824	22,396	Valid

Dari Tabel 4 disimpulkan bahwa dalam mengukur konstruk nilai *loading factor* untuk tiap indikatornya bernilai  $\geq 0,5$  dan memiliki nilai  $|CR| > 2$  sehingga semua indikator tersebut dikatakan valid.

#### b. *Composite Reliability*

Berdasarkan nilai *composite reliability*, untuk variabel laten Ekonomi Rumah Tangga memiliki nilai *Dillon-Glodstein's (D.G. rho)* sebesar 0,902; untuk variabel laten Pendidikan sebesar 0,946; untuk variabel laten Fasilitas Perumahan sebesar 0,921; dan untuk variabel laten Kesehatan Lingkungan Perumahan sebesar 0,756. Hasil tersebut menunjukkan nilai *composite reliability* atau *D.G. rho*  $\geq 0,7$  sehingga dapat disimpulkan bahwa konstruk memiliki reliabilitas atau keandalan yang tinggi sebagai alat ukur.

#### c. *Average Variance Extracted*

Berdasarkan *Average Variance Extracted* tiap variabel laten pada model II, untuk variabel laten Ekonomi Rumah Tangga memiliki nilai AVE sebesar 0,698; untuk variabel laten Pendidikan sebesar 0,855; untuk variabel laten Fasilitas Perumahan sebesar 0,854; dan untuk variabel laten Kesehatan Lingkungan Perumahan sebesar 0,515. Hasil tersebut menunjukkan bahwa nilai AVE untuk tiap variabel konstruk memiliki nilai yang lebih dari 0,5. Hal ini menunjukkan bahwa konstruk memiliki sifat *convergent validity* yang baik.

### 4.3 Evaluasi Model Struktural (*Inner Model*) II

Kebajikan model struktural dapat diukur dengan melihat nilai koefisien determinasi  $R^2$  yang menunjukkan besarnya *variability* variabel endogen yang mampu dijelaskan oleh variabel eksogen. Nilai  $R^2$  untuk variabel Pendidikan ( $\eta_1$ ) adalah 0,748 yang menunjukkan bahwa antar variabel menggambarkan hubungan yang substansial, nilai  $R^2$  untuk variabel Fasilitas Perumahan ( $\eta_2$ ) adalah 0,166 yang menunjukkan hubungan yang lemah, sedangkan nilai  $R^2$  untuk variabel Kesehatan Lingkungan Perumahan ( $\eta_3$ ) adalah 0,774 yang menunjukkan hubungan yang substansial.

Evaluasi berikutnya yaitu *goodness of fit* (GoF) untuk memvalidasi model secara keseluruhan. Berdasarkan hasil pengolahan diperoleh nilai GoF sebesar 0,635 yang berarti masuk dalam kategori GoF besar sehingga dapat diartikan bahwa model yang dihasilkan memiliki kemampuan yang sangat baik dalam menjelaskan data empiris.

### 4.4 Pengujian Hipotesis

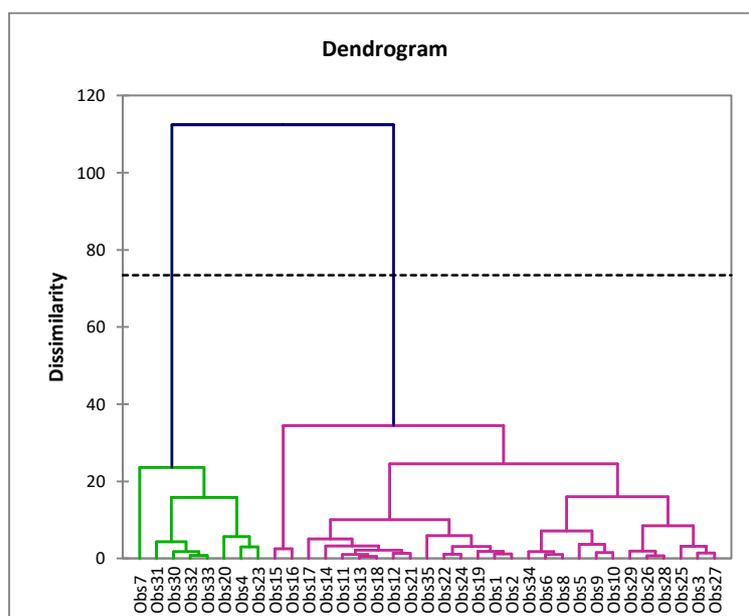
Dengan taraf signifikansi  $\alpha = 5\%$  dan kriteria uji :  $H_0$  ditolak jika  $|T - statistics| > 1,96$  atau  $p - value < 0,05$ , maka pengujian untuk masing-masing jalur antar variabel laten dapat dilihat pada Tabel 5 berikut.

**Tabel 5.** Uji Koefisien Jalur *Inner Model*

Hubungan Antar Variabel	Koefisien Parameter	Estimasi Nilai Koefisien Parameter	T-hitung statistik	P-value	Keputusan
ERT → PDK	$\gamma_{11}$	0,865	9,907	0,000	$H_o$ ditolak (ada pengaruh yang signifikan)
ERT → FP	$\gamma_{21}$	0,407	2,559	0,015	$H_o$ ditolak (ada pengaruh yang signifikan)
ERT → KLP	$\gamma_{31}$	0,259	1,519	0,139	$H_o$ diterima (tidak ada pengaruh yang signifikan)
PDK → KLP	$\beta_{31}$	0,214	1,221	0,231	$H_o$ diterima (tidak ada pengaruh yang signifikan)
FP → KLP	$\beta_{32}$	0,574	5,952	0,000	$H_o$ ditolak (ada pengaruh yang signifikan)

#### 4.5 Clustering pada REBUS-PLS

Pengelompokan unit observasi dalam REBUS-PLS berdasarkan hasil *hierarchical cluster analysis* menggunakan metode *Ward's* terhadap residual komunal dan residual struktural. Hasil dari analisis *cluster* terhadap residual komunal dan residual struktural adalah sebagai berikut:



**Gambar 2.** Dendrogram Hasil *Hierarchical Cluster Analysis*

Berdasarkan dendrogram yang telah dihasilkan, hanya dimungkinkan untuk mengelompokkan unit observasi menjadi dua kelas/segmen. Setiap unit observasi dimasukkan ke dalam kelas yang menunjukkan nilai CM yang lebih kecil. Komposisi kabupaten/kota dari 2 segmen hasil analisis *cluster* yaitu pada kelas/segmen 1 yang mana terdapat 23 kabupaten/kota terdiri dari Kabupaten Cilacap, Kabupaten Purbalingga, Kabupaten Kebumen, Kabupaten Magelang, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Wonogiri, Kabupaten Karanganyar, Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Blora, Kabupaten Rembang, Kabupaten Pati, Kabupaten Kudus, Kabupaten Demak, Kabupaten Kendal, Kabupaten Pekalongan, Kabupaten Pemalang,

Kabupaten Tegal, Kabupaten Brebes, Kota Pekalongan, dan Kota Tegal. Sedangkan pada kelas/segmen 2 yang mana terdapat 12 kabupaten/kota terdiri dari Kabupaten Banyumas, Kabupaten Banjarnegara, Kabupaten Purworejo, Kabupaten Wonosobo, Kabupaten Jepara, Kabupaten Semarang, Kabupaten Temanggung, Kabupaten Batang, Kota Magelang, Kota Surakarta, Kota Salatiga, dan Kota Semarang.

Selanjutnya, jika kabupaten/kota telah dikelompokkan berdasarkan segmen yang terbentuk, masing-masing segmen tersebut kemudian Mengestimasi model pada setiap kelas/segmen (model lokal) dengan menggunakan PLS biasa.

#### 4.6 Interpretasi Hasil REBUS-PLS

##### 4.6.1 Heterogenitas pada Persamaan Struktural

**Tabel 6.** Perbandingan Nilai Koefisien Jalur Model Global dan 2 Model Lokal

Hubungan Antar Variabel	Global Model		Local Model 1		Local Model 2	
	Koefisien Parameter	p-value	Koefisien Parameter	p-value	Koefisien Parameter	p-value
ERT → PDK	0,865*	0,000	0,790*	0,000	0,948*	0,000
ERT → FP	0,407*	0,015	0,728*	0,000	0,669*	0,017
ERT → KLP	0,259	0,139	0,229	0,227	-0,072	0,872
PDK → KLP	0,214	0,231	0,195	0,268	0,719	0,198
FP → KLP	0,574*	0,000	0,557*	0,002	0,327	0,176

Keterangan: tanda \* berarti signifikan pada taraf  $\alpha = 5\%$

Berdasarkan Tabel 6 diperoleh kesimpulan bahwa Ekonomi Rumah Tangga berpengaruh positif dan signifikan terhadap Pendidikan baik untuk model global, model lokal 1, maupun untuk model lokal 2. Ekonomi Rumah Tangga berpengaruh positif dan signifikan terhadap Fasilitas Perumahan baik untuk model global, model lokal 1, maupun untuk model lokal 2. Sedangkan pada model global dan model lokal 1 Fasilitas Perumahan berpengaruh positif dan signifikan terhadap kesehatan lingkungan perumahan, namun untuk model lokal 2 tidak berpengaruh signifikan.

##### 4.6.2 Heterogenitas pada Persamaan Pengukuran

**Tabel 7.** Perbandingan *Standardized Loading Factor* Model Global dan 2 Model Lokal

Latent variable	Manifest variables	Global Model		Local Model 1		Local Model 2	
		$\lambda$	Kesimpulan	$\lambda$	Kesimpulan	$\lambda$	Kesimpulan
Ekonomi Rumah Tangga	$X_2$	0,835	Valid	0,895	Valid	0,922	Valid
	$X_4$	0,764	Valid	0,876	Valid	0,838	Valid
	$X_5$	0,837	Valid	0,728	Valid	0,979	Valid
	$X_6$	0,899	Valid	0,865	Valid	0,959	Valid
Pendidikan	$Y_1$	0,827	Valid	0,691	Valid	0,957	Valid
	$Y_2$	0,979	Valid	0,979	Valid	0,987	Valid
	$Y_3$	0,961	Valid	0,951	Valid	0,989	Valid
Fasilitas Perumahan	$Y_6$	0,891	Valid	0,840	Valid	0,957	Valid
	$Y_7$	0,956	Valid	0,885	Valid	0,954	Valid
Kesehatan Lingkungan Perumahan	$Y_8$	0,532	Valid	0,261	Tidak Valid	0,782	Valid
	$Y_9$	0,764	Valid	0,865	Valid	0,865	Valid
	$Y_{10}$	0,824	Valid	0,910	Valid	0,866	Valid

Berdasarkan Tabel 7 disimpulkan bahwa untuk model lokal atau segmen 1, hampir seluruh nilai *loading factor* memiliki nilai  $\geq 0,5$ ; hanya ada satu variabel *manifest* yang

memiliki nilai *loading factor*  $\leq 0,5$  yaitu  $Y_8$  (Persentase Rumah Tangga yang Memiliki Akses terhadap Air Minum Layak). Sedangkan pada model lokal atau segmen 2, seluruh nilai *loading factor* memiliki nilai  $\geq 0,5$  sehingga seluruh indikator dikatakan memiliki validitas yang baik.

#### 4.6.3 Evaluasi Model

**Tabel 8.** Perbandingan Nilai  $R^2$  Model Global dan 2 Model Lokal

Variabel Laten Endogen	$R^2$		
	Global Model	Local Model	Local Model 2
Pendidikan ( $\eta_1$ )	0,748	0,625	0,898
Fasilitas Perumahan ( $\eta_2$ )	0,166	0,531	0,448
Kesehatan Lingkungan Perumahan ( $\eta_3$ )	0,774	0,804	0,868

Dengan demikian, dapat disimpulkan bahwa nilai  $R^2$  pada masing-masing segmen yang terbentuk (model lokal) menjadi lebih besar untuk tiap laten endogennya (kecuali nilai  $R^2$  untuk Pendidikan pada model lokal 2) daripada nilai  $R^2$  pada model global.

Selain nilai *R square*, evaluasi berikutnya yaitu *goodness of fit* (GoF) untuk memvalidasi model secara keseluruhan dan untuk mendapatkan model lokal yang lebih baik daripada model global. Untuk model global diperoleh nilai *Goodness of Fit* sebesar 0,635; untuk model lokal 1 diperoleh nilai *Goodness of Fit* sebesar 0,673; dan untuk model lokal 2 diperoleh nilai *Goodness of Fit* sebesar 0,794. Dengan demikian, dapat disimpulkan bahwa nilai GoF baik untuk model global, model lokal 1, dan model lokal 2 masuk dalam kategori GoF besar. Nilai GoF yang tinggi juga menunjukkan model struktural dan pengukuran yang lebih baik. Selain itu, nilai GoF untuk masing-masing model lokal menjadi lebih baik dibandingkan nilai GoF untuk model global yang mengindikasikan kebaikan model pada model lokal yang lebih baik sehingga heterogenitas dalam penelitian ini dapat dideteksi dengan baik.

## 5. KESIMPULAN

Kesimpulan yang diperoleh dari hasil dan pembahasan pada penelitian ini yaitu pada model hubungan untuk variabel laten ekonomi rumah tangga, pendidikan, fasilitas perumahan, dan kesehatan lingkungan perumahan, terdapat 12 indikator yang signifikan dalam struktur model. Pada model persamaan struktural terdapat jalur yang signifikan yaitu pengaruh ekonomi rumah tangga terhadap pendidikan, pengaruh ekonomi rumah tangga terhadap fasilitas perumahan, dan pengaruh fasilitas perumahan terhadap kesehatan lingkungan perumahan. Berdasarkan dendrogram yang telah dihasilkan, ditetapkan untuk mengelompokkan unit observasi menjadi dua kelas/segmen. Berdasarkan nilai CM, dapat diketahui bahwa ketika dibentuk 2 kelas/segmen menunjukkan komposisi kelas/segmen 1 terdiri dari 23 kabupaten/kota, sedangkan untuk kelas/segmen 2 terdiri dari 12 kabupaten/kota. REBUS-PLS mampu mendeteksi heterogenitas pada model SEM-PLS dengan nilai  $R^2$  pada masing-masing segmen yang terbentuk (model lokal) menjadi lebih besar untuk tiap laten endogennya (kecuali nilai  $R^2$  untuk Pendidikan pada model lokal 2) daripada nilai  $R^2$  pada model global, yang mengindikasikan kebaikan model pada model lokal lebih baik dari model global. Sedangkan untuk nilai GoF baik untuk model global, model lokal 1, dan model lokal 2 masuk dalam kategori GoF besar yang menunjukkan model struktural dan pengukuran yang lebih baik. Selain itu, nilai GoF untuk masing-masing model

lokal menjadi lebih baik dibandingkan nilai GoF untuk model global sehingga heterogenitas dalam penelitian ini dapat dideteksi dengan baik.

#### **DAFTAR PUSTAKA**

- Afifah I.N., Sunaryo S. 2013. “Analisis Structural Equation Modelling (SEM) dengan Finite Mixture Partial Least Square (FIMIX-PLS) (Studi Kasus: Struktur Model Kemiskinan di Provinsi Jawa Tengah Tahun 2011).” *Yogyakarta : Seminar Nasional Matematika dan Pendidikan Matematika FMIPA UNY*. ISBN : 978-979-16353-9-4
- [BPS] Badan Pusat Statistik. 2019. *Profil Tempat Tinggal Provinsi Jawa Tengah 2018*. Semarang : Badan Pusat Statistik Provinsi Jawa Tengah.
- Mayasari, M., Ritohardoyo, S. 2012. “Kualitas Pemukiman di Kecamatan Pasarkliwon Kota Surakarta”. *Jurnal Bumi Indonesia Vol. 1, No. 3*
- Sanchez, G. 2013. *PLS Path Modeling with R*. Tersedia : [https://www.gastonsanchez.com/PLS\\_Path\\_Modeling\\_with\\_R.pdf](https://www.gastonsanchez.com/PLS_Path_Modeling_with_R.pdf) (diakses pada tanggal 28 April 2020)
- Trincherà, L. 2007. “*Unobserved Heterogeneity in Structural Equation Models: a new approach to latent class detection in PLS Path Modeling*”. *Doctoral Thesis, Universita degli Studi di Napoli Federico II*
- Yamin, S., Kurniawan H. 2011. *Generasi Baru Mengolah Data Penelitian dengan Partial Least Square Path Modeling : Aplikasi dengan Software XLSTAT, SmartPLS, dan Visual PLS*. Jakarta : Salemba Infotek